

Multi-Joint Adaptive Motion Imitation in Robot-Assisted Physiotherapy with Dynamic Time Warping and Recurrent Neural Networks

Ali Ashary, *Student Member IEEE*, Madan M. Rayguru *Member IEEE*, Payman SharafianArdakani, *Student Member IEEE*, Irina Kondaurova, and Dan O. Popa, *Senior Member IEEE*.

Abstract—Robot-assisted physiotherapy offers a promising avenue for easing the burden on healthcare professionals and providing treatment in the comfort of one's home. Typically, physiotherapy requires the repetitive movements until a certain efficiency metric is achieved. In the field of robot-assisted physiotherapy challenges include accurately determining the quality of imitation between robot and human movements, and tailoring the robot behavior to match the subject's abilities. This paper presents an innovative modular framework for Adaptive Motion Imitation (AMI) in the context of multi-joint robot-assisted physiotherapy. The proposed framework utilizes a deep Gated Recurrent Unit (GRU) Neural Network and Segment Online Dynamic Time Warping (SODTW). The SODTW cost is employed as a measure to determine the closeness between the movements of the robot and the subject. The GRU, which uses the range of motions and the fundamental frequency components of joint trajectories as inputs, forecasts dynamic and periodic reference trajectories for the robot joints. By modifying the input frequency coefficients according to the subject's SODTW cost, the output of the GRU is adapted to adapt the robot's motion with the subject's imitation capabilities. The division of the prediction and adaptation elements of our framework greatly streamlines testing and coding, and boosts the scalability of the algorithm. The efficacy of the proposed AMI framework was experimentally assessed with a group of 15 participants and the social robot Zeno in our lab. The results demonstrate the validity of the proposed framework in adapting the behavior of the robot according to the subject's imitation abilities.

I. INTRODUCTION

Physiotherapy and rehabilitation play a pivotal role in healthcare, aiding patients in their recovery from surgical aftereffects, injuries, and neurological conditions [1]. Over the past few decades, the exploration of various types of social and medical robots, as well as learning algorithms for robot-assisted physiotherapy, has been a significant focus of research [2]. These advancements hold the potential to enhance patient outcomes in several ways, such as improved mobility, increased strength, and facilitated independent living [3]. Motivated by these potential benefits, some healthcare professionals have begun to incorporate robot-enabled solutions into their physiotherapy programs [4]. Robot-assisted physiotherapy programs have particularly garnered the interest of medical professionals specializing in Autism Spectrum Disorders (ASD). Commercially available social robots like Nao and Milo, which offer interactive functionality, are being utilized to assist special educators in enhancing the social, physical, and communication skills

of children with autism [5], [2]. Current research evidence indicates promising outcomes for ASD treatment programs that incorporate social robots [6], [7]. However, challenges persist in areas such as technology integration, training methodologies, additional costs, and the standardization of treatment protocols for robot-assisted physiotherapy, presenting ongoing obstacles for researchers and educators.

Imitation learning has emerged as a potent instrument in robot-assisted physiotherapy, particularly for patients with Autism Spectrum Disorder (ASD). A variety of imitation learning algorithms have been developed to equip robots with the ability to emulate the movements necessary for physiotherapy and rehabilitation sessions [8], [9]. Once trained, these robots can demonstrate the learned movements to patients and record their responses for subsequent evaluation [10]. In a study by Zheng et al., an architecture was proposed for robot-assisted skill training specifically designed for children with ASD [11]. Techniques such as deterministic policy gradient, approximate dynamic programming, and recurrent neural networks are frequently employed for imitation learning in social robots [12], [13], [14].

Xu et al. introduced a novel shared control technique grounded in reinforcement learning, aimed at assisting individuals in walking tasks [15]. A master-slave robotics system, underpinned by reinforcement learning, was presented in [16] for the purpose of mirror therapy for patients with limb impairments. In another study, an adaptive ankle exoskeleton control was validated, demonstrating its potential to enhance the walking capabilities of patients [17]. Deep Reinforcement Learning (DRL) approaches have demonstrated their capacity to tackle high-dimensionality problems, a longstanding limitation of traditional RL techniques. In a recent study by Taghavi et al., a Deep Deterministic Policy Gradient (DDPG) method was proposed for adaptive motion imitation, specifically designed for training children with ASD [18]. This approach predicted the so-called "shape" (e.g., magnitude) and "speed" (e.g., frequency) factors derived from the periodic recorded joint motion of subjects. However, this DDPG method was primarily designed to learn the map from trajectories on a single joint, and it batch trains the sequenced motion data of the robot and child. This could potentially lead to an increase in the input dimension for DDPG with variations in motion. Consequently, the approach faces challenges when generalizing to multiple joints and longer training episodes [12]. Moreover, Reinforcement Learning (RL) and Deep Reinforcement Learning (DRL)-based approaches, while powerful, are known to be com-

This work was supported by any NSF grants #1838808 and #1849213. Authors are with Louisville Automation and Robotics Research Institute (LARRI), University of Louisville, Kentucky, U.S.A., Contact: ali.ashary@louisville.edu, madanmohan.rayguru@louisville.edu

putationally demanding and time-consuming. This is largely due to their requirement for extensive interaction with the environment to explore optimal policies [13]. These factors can pose significant challenges in real-world applications, particularly in scenarios where computational resources or time are limited [14].

In contrast to Reinforcement Learning (RL), Recurrent Neural Network (RNN)-based techniques have demonstrated superior performance in predicting temporal data sequences in dynamically evolving tasks, such as physiotherapy exercises [19], [20]. Doering et al. developed an RNN-based imitation learning approach for social interaction, which was grounded in data from human-to-human interactions [21]. In another study, a gaze-based imitation learning and master-to-robot policy transfer mechanism was proposed for instructing tasks that necessitate force sensor data [22]. The Long Short-Term Memory RNN (LSTM-RNN) was leveraged for imitation learning, based on a visual change-based image representation [23]. A unique objective function and LSTM-RNN were employed to predict the future behavior of pedestrians [24]. Kawaharazuka et al. proposed a novel methodology for adaptive imitation learning, which incorporated dynamic constraints and parametric bias [25]. A modified RNN was trained using demonstration images, control inputs, and an additional parameter known as parametric bias. This approach enabled the RNN to adapt to variations in task execution.

Another challenge in the current state of adaptive robot imitation [17], [22], [18], [25] is the absence of generalization and standardization across different settings. Given that these studies employ diverse training strategies and network architectures, they may yield varying results, even for similar robot-assisted physiotherapy routines. The adoption of a uniform metric to quantify the similarity or discrepancy in the task executions of the robot and the subject, and the use of the same metric to train or adapt the neural networks, could address the issue of standardization. The Segment Online Dynamic Time Warping (SODTW) algorithm, as presented in our recent work [2], could be an appropriate choice due to its invariance of signal temporal alignment. In the context of robot-assisted physiotherapy sessions, it's crucial not only to capture the difference in amplitude between the subject's and the robot's motion, but also the speed at which the subject repeats the motion. The SODTW algorithm is capable of efficiently capturing both of these aspects, thereby offering the potential to standardize the evaluation of the subject's physical ability. However, further research is needed to fully explore and address these challenges.

Contributions: In order to address the challenges associated with multi-joint adaptive motion imitation, we propose the use of a deep Gated Recurrent Unit (GRU) network for predicting periodic physiotherapy motions of a robot, and the SODTW distance as a foundation for adapting the robot's movements based on the responses of the subject. This approach involves reversing the input-output mapping used in our previous work on DDPG [26], which helps maintain a constant input dimension in the presence of

motion variations.

The key contributions of this paper are as follows:

- We introduce a novel modular framework for adaptive motion imitation in robot-assisted physiotherapy. This framework stands apart from previous approaches in that the reference trajectory predictor (GRU) and the adaptation algorithm (SODTW Module) can be independently modified for different types of physiotherapy sessions. This modularity simplifies the extension of our framework to complex motions involving multiple joints and standardizes motion adaptation.
- We present a new training strategy for a multivariate GRU capable of generating reference trajectories for the robot based on the subject's imitation ability. The network is trained using only the fundamental Fourier coefficients (amplitude and frequency component) and the range of periodic motion as inputs. Once trained, the GRU can generate a suitable reference trajectory for the Zeno robot, which can adapt to different variations in shape, speed, or their combination according to a fuzzy inference engine.
- We validate the proposed GRU + SODTW Framework through subject trials on our Zeno robot. During the training phase, we collected joint trajectory data from 15 subjects imitating an upper body motion demonstrated by Zeno. The data from ten subjects were processed, re-sampled, and normalized for training the GRU. We then asked five subjects to intentionally modify the motion and used the data to test our proposed framework. The experiments show that Zeno's motion adapts well to all the variations in motion imitation from the test subjects.

The paper is organized as follows: in section II, we discuss the proposed AMI scheme, and in section III, we present a new deep GRU-RNN architecture. Finally, section IV presents our conclusions and discusses future work.

II. PROPOSED AMI FRAMEWORK

Our proposed framework for multi-joint adaptive motion imitation encompasses a series of tasks executed in a synchronized manner. A robot-assisted physiotherapy session commences with a specific social robot, in our case Zeno [27], demonstrating a specific periodic upper arm motion to human subjects. The subjects are instructed to replicate the robot's upper body motion as accurately as possible. The movements of the subjects are captured through an RGBD camera, and the joint trajectories for both the human and Zeno are computed. The periodic joint trajectories of the subject are processed through a Discrete Fourier Transform (DFT) to generate the fundamental components of amplitude (shape) and frequency (speed). Simultaneously, the multi-joint SODTW cost, which measures the similarity between the recorded human motion and the reference Zeno motion, is computed. The fundamental DFT coefficients are then normalized to ensure uniform training across all ranges of motion. These coefficients, along with the range of motion, are then input into our GRU for training, enabling it to predict suitable reference joint trajectory commands for Zeno.

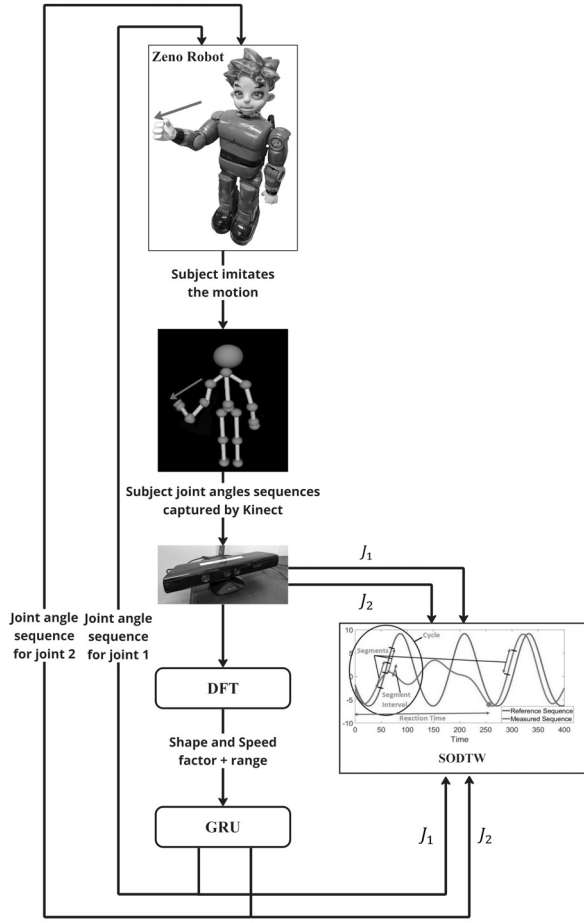


Fig. 1: Multi Joint Adaptive Motion Imitation (AMI) Architecture.

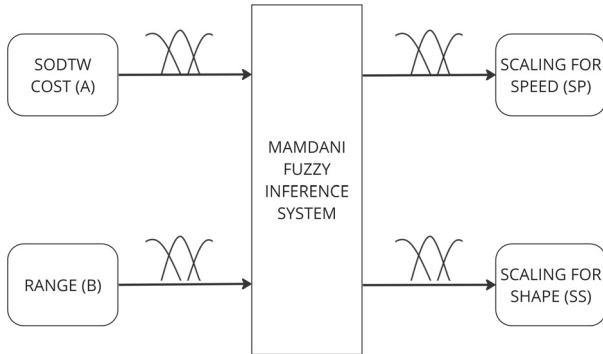


Fig. 2: Fuzzy Engine to Modify GRU-RNN Inputs.

The multi-joint SODTW cost is utilized as a metric to modify the inputs (DFT coefficients) so that the trained GRU can adjust the reference commands for Zeno proportionally. The decision to not use the SODTW cost as another input for the GRU stems from the desire to simplify the framework. While RNNs are generally effective in mapping the temporal relationship in the input data, the SODTW cost is not explicitly linked to the input data, which could result in the GRU requiring more computation time to learn the map for all motion variations. By keeping the SODTW cost computation separate and using it to directly change the inputs, we can achieve a similar adaptation to motion variations in speed, amplitude, or a combination of both in a more efficient

manner [19], [20], [21], [22], [23], [24], [25].

The Segment Online Dynamic Time Warping (SODTW) algorithm, a variant of the Dynamic Time Warping (DTW) algorithm, is utilized to compare and align two-time series sequences that may exhibit variations in time and speed [26]. Given a reference signal of M samples and a measured signal of N samples, the SODTW algorithm [2] computes the difference between the samples of the two signals using a dynamic error computation approach. The mathematical representation of a single joint SODTW process is as follows:

$$D_{i,j} = ||x_i - y_j|| + \min(D_{i-1,j-1}, D_{i-1,j}, D_{i,j-1}) \quad (1)$$

Here, $i=1, \dots, M$; $j=1, \dots, N$. $D_{i,j}$ represents the Euclidean norm of the difference between the i th sample of the reference trajectory and the j th sample of the measured trajectory. The SODTW algorithm recursively calculates the similarity cost between two samples based on the cost of previous samples. The final cost for the entire motion sequences is given by $D_{M,N}$.

The multi-joint SODTW is a convex combination of single-joint SODTWs for each joint in a particular motion. Mathematically

$$D_o = \sum_{x=1}^n \kappa_x D_{i,j}^x \quad (2)$$

where $x = 1, \dots, n$ is the number of joints involved in that specific motion and $\kappa_x < 1$ is a positive constant decided by the weightage of a particular joint in the overall motion.

When the subject is unable to satisfactorily mimic a specific motion demonstrated by the robot, the SODTW cost escalates. This high SODTW cost signals the need to simplify the physiotherapy until the subject achieves a satisfactory cost. Conversely, a low SODTW cost indicates that the subject has perfectly imitated the motion. In such instances, the complexity of the physiotherapy should be increased. Generally, it has been observed that slower motions are more challenging to imitate compared to faster ones (refer to Table II). Therefore, the reference trajectory of the Zeno robot is slowed down when the user has a lower SODTW cost, and vice versa.

III. GRU FOR REFERENCE GENERATION

A Gated Recurrent Unit (GRU) based deep network for motion imitation would consist of multiple GRU layers followed by dense layers. The architecture would look like this:

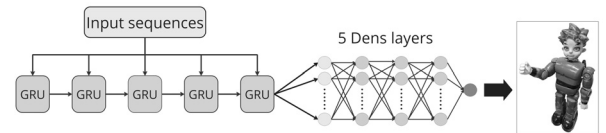


Fig. 3: Proposed five-layer GRU-RNN Architecture.

Input Layer: The input to the network would be the sequence of joint angles or positions. **GRU Layers:** There would be 5 GRU layers. Each GRU layer would have a certain number of hidden units. These layers are responsible for capturing the temporal dependencies in the input data. **Dense Layers:** After the GRU layers, there would be 5 dense

(or fully connected) layers. These layers are used for further processing of the features extracted by the GRU layers. The forward pass of a GRU involves two types of gates: update gates and reset gates. For each time step t , the GRU computes the output $\hat{y}_t$ using the input x_t and the previous internal state s_{t-1} .

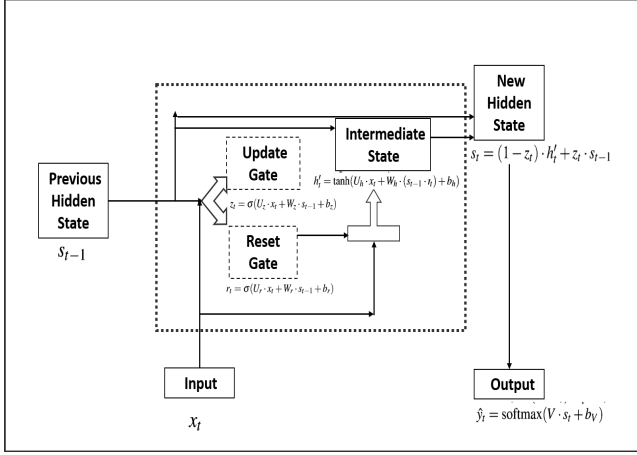


Fig. 4: GRU States and Transition Diagram.

Where the parameters of the GRU are: $U_z, U_r, U_h \in \mathbb{R}^{n_i \times n_v}$, $W_z, W_r, W_h \in \mathbb{R}^{n_i \times n_i}$, $b_z, b_r, b_h \in \mathbb{R}^{n_i \times 1}$, $V \in \mathbb{R}^{n_v \times n_i}$, $b_V \in \mathbb{R}^{n_v \times 1}$ and n_i, n_v are the sizes of internal memory and vocabulary.

This GRU architecture is pretty similar to the Long Short-Term Memory (LSTM) architecture, which retain important features while being more efficient computationally. The input signals and the previous hidden state signals pass through the reset and update gates, to decide the new hidden states and output. The reset gate state r_t and update gate state z_t decide how much of the information in previous hidden state and how much of the similarity have to be retained. The intermediate state h'_t integrates the reset gate state with the previous hidden state, before passing on to compute the new hidden state s_t and the output. For this work, the goal of the suggested GRU is to adjust its output to match both the shape and speed of the subject's motion. To achieve this, we establish a convex loss function:

$$E_{loss_x} = \frac{(\sum_{j=1}^N (||o_j - y_j||^2 + ||(o_j - o_{j-1}) - (y_j - y_{j-1})||^2))}{NT} \quad (3)$$

$$E_{loss} = \sum_{x=1}^n \kappa_x E_{loss_x} \quad (4)$$

In this equation, T represents the time period of the joint trajectory, o_j is the j th output sample from the network, and y_j is the j th sample of recorded subject data. The first term in the equation represents the shape error, which is equivalent to a simple RMSE. The second term accounts for the speed discrepancy between the RNN output and the actual subject trajectory. To train the GRU, we want to know the values of all parameters that minimize the total loss. We used Stochastic Gradient Descent (SGD) to solve this problem.

IV. EXPERIMENTAL RESULTS

To validate the proposed framework, a series of experiments were conducted involving 15 human subjects. Each participant was instructed to replicate a predefined hand motion guided by our Zeno robot. The identical motion was performed at three distinct speed settings: Slow, Normal, and Fast. Utilizing a Kinect camera, the subjects' motions were captured, and the recorded data underwent processing to calculate the Discrete Fourier Transform (DFT) of the motion waveform, along with its Multi-Joint SODTW cost relative to the robot's motion. Given that the Zeno robot's arm possesses four degrees of freedom—shoulder joints α and β , and elbow joints θ and γ —a preliminary analysis involved determining the percentage involvement of each joint in executing the fist bump motion. For the specific motion chosen, it was found that the joint angles γ and α played dominant roles. Consequently, the SODTW cost for similarity was computed based solely on these joint angles.



Fig. 5: Imitation Experiment Setup.

TABLE I: Weight (in terms of range) of joint angle contribution during a fist bump motion exercise.

Joint Angle	Right-hand Hammer motion
α	0.795548411
β	0.048682412
γ	0.852457816
θ	0.131227326

Subsequently, the average Multi-Joint SODTW cost was calculated for the subjects imitating motions at fast, slow, and normal speeds. The results indicated that fast motion was relatively easier to replicate, while the slow mode posed greater difficulty for the subjects.

For the implementation of the AMI training process, widely-used Python libraries TensorFlow and Keras were

employed. The deep networks were constructed, and the training process was executed on an NVIDIA V100 TENSOR CORE GPU, connected to an ASUS VivoBook Pro laptop equipped with an Intel Core i7 processor running at 2.80 GHz and 16 GB of RAM. The setup involved training durations ranging from 3 to 5 minutes for each subject, depending on factors such as the number of joint sequences and epochs.



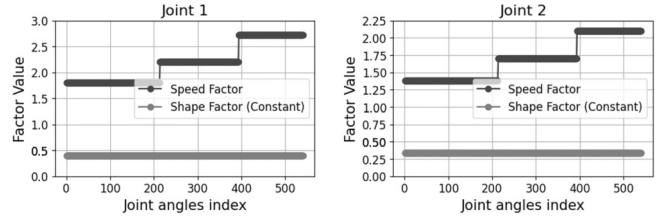
Fig. 6: Zeno's arm with its four degrees of freedom [27].

TABLE II: SODTW cost statistics for all subjects.

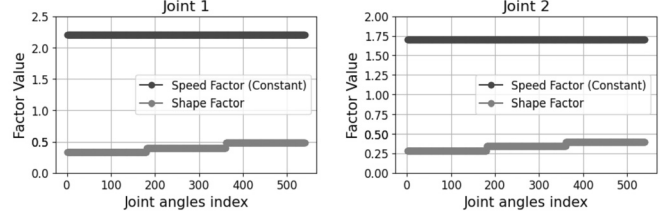
Speed Mode	Average SODTW cost	Std Dev.
Normal	32.46	3.21
Slow	70.07	6.18
Fast	19.21	1.40

To validate the efficacy of the proposed approach in forecasting future joint trajectories based on Fourier coefficients, we employed an identical dataset comprising data from 15 subjects. Fourier coefficients and the motion range of each subject were computed, forming the basis for training the neural network. To accommodate various shapes and speeds exhibited by subjects, we employed resampling techniques, adjusting Fourier coefficients accordingly (Fig 7). Our approach leveraged the two most significant amplitude components as the shape factor and frequency for the speed factor.

The deep network architecture underwent modulation by manipulating the number of GRU layers, dense layers, and learning rate, leading to the achievement of optimal results. The final architecture is depicted in Figure 3. Each subject's trajectory data was segmented into three cycles, with each cycle consisting of 60 samples. This fixed cycle length ensures consistent data processing. It is crucial to note that the inputs for each subject remained consistent across all output samples, encompassing Fourier coefficients and the range of motion.



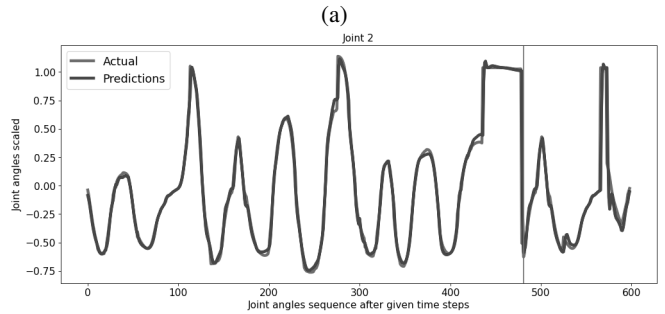
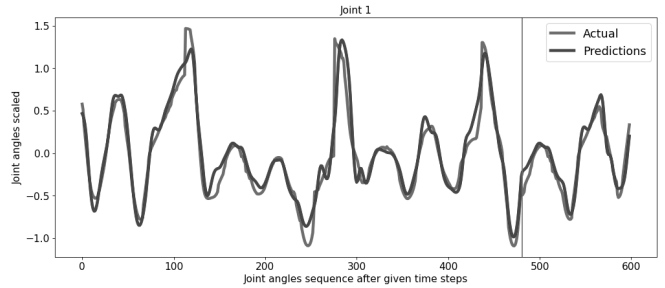
(a) Oversampled (slower), Normal and Undersampled (faster) Version of One Subject Data with Fixed Range.



(b) Lower range, Normal, and Higher range motion of One Subject Data With Fixed Speed

Fig. 7: Fourier Coefficients (shape and speed)

The training of our network employed the widely-used Adam optimizer, facilitating weight updates through back-propagation in time. The Adam optimizer is particularly well-suited for optimizing convex and stochastic loss functions, employing first-order gradient descent. The average joint error during training is 0.03, and during testing, it is 0.04 (see Fig 8).



(b) Fig. 8: GRU-RNN Proposed Output.

To assess the robustness of our framework, we computed the SODTW cost for each subject. Different subjects were selected and deliberately instructed to deviate from Zeno's motion in various ways, including moving slower, faster, with a lower range, or a higher range. The reference trajectory, generated from the GRU architecture, was then compared to each subject's joint trajectory (Fig. 9, 10, 11, 12), and the corresponding RMSE values are summarized in Table III.

In one specific test scenario, a subject was directed to precisely follow a fist bump motion, resulting in a minimal SODTW cost of 25.8, lower than the average. Leveraging this cost and the observed range, we inputted these values into our Fuzzy engine. The engine recommended adjusting the amplitude of Fourier coefficients by a factor of 1.15 and decreasing the frequency by a factor of 0.8, effectively elevating the difficulty level. Subsequently, we adjusted the inputs to the network accordingly, and the network's outputs are illustrated in Figure 13. The RMSE values for the first and second joints were found to be 0.06 and 0.07, respectively.

In summary, our proposed framework showcases the capability to accurately adapt to both speed and shape changes in subjects' motions, as demonstrated by the experimental results. Furthermore, the framework excels in scenarios where subjects' motions exhibit simultaneous changes in both speed and shape, emphasizing its versatility and effectiveness.

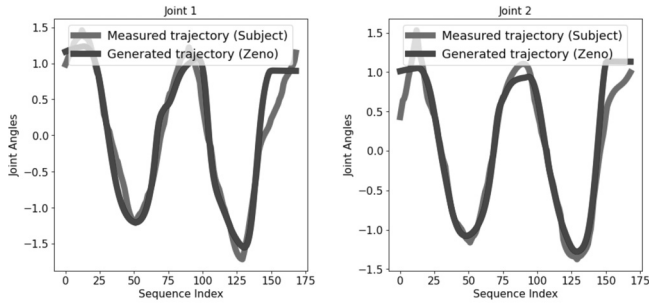


Fig. 9: Generated sequence after training for the subject performing the motion with a slow speed.

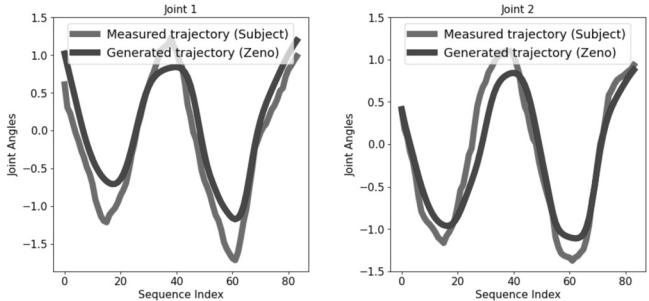


Fig. 10: Generated sequence after training for the subject performing the motion with a fast speed.

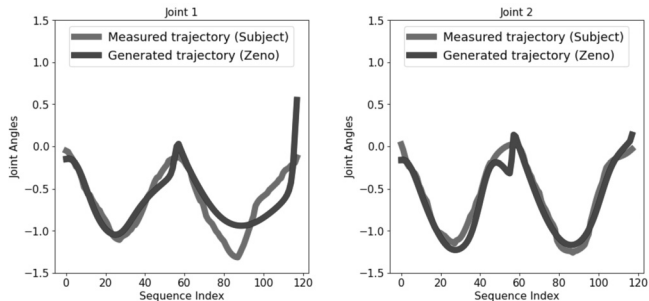


Fig. 11: Generated sequence after training for the subject performing the motion with a lower range.

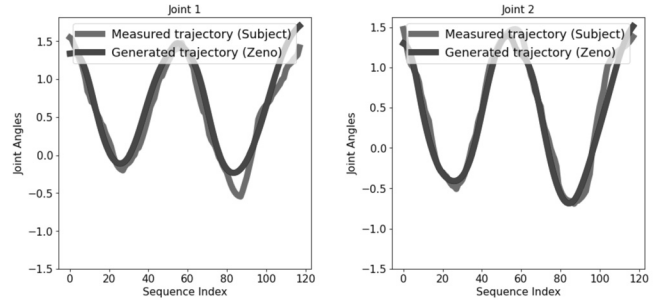


Fig. 12: Generated sequence after training for the subject performing the motion with a higher range.

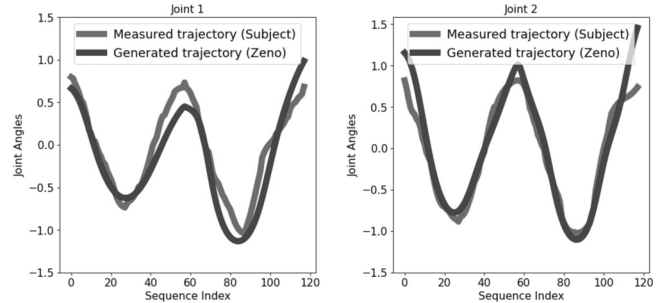


Fig. 13: Generated sequence after training for the subject performing the motion with a higher range and slower speed.

TABLE III: RMSE values for different task variations.

Task Variation	First joint RMSE	Second joint RMSE
Slower Speed	0.05	0.06
Faster Speed	0.04	0.05
Lower Range	0.02	0.04
Higher Range	0.03	0.04
Higher Range+ Slower Speed	0.06	0.07

V. CONCLUSIONS AND FUTURE WORK

In this study, we introduced a modular AMI framework for robot-assisted physiotherapy aimed at individuals with motion disparity. This framework employs a deep GRU network to predict the reference joint trajectory for our Zeno robot, using the Fourier coefficients and range of motion of a specific human subject as input. The network leverages the SODTW algorithm to adjust the inputs based on the subject's ability to mimic Zeno's actions, thereby constructing a fuzzy inference engine. Our research, which involved 15 adult subjects, validated the effectiveness of our AMI framework by predicting suitable robot exercises that mirror human subject movements. The AMI framework we developed is scalable and can be expanded to include motions with more degrees of freedom and longer motion sequences. In addition to these findings, we discovered that the GRU network provided a more robust prediction of the joint trajectory, enhancing the overall performance of the Zeno robot. This has significant implications for the future of robot-assisted physiotherapy, particularly for individuals with ASD. Looking ahead, we plan to conduct more comprehensive experiments with this framework, including trials with children with ASD and various physiotherapy programs. We also aim to explore the potential of integrating other deep learning models to further enhance the prediction accuracy and effectiveness of the therapy.

REFERENCES

- [1] R. M. Andrade, P. H. F. Ulhoa, and C. B. S. Vimieiro, "Designing a highly backdrivable and kinematic compatible magneto-rheological knee exoskeleton," in *2022 International Conference on Robotics and Automation (ICRA)*, 2022, pp. 5724–5730.
- [2] E. Nazzi, E. Canzi, G. Piga, A. Galassi, G. Lippi, and G. Benassi, "Segment online dtw for smart rehabilitation of asd children: A preliminary study," in *Proceedings of the 4th EAI International Conference on Smart Objects and Technologies for Social Good (GOODTECHS)*, 2015, pp. 122–126.
- [3] A. Frolov, I. Kozlovskaya, E. Biryukova, and P. Bobrov, "Use of robotic devices in post-stroke rehabilitation," *Neuroscience and behavioral physiology*, vol. 48, pp. 1053–1066, 2018.
- [4] T. Goyal, S. Hussain, E. Martinez-Marroquin, N. A. T. Brown, and P. K. Jamwal, "Stiffness-observer-based adaptive control of an intrinsically compliant parallel wrist rehabilitation robot," *IEEE Transactions on Human-Machine Systems*, vol. 53, no. 1, pp. 65–74, 2023.
- [5] M. Kirtay, J. Chevalère, R. Lazarides, and V. V. Hafner, "Learning in social interaction: Perspectives from psychology and robotics," in *2021 IEEE International Conference on Development and Learning (ICDL)*, 2021, pp. 1–8.
- [6] H. Mahdi, S. A. Akgun, S. Saleh, and K. Dautenhahn, "A survey on the design and evolution of social robots — past, present and future," *Robotics and Autonomous Systems*, vol. 156, p. 104193, 2022.
- [7] I. B. Wijayasinghe, I. Ranatunga, N. Balakrishnan, N. Bugnariu, and D. O. Popa, "Human-robot gesture analysis for objective assessment of autism spectrum disorder," in *Int. J. Soc. Robot.*, vol. 8, no. 5, 2016, p. 695–707.
- [8] P. Abbeel and A. Y. Ng, "Apprenticeship learning via inverse reinforcement learning," in *Proceedings of the Twenty-first International Conference on Machine Learning*, 2004, pp. 1–8.
- [9] B. D. Argall, S. Chernova, M. Veloso, and B. Browning, "A survey of robot learning from demonstration," in *Robotics and Autonomous Systems*, vol. 57, no. 5, 2009, pp. 469–483.
- [10] A. Hussein, M. M. Gaber, and E. Elyan, "Imitation learning: A survey of learning methods," in *Artificial Intelligence Review*, vol. 48, no. 1, 2017, pp. 31–60.
- [11] Z. Zheng, E. M. Young, A. R. Swanson, A. S. Weitlauf, Z. E. Warren, and N. Sarkar, "Robot-mediated imitation skill training for children with autism," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 6, pp. 682–691, 2016.
- [12] W. Liu, L. Peng, J. Cao, X. Fu, Y. Liu, Z. Pan, and J. Yang, "Ensemble bootstrapped deep deterministic policy gradient for vision-based robotic grasping," *IEEE Access*, vol. 9, pp. 19 916–19 925, 2021.
- [13] P. Sharma, D. Pathak, and A. Gupta, "Third-person visual imitation learning via decoupled hierarchical controller," in *Advances in Neural Information Processing Systems*, H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, Eds., vol. 32. Curran Associates, Inc., 2019.
- [14] S. Park, J. H. Park, and S. Lee, "Direct demonstration-based imitation learning and control for writing task of robot manipulator," in *2023 IEEE International Conference on Consumer Electronics (ICCE)*, 2023, pp. 1–3.
- [15] W. Xu, J. Huang, Y. Wang, C. Tao, and L. Cheng, "Reinforcement learning based shared control for walking aid robot and its experimental verification," *Advanced Robotics*, vol. 29, no. 22, pp. 1463–1481, 2015.
- [16] J. Xu, L. Xu, Y. Li, G. Cheng, J. Shi, J. Liu, and S. Chen, "A multi-channel reinforcement learning framework for robotic mirror therapy," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 5385–5392, 2020.
- [17] S. S. P. A. Bishe, T. Nguyen, Y. Fang, and Z. F. Lerner, "Adaptive ankle exoskeleton control: Validation across diverse walking conditions," *IEEE Transactions on Medical Robotics and Bionics*, vol. 3, no. 3, pp. 801–812, 2021.
- [18] N. Taghavi, M. H. A. Alqatamin, and D. O. Popa, "Ami: Adaptive motion imitation algorithm based on deep reinforcement learning," in *2022 International Conference on Robotics and Automation (ICRA)*, 2022, pp. 798–804.
- [19] Z. C. Lipton, J. Berkowitz, and C. Elkan, "A critical review of recurrent neural networks for sequence learning," *arXiv Preprint arXiv:1506.00019*, 2015.
- [20] S. Yang, W. Zhang, R. Song, J. Cheng, H. Wang, and Y. Li, "Watch and act: Learning robotic manipulation from visual demonstration," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2023.
- [21] M. Doering, D. F. Glas, and H. Ishiguro, "Modeling interaction structure for robot imitation learning of human social behavior," *IEEE Transactions on Human-Machine Systems*, vol. 49, no. 3, pp. 219–231, 2019.
- [22] H. Kim, Y. Ohmura, A. Nagakubo, and Y. Kuniyoshi, "Training robots without robots: deep imitation learning for master-to-robot policy transfer," *IEEE Robotics and Automation Letters*, vol. 8, no. 5, pp. 2906–2913, 2023.
- [23] S. Yang, W. Zhang, R. Song, W. Lu, H. Wang, and Y. Li, "Explicit-to-implicit robot imitation learning by exploring visual content change," *IEEE/ASME Transactions on Mechatronics*, vol. 27, no. 6, pp. 4920–4931, 2022.
- [24] D. X. V. R. Johnson-Roberson, "M bio-ilstm: a biomechanically inspired recurrent neural network for 3-d pedestrian pose and gait prediction," *Autom. Lett.*, vol. 4, no. 2, p. 1501, 2019.
- [25] K. Kawaharazuka, Y. Kawamura, K. Okada, and M. Inaba, "Imitation learning with additional constraints on motion style using parametric bias," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 5897–5904, 2021.
- [26] N. Taghavi, J. Berdichevsky, N. Balakrishnan, K. C. Welch, S. K. Das, and D. O. Popa, "Online dynamic time warping algorithm for human-robot imitation," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 3843–3849.
- [27] N. A. Torres, N. Clark, I. Ranatunga, and D. Popa, "Implementation of interactive arm playback behaviors of social robot zeno for autism spectrum disorder therapy," in *Proceedings of the 5th International Conference on Pervasive Technologies Related to Assistive Environments*, 2012, pp. 1–7.