

Analyzing Data Efficiency and Performance of Machine Learning Algorithms for Assessing Low Back Pain Physical Rehabilitation Exercises

Aleksa Marusic¹, Louis Annabi¹, Sao Mai Nguyen^{1,2} and Adriana Tapus¹

Abstract—Physical rehabilitation focuses on the improvement of body functions, usually after injury or surgery. Patients undergoing rehabilitation often need to perform exercises at home without the presence of a physiotherapist. Computer-aided assessment of physical rehabilitation can improve patients' performance and help in completing prescribed rehabilitation exercises. In this work, we focus on human motion analysis in the context of physical rehabilitation for Low Back Pain (LBP). As 2D and 3D human pose estimation from RGB images had made impressive improvements, we aim to compare the assessment of physical rehabilitation exercises using movement data acquired from RGB videos and human pose estimation from those. In this work, we provide an analysis of two types of algorithms on a Low Back Pain rehabilitation datasets. One is based on a Gaussian Mixture Model (GMM), with performance metrics based on the log-Likelihood values from GMM. Furthermore, with the recent development of Deep Learning and Graph Neural Networks, algorithms based on Spatio-Temporal Graph Convolutional Networks (STGCN) are taken as a novel approach. We compared the algorithms in terms of data efficiency and performance, with evaluation performed on two LBP rehabilitation datasets: KIMORE and Keraal. Our study confirms that Kinect, OpenPose, and BlazePose data yield similar evaluation scores, and shows that STGCN outperforms GMM in most configurations.

I. INTRODUCTION

Physical rehabilitation has a very important role in post-operative recovery and in the restoration of body functions [3]. Usually, during the rehabilitation process, patients performing exercises are monitored in a clinical setting by a medical professional, such as a physiotherapist. During a rehabilitation exercise session, patients' behavior reflects their health status and is an important indicator of the treatment outcome. However, patients often have a limited number of supervised sessions, and they need to continue the rehabilitation process at home without any supervision. In these cases, a physiotherapist makes a rehabilitation plan consisting of several recommended exercises. Patients are typically responsible for performing their exercises regularly at home and periodically visiting the hospital for progress assessment. However, a lack of supervision and timely feedback from healthcare professionals can reduce patient's engagement during the rehabilitation process. Lower motivation

and poor supervision can increase the chances of incorrect exercise performance, which can slow down the recovery process and increase the risk of re-injury [22].

Low back pain (LBP) is a major cause of disability worldwide, with more than 50% of the global population experiencing LBP at some point in their lives [3]. This is especially concerning as LBP disproportionately affects elderly individuals, whose percentage in European societies is steadily increasing. As a result, medical staff are under significant strain to manage the growing number of patients suffering from LBP.

Automatic physical rehabilitation monitoring can significantly improve patients' progress during at-home rehabilitation. The goal of such a system is to recognize the activity being performed, the intensity with which it is performed, and its quality, thus helping monitor patients' progress. In general, human activity analysis is a very active research topic today and one of the most important and challenging areas in AI. It involves analyzing human body movements based on the motions of different body joints, skeletons, and muscles [30]. It also has applications in several domains such as sports sciences, action or gesture recognition [10], [17], [9], [2], and range-of-motion estimation [1].

Developing an effective system for movement assessment highly depends on a few factors including motion sensors, precise movement data and its pre-processing, and evaluation techniques. In recent years, there were several studies that employed machine learning methods to classify individual repetitions into correct or incorrect classes of movements. Some of the first methods proposed for this task included distance function-based algorithms such as Dynamic Time Warping and Mahalanobis distance or probabilistic models such as hidden Markov models and Gaussian mixture models [32], [7]. The outputs in these approaches are discrete class values of 0 or 1 (i.e., incorrect or correct).

Naturally, with recent developments in Neural Networks and Deep Learning (DL), there is a big interest in their application for modeling and analysis of human motions. There are already numerous papers on general Human Action Recognition (HAR) systems that utilize various DL frameworks ranging from Convolutional Networks and Long Short-Term Memory (LSTM) [15] and encoder-decoder networks to even more novel architectures such as Spatio-Temporal Graphs [35], and Attention models [23].

Furthermore, a large number of datasets related to HAR fields are freely available for analysis. These datasets are

*This work was supported by ENSTA Paris

¹Autonomous Systems and Robotics Lab, Computer Science and System Engineering (U2IS), ENSTA Paris, Institut Polytechnique de Paris, 828 Blvd des Maréchaux, 91120 Palaiseau, France, name.surname@ensta-paris.fr

²Dep. Informatique, IMT Atlantique, nguyensmai@gmail.com
979-8-3503-0704-7/23/\$31.00 ©2023 European Union

extensively used for benchmarking algorithms for action recognition, gesture recognition, or pose estimation. However, in the medical domain, collecting large data sets of rehabilitation exercise data from patients faces multiple challenges such as impairment, unlabeled data, or privacy and safety concerns. Consequently, only a few public datasets for rehabilitation evaluation are currently available, and they are still quite smaller than more general ones, e.g. the ones used for benchmarking HAR algorithms.

Our study focuses on evaluating the performance and data efficiency of two distinct algorithms used to assess the effectiveness of rehabilitation exercises. Specifically, we investigate the Gaussian Mixture Model and the Spatio-Temporal Graph Convolutional Network algorithm. Our evaluation is conducted on two datasets that contain rehabilitation exercises for Low Back Pain patients, namely Kimore [8] and Keraal dataset [28].

The rest of the paper is structured as follows. Section II presents an overview of the different approaches for motion analysis and physical rehabilitation assessment. Section III describes the used datasets, while Section IV details the implemented methods for rehabilitation assessment. The results are summarized in Section V. Finally, the conclusions and discussions are presented in Section VI.

II. RELATED WORK

This section presents related work in deep learning for motion analysis in general and in the assessment of physical rehabilitation exercises.

A. Deep learning for motion analysis

Several deep learning approaches have been applied on skeleton data, in particular on the task of action recognition. Motion data can be seen as 3D tensors with one temporal dimension (the timeframe of the movement), one spatial dimension (the skeleton joint), and one feature dimension (the XYZ Euclidean position). Early deep learning approaches for motion processing focused on the temporal processing, using recurrent neural networks [12], or 1D convolutional neural networks [19]. Other approaches explored the idea of representing motion as images, in order to exploit the performances of 2D and 3D convolutional neural network for image processing [34], [18], [20], [6], [14].

With the recent development of graph neural networks, there is a new wave of algorithms that are able to properly take into account the skeleton structure using graph convolutions [35], [31], [24]. These neural network layers perform an operation that can be seen as a message passing between adjacent joints in the skeleton graph, thus properly exploiting this prior knowledge. More recently, self-attention mechanisms have been added to allow graph convolutions to span across non-adjacent joints based on dynamically computed attention coefficients [29], [27].

B. Assessment of physical rehabilitation exercises

Movement assessment is typically accomplished by comparing a patient's performance of an exercise to the desired

performance as specified by therapists. A sequence of body movements is provided as input for a machine or deep learning algorithm, which should assess that exercise with a quality score. This thus requires a more precise model of the movement than most gesture classification models.

At first, studies on exercise evaluation employed more traditional machine learning methods for classification, such as Adaboost classifier, K-Nearest Neighbors, Bayesian classifier, or ANNs [4]. Others tried using distance function based models like [16]. However, classifiers only provide correct or incorrect labels, not providing any additional information or score, while distance functions solve that problem but are not able to learn from the rehabilitation data.

Further, some of the research tried using probabilistic approaches, like Hidden Markov models [7] or Gaussian Mixture Model [25]. Such models provide an assessment that is based on the likelihood that the given exercises are being drawn from a trained model. These models were able to solve previous problems, and stochastic character of human movements goes hand in hand with models nature, but they are not able to extract all the information from the data, such as joint or spatial connections among body parts.

Liao et al. [21] created a deep neural network model to generate quality scores of input movements. They proposed deep learning architecture for hierarchical spatio-temporal modeling combining GMMs, CNNs, and LSTM to provide a quality score. However, with recent development of Graph Neural Networks, it is possible to extract even more information from spatio-temporal features of the exercise. In [11] and [13], Graph Convolutional Networks (GCN) are used to assess physical rehabilitation, obtaining state-of-the-art scores on commonly used KIMORE and UI-PRMD datasets. Last but not least, [37] et al. used an ensemble of two GCN, one for position and for orientation features of the skeleton joints.

III. DATASET

This section explains the type of data used and presents two datasets used for the evaluation of the algorithms.

A. Skeleton data

A Human Pose Skeleton represents the orientation of a person in a graphical format. Depth cameras, like the Microsoft Kinect, can provide position and orientation of skeleton joints. They had become very popular due to their price and ease of use over optical motion tracking systems, which place a set of markers on the body. More recently, a standard vision camera can be used with deep learning techniques that estimate skeleton joints positions from plain RGB images. In this work, we will consider the algorithms OpenPose and BlazePose [38], [33]. Figure 1 displays the joints of Kinect ¹, OpenPose ² and BlazePose ³ skeletons.

¹<https://www.sealeftstudios.com/blog/blog20160708.php>

²<https://maelfabien.github.io/tutorials/open-pose/>

³<https://ai.googleblog.com/2020/08/on-device-real-time-body-pose-tracking.html>

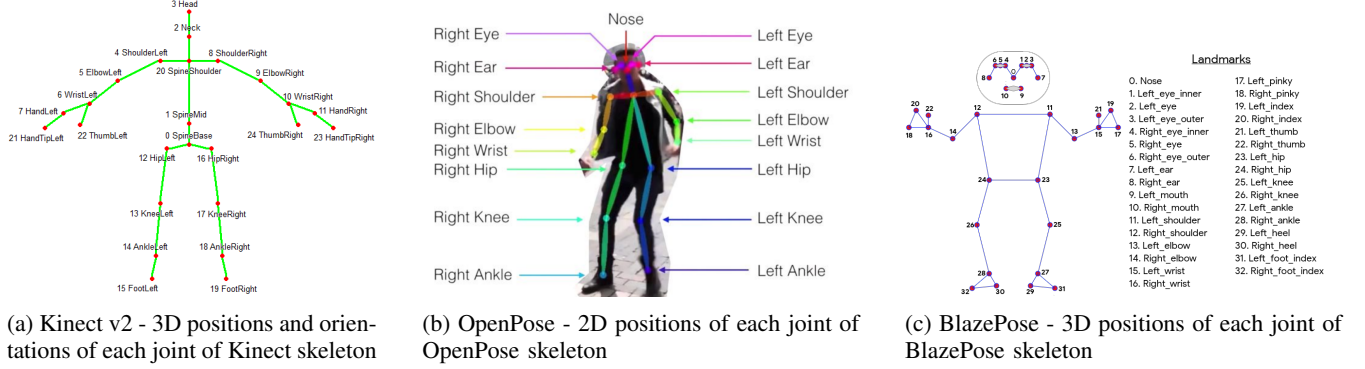


Fig. 1: Skeleton format for the three pose estimation algorithms used in the Keraal dataset. For the Kimore dataset, we only have Kinect data.

B. Keraal dataset

The Keraal dataset is a medical database of clinical patients carrying out low back-pain rehabilitation exercises. The data includes recordings from healthy subjects but, more importantly, of rehabilitation patients, extracted from a 4 weeks evolution of each patient. The centrally randomized, controlled, single-blind, and bi-centric study was conducted from October 2017 to May 2019. The rehabilitation program includes a group of 31 patients, aged 18 to 70 years, recruited in the double-blind study. 12 patients suffering from low-back pain were included in the Robot Supervised Rehabilitation Group, and were asked by a humanoid robot coach to perform each of the three predefined exercises the best they can from its demonstration. The details of this clinical trial, including the patient care, the rehabilitation sessions, the robot coach, the inclusion and exclusion criteria, the characteristics of the patients, and the efficiency of the care have been reported in [5]. Details can be read on <http://nguyensmai.free.fr/KeraalDataset.html>. A list of three exercises has been chosen in conjunction with therapists as common rehabilitation exercises that are also used for low-back pain treatment.

Videos collected from patients and healthy subjects were annotated by two physiotherapists. In this study, the labels are obtained by merging the assessments of two physicians, and we process the videos to obtain the BlazePose and OpenPose skeletons. Each exercise was labeled as either correct or incorrect. The dataset used in this study comprises Kinect (3D) v2 skeleton data, Blazepose (3D) and OpenPose (2D) skeletons obtained from videos and annotations.

C. Kimore dataset

The Kimore dataset [8] includes RGB-D videos and score annotations of five exercises for LBP rehabilitation, selected by physicians. The exercises are performed by two groups of participants: a control group (44 participants) and a group of patients (34 participants). The dataset also contains an assessment of the performed exercises, provided by two physicians. More details can be found here <https://vrai.dii.univpm.it/content/kimore-dataset>.

IV. METHODOLOGY

This section provides the technical overview of two algorithms used in this study.

A. Gaussian Mixture Model

Gaussian mixture models (GMMs) belong to a group of probabilistic models used to classify data into different categories based on probability distribution. GMM models the dataset as a mixture of several Gaussian distributions. As in [25], we encode the movement point positions as a Gaussian Mixture Model (GMM): $\theta = [t, x]$, where t is the timestamp and x the joints positions.

$$p(\theta) = \sum_{i=1}^K \phi_i \mathcal{N}(\mu_i, \Sigma_i) \quad (1)$$

where the i^{th} vector component is characterized by normal distributions with weights ϕ_i , means μ_i , and covariance matrices Σ_i . Each Gaussian of the mixture is thus defined by:

$$\mu_i = \begin{bmatrix} \mu_i^t \\ \mu_i^x \end{bmatrix}, \Sigma_i = \begin{bmatrix} \Sigma_i^t & \Sigma_i^{xt} \\ \Sigma_i^{xt} & \Sigma_i^x \end{bmatrix} \quad (2)$$

where the indices t and x refer to respectively time and position.

The parameters ϕ_i, μ_i, Σ_i are learned by Expectation-Maximisation (EM) from the skeleton data of the movements captured by the Kinect or estimated with OpenPose or BlazePose.

B. Graph Convolutional Networks

Graph Neural Networks (GNNs) are a class of deep learning models that are specifically designed to operate on graph-structured data [13]. These models leverage the graph topology to learn meaningful representations of the nodes and edges of the graph.

Given our focus on skeletons, let us examine how they can be integrated into graph data. Skeleton-based data can be obtained from motion-capture devices or pose estimation algorithms from videos. Usually, the data is a sequence of frames, each frame will have a set of joint coordinates. Each joint in the given skeleton can be represented as a node in

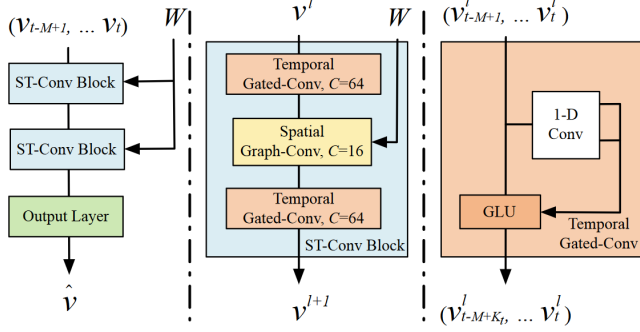


Fig. 2: Architecture of spatio-temporal graph convolutional networks. The network chains two spatio-temporal convolutional blocks (ST-Conv blocks) and a fully-connected output layer. Each ST-Conv block contains two temporal gated convolution layers and one spatial graph convolution layer in the middle. Image taken from [36].

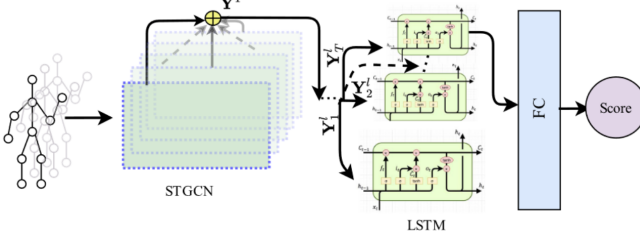


Fig. 3: Architecture of STGCN for physical rehabilitation assessment. The network combines STGCNs with LSTM layers as suggested in [11]. Image taken from [11].

a graph, while connections between the joints represent the edges in the graph (e.g., right hip to right knee). In this way, the graph provides information about the hierarchy of the human skeleton, starting from one joint as a root (e.g., mSpine) and expanding further to hands and feet, which would be the leaves of the graph. Although GNNs have been extensively used in various domains, they were first used on static graph data, where graph structure does not change once data is fitted. In recent years, there has been an increased interest in systems with temporal dimension, meaning graph data would change over time. To address this need, a new family of GNNs has emerged: Spatio-Temporal GNNs, which take into account both the spatial and temporal dimensions of the data by learning temporal representations of the graph structure.

Such architecture was first introduced in [36]. The authors proposed the architecture of spatio-temporal graph convolutional networks (STGCN). As shown in Figure 6, STGCN is composed of several spatio-temporal convolutional blocks, each of which is formed as a “sandwich” structure with two gated sequential convolution layers and one spatial graph convolution layer in between.

Further, in [35] this model was applied to skeleton-based action recognition, while [11] modified that algorithm for the task of assessing physical rehabilitation exercises. Since it is

a very novel approach obtaining state-of-the-art results, we decided to their algorithm as the base for our analysis here. An overview of this model can be seen in Figure 3.

V. RESULTS

We experimented with the STGCN and GMM algorithms and compared their performances and sample efficiency on the two physical rehabilitation exercises datasets. The GMM is trained on correct demonstrations of the exercises, and then a classification threshold is determined based on validation data containing both correct and incorrect demonstrations. In contrast, the STGCN method needs to be trained on both correct and incorrect demonstrations. Even though validation data could be used to optimize hyperparameters or to perform early stopping, we did not use it with this method. To compare the data efficiency of both algorithms, we thus need to take into account both the number of training and validation examples needed for the GMM method and compare it with the number of training examples needed for the STGCN based method. We report the scores after training, which takes a couple of seconds for GMM to train, while STGCN, for one training of 250 epochs, takes 20 - 70 minutes depending on the setup (which dataset, skeleton type, and number of training examples). All models have been trained on CPU Intel Core i9-9900KF.

Figure 4 shows the F1 scores obtained with two methods on the Keraal dataset, while 5 provides the F1 scores. The scores are averaged across the 3 exercises of the dataset, with Kinect, OpenPose, and BlazePose poses. We can notice a slight but not significant improvement in these scores as the training set size increases. For both GMM and STGCN, we note that the scores with Kinect, Openpose and BlazePose are similar : the use of depth sensors (Kinect v2) does not seem to improve significantly the performance of the algorithm, which corroborates the conclusions presented in [26]: for low-back rehabilitation exercises, previous GMM obtained through Kinect, OpenPose and BlazePose data revealed comparable results. These new results extend the same conclusion to variations in the sizes of the training and validation sets, and to another evaluation algorithm : STGCN. This indicates that independently of the size of the dataset and the machine learning algorithm, simple RGB cameras have the potential to be used as the main sensor for collecting movement data. On Kinect and BlazePose data, the STGCN method seems to outperform the GMM method, especially when a large number of training examples are available. These results advocate in favor of using the STGCN method, even when few training examples are available.

Results for the Kimore dataset are presented in Figure 6. While GMM can achieve an F1 score of nearly 0.9 with 100 training examples, we can see that more than 200 examples are needed to achieve such classification precision with STGCN. Even when taking into consideration the additional validation examples needed for GMM, this model seems to perform significantly better when a small number of examples is available. This result contradicts what was observed for the Keraal dataset, where the STGCN method

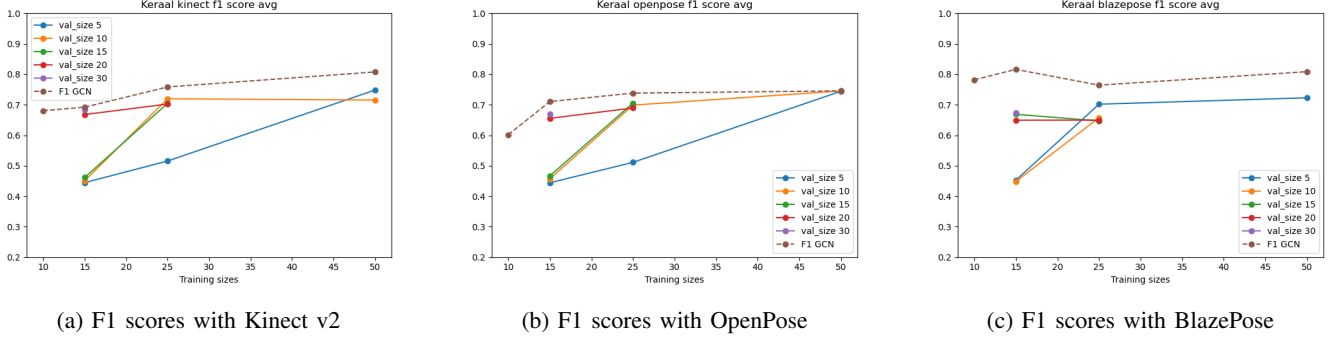


Fig. 4: F1 scores of GCN and GMM on the Keraal dataset, with different training sizes used. The data is averaged across exercises or groups. For GMM various sizes of validation sizes (for threshold defining) are deployed.

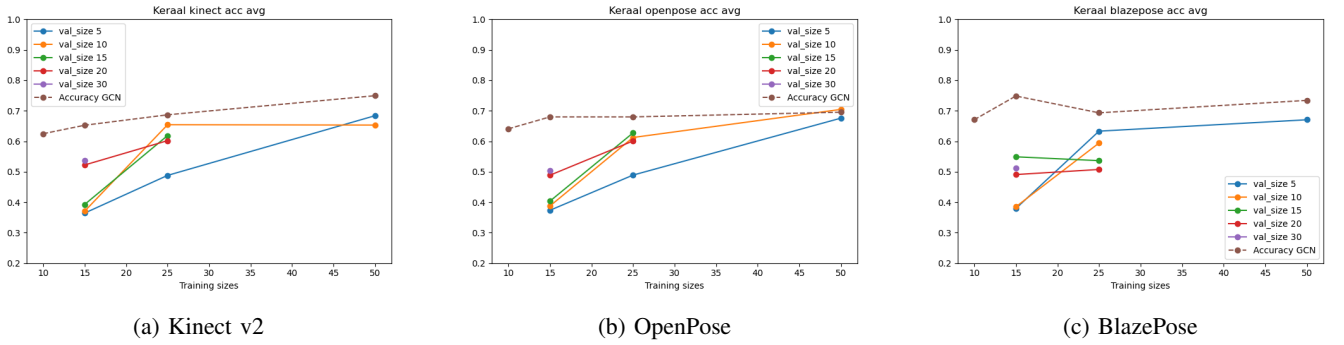


Fig. 5: Accuracy of STGCN and GMM on the Keraal dataset.

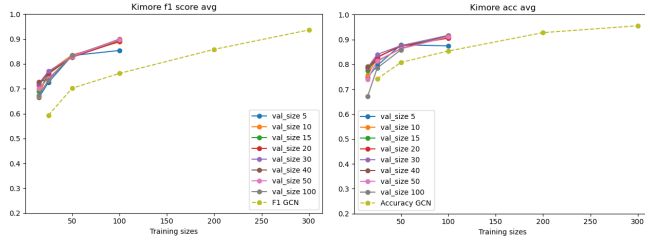


Fig. 6: F1 score and accuracy scores of STGCN and GMM on the Kimore dataset.

outperforms the GMM method even with a few training examples. Our supposition is that this could be due to the absence of a strong agreement between the physicians for Keraal (Cohen's $\kappa = 0.63$ and Krippendorff's $\alpha = 0.62$). In comparison, the Kimore dataset uses a questionnaire containing 10 questions to assess the quality of the performed exercises, which could lead to more robust labeling.

VI. CONCLUSIONS

In this work, we have compared two algorithms for LBP physical rehabilitation assessment on two datasets with several human pose estimation methods. We can draw several conclusions from the observed results, that can provide useful insights in order to further develop the use of automated physical rehabilitation methods :

- While more experiments could be done in order to confirm this result, we observed that using more expensive

depth cameras does not seem to impact the performance of the assessment method. This study confirms the conclusion presented in [26] over a more extensive study using more sizes of training and validation sets and, additionally, using a more efficient evaluation algorithm. Similarly to that evaluation, we see that the use of 3D inputs (Kinect and BlazePose), compared to 2D inputs (OpenPose) does not improve the results obtained on the Keraal dataset.

- More training examples lead to a better assessment. We recommend collecting data from as many participants as possible when recording exercises.
- Label quality is essential. We observe significantly better accuracy on the Kimore dataset, where the labels were obtained by merging the answers to ten questions given by two physicians, compared to the Keraal dataset, where only two evaluations are combined. Although this reveals that assessing rehabilitation movements is a difficult task, we suggest having as many annotators as possible and monitoring a measure of their agreement to ensure high label quality.
- Finally, we recommend using the STGCN algorithm instead of the GMM algorithm in most situations. The GMM algorithm should still be useful in special cases when we need a fast (real-time) learning system or when gathering incorrectly performed exercises is difficult. It can be trained using only correct demonstrations and only needs a few incorrect demonstrations to optimize the threshold value used for classification.

REFERENCES

- [1] Miron A, Sadawi N, Waidah I, Hussain H, and Grosan C. Intellirehabds (irds)—a dataset of physical rehabilitation movements. *Data*, 6(5), 2021.
- [2] Tapus A., Bandera A., Vazquez-Martin R., and Calderita L. V. Perceiving the person and their interactions with the others for social robotics - a review. *Pattern Recognition Letters*, 2018.
- [3] Wu A, March L, Zheng X, Huang J, Wang X, Zhao J, Blyth FM, Smith E, Buchbinder R, and Hoy D. Global low back pain prevalence and years lived with disability from 1990 to 2017: estimates from the global burden of disease study 2017. *Ann Transl Med.*, 8(6):299, January 2020.
- [4] Ilktan Ar and Yusuf Sinan Akgul. A computerized recognition system for the home-based physiotherapy exercises using an rgbd camera. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 22(6):1160–1171, 2014.
- [5] Agathe Blanchard, Sao Mai Nguyen, Maxime Devanne, Mathieu Simonnet, Myriam Le Goff-Pronost, and Olivier Rémy-Néris. Technical feasibility of supervision of stretching exercises by a humanoid robot coach for chronic low back pain: The r-cool randomized trial. *BioMed Research International*, 2022:1–10, mar 2022.
- [6] Carlos Caetano, François Brémont, and William Robson Schwartz. Skeleton image representation for 3d action recognition based on tree structure and reference joints. In *2019 32nd SIBGRAPI conference on graphics, patterns and images (SIBGRAPI)*, pages 16–23. IEEE, 2019.
- [7] Marianna Capecci, Maria Gabriella Ceravolo, Francesco Ferracuti, Sabrina Iarlori, Ville Kyrki, Andrea Monteriù, Luca Romeo, and Federica Verdini. A hidden semi-markov model based approach for rehabilitation exercise assessment. *Journal of Biomedical Informatics*, 78:1–11, 2018.
- [8] Marianna Capecci, Maria Gabriella Ceravolo, Francesco Ferracuti, Sabrina Iarlori, Andrea Monteriu, Luca Romeo, and Federica Verdini. The kimore dataset: Kinematic assessment of movement and clinical scores for remote monitoring of physical rehabilitation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 27(7):1436–1448, 2019.
- [9] Glowinski D, Dael N, Camurri A, Volpe G, Mortillaro M, and Scherer K. Toward a minimal representation of affective gestures. *IEEE Transactions on Affective Computing*, 2(2):106–118, 2011.
- [10] Kulić D, Venture G, and Nakamura Y. Detecting changes in motion characteristics during sports training. *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4011–4014, 2009.
- [11] Swakshar Deb, Md Fokhrul Islam, Shafin Rahman, and Sejuti Rahman. Graph convolutional networks for assessment of physical rehabilitation exercises. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 30:410–419, 2022.
- [12] Maxime Devanne, Panagiotis Papadakis, et al. Recognition of activities of daily living via hierarchical long-short term memory networks. In *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, pages 3318–3324. IEEE, 2019.
- [13] Chen Du, Sarah A Graham, Colin A. Depp, and Truong Q. Nguyen. Assessing physical rehabilitation exercises using graph convolutional network with self-supervised regularization. *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 281–285, 2021.
- [14] Haodong Duan, Yue Zhao, Kai Chen, Dahua Lin, and Bo Dai. Revisiting skeleton-based action recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2969–2978, June 2022.
- [15] Fuqiang Gu, Mu-Huan Chung, Mark Chignell, Shahrokh Valaee, Baoding Zhou, and Xue Liu. A survey on deep learning for human activity recognition. *ACM Computing Surveys*, 54, 08 2021.
- [16] Roshanak Houmanfar, Michelle Karg, and Dana Kulić. Movement analysis of rehabilitation exercises: Distance metrics for measuring patient progress. *IEEE Systems Journal*, 10:1014–1025, 2016.
- [17] Aggarwal J.K. and Ryoo M.S. Human activity analysis: A review. *ACM Comput. Surv.*, 43(3), apr 2011.
- [18] Qiuhong Ke, Mohammed Bennamoun, Senjian An, Ferdous Sohel, and Farid Boussaid. A new representation of skeleton sequences for 3d action recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3288–3297, 2017.
- [19] Tae Soo Kim and Austin Reiter. Interpretable 3d human action analysis with temporal convolutional networks. In *2017 IEEE conference on computer vision and pattern recognition workshops (CVPRW)*, pages 1623–1631. IEEE, 2017.
- [20] Chuankun Li, Yonghong Hou, Pichao Wang, and Wanqing Li. Joint distance maps based action recognition with convolutional neural networks. *IEEE Signal Processing Letters*, 24(5):624–628, 2017.
- [21] Yalin Liao, Aleksandar Vakanski, and Min Xian. A deep learning framework for assessing physical rehabilitation exercises. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, PP:1–1, 01 2020.
- [22] Yalin Liao, Aleksandar Vakanski, Min Xian, David Paul, and Russell Baker. A review of computational approaches for evaluation of rehabilitation exercises. *Computers in biology and medicine*, 119:103687, 2020.
- [23] Yanan Liu, Hao Zhang, Dan Xu, and Kangjian He. Graph transformer network with temporal kernel attention for skeleton-based action recognition. *Knowledge-Based Systems*, 240:108146, 2022.
- [24] Ziyu Liu, Hongwen Zhang, Zhenghao Chen, Zhiyong Wang, and Wanli Ouyang. Disentangling and unifying graph convolutions for skeleton-based action recognition. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 143–152, 2020.
- [25] Nguyen S M and Tanguy P. Cognitive architecture of a humanoid robot for coaching physical exercises in kinaesthetic rehabilitation. In *International Workshop on Cognitive Robotics*, 2016.
- [26] Aleksa Marusic, Sao Mai Nguyen, and Adriana Tapus. Evaluating kinect, openpose and blazepose for human body movement analysis on a low back pain physical rehabilitation dataset. In *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*, pages 587–591, 2023.
- [27] Vittorio Mazzia, Simone Angarano, Francesco Salvetti, Federico Angelini, and Marcello Chiaberge. Action transformer: A self-attention model for short-time pose-based human action recognition. *Pattern Recognition*, 124:108487, 2022.
- [28] Sao Mai Nguyen, Maxime Devanne, Olivier Remy-Neris, Mathieu Lempeur, and Andre Thepaut. A medical low-back pain physical rehabilitation dataset for human body movement analysis. In *International Joint Conference on Neural Networks*, 2024.
- [29] Chiara Plizzari, Marco Cannici, and Matteo Matteucci. Skeleton-based action recognition via spatial and temporal transformer networks. *Computer Vision and Image Understanding*, 208:103219, 2021.
- [30] Sara Sardari, Sara Sharifzadeh, Alireza Daneshkhah, Bahareh Nakisa, Seng W. Loke, Vasile Palade, and Michael J. Duncan. Artificial intelligence for skeleton-based physical rehabilitation action evaluation: A systematic review. *Computers in Biology and Medicine*, 158:106835, 2023.
- [31] Lei Shi, Yifan Zhang, Jian Cheng, and Hanqing Lu. Skeleton-based action recognition with directed graph neural networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 7912–7921, 2019.
- [32] Chuan-Jun Su, Chang-Yu Chiang, and Jing-Yan Huang. Kinect-enabled home-based rehabilitation system using dynamic time warping and fuzzy logic. *Appl. Soft Comput.*, 22:652–666, 2014.
- [33] Bazarevsky V, Grishchenko I, Raveendran K, Zhu T, Zhang F, and Grundmann M. Blazepose: On-device real-time body pose tracking. *CoRR*, abs/2006.10204, 2020.
- [34] Pichao Wang, Zhaoyang Li, Yonghong Hou, and Wanqing Li. Action recognition based on joint trajectory maps using convolutional neural networks. In *Proceedings of the 24th ACM international conference on Multimedia*, pages 102–106, 2016.
- [35] Sijie Yan, Yuanjun Xiong, and Dahua Lin. Spatial temporal graph convolutional networks for skeleton-based action recognition. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32, 2018.
- [36] Bing Yu, Haoteng Yin, and Zhanxing Zhu. Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, pages 3634–3640, 2018.
- [37] Bruce X.B. Yu, Yan Liu, Xiang Zhang, Gong Chen, and Keith C.C. Chan. Egcnn: An ensemble-based learning framework for exploring effective skeleton-based rehabilitation exercise assessment. In Lud De Raedt, editor, *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI-22*, pages 3681–3687. International Joint Conferences on Artificial Intelligence Organization, 7 2022. Main Track.

Marusic, A., Annabi, L., Nguyen, S. M., and Tapus, A. (2023). Analyzing Data Efficiency and Performance of Machine Learning Algorithms for Assessing Low Back Pain Physical Rehabilitation Exercises. European Conference on Mobile Robots. <https://dx.doi.org/10.1109/ECMR59166.2023.10256318>

- [38] Cao Z, Hidalgo Martinez G, Simon T, Wei S, and Sheikh Y. A. Openpose: Realtime multi-person 2d pose estimation using part affinity fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.