Deep Active Inference for Engagement Recognition in Robot-Assisted Autism Therapy

Shyrailym Shaldambayeva[1], Saparkhan Kassymbekov[1], Anara Sandygulova[1], and Almas Shintemirov[1,2], *Senior Member, IEEE*

***Abstract*—Robot-Assisted Autism Therapy (RAAT) is becoming increasingly popular due to its ability to enhance therapeutic outcomes for children with autism spectrum disorders (ASD). RAAT offers consistent, personalized, and engaging interventions, complementing traditional therapies and supporting social and cognitive development. However, the rising use of RAAT also brings challenges for therapists, who must make real-time, personalized decisions during sessions. This demands a deep understanding of individual needs and effective strategies, placing significant cognitive pressure on therapists.**

To address these challenges, our research aims to develop an AI-driven RAAT system that supports therapists by assisting with decision-making during sessions. By analyzing real-time data and leveraging prior knowledge, the AI system can suggest appropriate interventions and adapt strategies to each child's specific needs. In this study we apply Deep Active Inference (dAIF) model to enable the robot to learn and improve its interventions over time, enhancing the effectiveness of RAAT. The proposed approach aims to ultimately offer a more personalized and dynamic therapeutic experience for children with ASD.

***Index Terms*—Deep Active Inference, Autism Spectrum Disorder (ASD), Robot-Assisted Autism Therapy**

I. Introduction

The use of Robot-Assisted Autism Therapy (RAAT) is rapidly increasing due to its potential to enhance therapeutic outcomes for children with autism spectrum disorders (ASD). Its growing popularity can be attributed to its potential to provide consistent, personalized, and engaging therapeutic interventions that can complement traditional therapies [1]. Robots used in RAAT are designed to interact with children in a controlled and repeatable manner, offering unique opportunities for social and cognitive development. As a result, RAAT is being increasingly adopted in therapeutic settings as an innovative tool to support children with ASD.

However, the increasing use of RAAT presents significant challenges for specialists, who are often overwhelmed by the need to make real-time decisions during therapy sessions. These decisions must account for the child's immediate responses, their individual needs, and broader therapeutic goals. Additionally, therapists must possess a deep understanding of what strategies would work best for different children, given the considerable variability in how children with ASD respond to interventions [2]. This combination of real-time decision-making and the need for personalized approaches could place considerable cognitive demands on therapists, potentially limiting the effectiveness of RAAT.

Given the above stated challenges, there is a growing need for AI-driven tools that can assist therapists by supporting decision-making processes during RAAT sessions. Such assistive tools could be deployed for analyzing data in real time, suggesting appropriate interventions, and/or adapting strategies based on the individual needs of each child. By leveraging AI, therapists can focus more on engaging with the child, while the AI system provides insights based on prior knowledge and accumulated data on effective interventions.

Building on this need for AI assistance, our research work focus on developing an AI-driven RAAT system that incorporates both prior knowledge and available data from past research. For this purpose, we are utilizing an extensive dataset that comprises 194 therapy sessions involving 34 children [3]. By leveraging the existing data, we could train our AI model to recognize patterns, which would facilitate further tailoring interventions to the specific needs of individual child, thus making therapy sessions more effective and personalized.

Deep Active Inference (dAIF) has emerged as a powerful framework providing a robust approach to modeling intelligent behavior in uncertain environments. The Active Inference approach, first introduced by Friston in [4], is based on the idea that an agent will perceive and interact with its environment in a way that minimizes its free energy. This principle integrates perception, action, and learning, allowing an agent to use a hierarchical generative model to predict sensory inputs and adjust its actions to reduce the discrepancy between these predictions and actual observations. By continuously refining its generative model and minimizing prediction errors, Active Inference enables the agent to maintain a representation of the environment that is both robust and adaptive. dAIF combines the principles of Active Inference with deep learning, enabling Active Inference to be scaled up to tasks that are significantly larger and more complex than those previously addressed solely by Active Inference [5].

We propose applying dAIF to the RAAT scenario, aiming to enabling the robot to not only assist in real-time decision-making but also to learn and improve its interventions over time. This paper presents initial study of integrating dAIF into RAAT, focusing on enhancing effectiveness of robot-assisted therapies for children with autism. Additionally, we explored different sets of features from our dataset to determine which combination would yield the best performance. This was done

[1] Department of Robotics and Mechatronics, School of Engineering and Digital Sciences, Nazarbayev University, Astana, Kazakhstan.

[2] Department of Electrical Engineering and Automation, School of Electrical Engineering, Aalto University, Espoo, Finland.

This work was funded under the Nazarbayev University CRP project #11022021CRP1502 and FDCRGP project #021220FD1751.

to ensure that the most relevant and informative features were used to optimize the model's ability to make accurate predictions and adapt its interventions. Feature selection is crucial in minimizing noise, improving generalization, and enhancing the overall effectiveness of the robot's decision-making process, particularly in a complex and dynamic environment like RAAT.

II. Deep Active Inference

In our study, we adopted the deep active inference model presented in [6]. The model couples the free energy principle with neural networks and utilizes a Partially Observable Markov Decision Process (POMDP) framework, assuming the agent does not have full access to the underlying state of the environment, thus enabling the agent to operate effectively in complex and uncertain environments. A POMDP is characterized by a set of hidden states, actions, observations, a transition model, an observation model, and a reward function. The agent maintains a belief, or a probability distribution over possible hidden states, which it updates based on the observations it receives over time. This belief-update process enables the agent to make informed decisions in the face of uncertainty. The agent's objective is to minimize its expected free energy (EFE) into the future. To do this, the agent must evaluate the EFE for each possible policy, which involves projecting future states and observations based on its generative model. This requires running the model forward through time to estimate the potential free energy of each policy. However, computing the EFE for every policy across large policy spaces or long time horizons involves significant computation, making the process intractable for complex environments. To overcome this, deep Active Inference (dAIF) employs deep neural networks to approximate the necessary components of EFE calculation, providing a general scalable approach.

Formally, the agent's objective can be expressed as:

$$-F_t = -E_{q(s_t)}[lnp(o_t|s_t)] + D_{KL}[q(s_t)||p(s_t|s_{t-1}, a_{t-1})] + D_{KL}[q(a_t|s_t)||p(a_t|s_t)] \tag{1}$$

where o_t is the observation at time t, s_t is the environment state, a_t is the agent's action and $E_{q(s_t)}$ is the expectation over the variational density $q(s_t)$.

In [6], each term of Eq. 1 is approximated using neural networks. The first term is the perception model that represents mapping of observations to states and is estimated by a variational autoencoder (VAE). The second term is the state prediction error, calculated as the Kullback-Leibler (KL) divergence between the state at time t and the state predicted at $t-1$. To compute this, the agent needs a transition model, which gives the probability of the current state based on the previous state and action, and is trained to minimize the prediction error using a feedforward network. The last term contains two densities: a value network, which maps a state action pair to an estimated EFE, and an action model, which returns a distribution over actions given states.

The complete model is made up of four neural networks that approximate the components of the variational free energy. This model not only scales effectively to high-dimensional spaces but also enhances the agent's decision-making capabilities in scenarios where uncertainty plays a significant role. As a result, the deep active inference model proposed in [6] is particularly well-suited for human-robot interaction applications, where inferring internal state of a human for effective response to inherent uncertainty of its future actions is a critical challenge.

III. Experimental Setup

A. Dataset

For this project we used the QAMQOR dataset which was developed by Zhanatkyzy et al. [7] using video recordings of 34 children with autism during robot-assisted therapy. Their study involved the development and implementation of 24 diverse robot activities, each with varying levels of social interaction. These activities were thoroughly analyzed in total of 194 therapy sessions and about 48 hours of videos. The dataset includes 2D data of 25 body and leg keypoints, 21 keypoints per hand, and 70 facial keypoints, all extracted using the OpenPose library [8]. The main aim of the research was to identify which types of robot activities were most effective in meeting individual needs and promoting social and behavioral development among the children. Additionally, the study explored the relationship between specific child characteristics and the behavioral outcomes of each activity. Video data from the sessions were meticulously annotated frame by frame, with each frame representing one second. The annotations provide detailed session information, child attributes, pose landmarks, activity descriptions, and two types of engagement scores (binary and a five-point scale).

B. Environment

To evaluate the proposed algorithm we developed a customized environment implemented using the OpenAI Gym library [9] and the QAMQOR dataset [7]. The state space of our environment consists of two distinct elements. The first element includes variables that remain unchanged throughout the interaction with a child. These variables - such as age, ADOS-2 score, the presence of Attention Deficit Hyperactivity Disorder (ADHD), the child's compliance status (compliant or non-compliant), and their verbal proficiency (verbal or nonverbal) - are established before the interaction begins. The second element consists of variables that change dynamically during the interaction. Specifically, this involves pose landmarks - body, face, and hands - generated by the OpenPose Library.

The action space is represented by 26 distinct actions. These categories are grouped into eight main blocks of interaction: "Dances," "Songs," "Touch Me," "Social Acts," "Storytelling," "Emotions," "Imitations," and "Hello and Bye". Although these categories were not used in the current environment, with distinct activities being employed instead, it may be worth exploring the use of these categories in future work. It is important to note that the specific action space can vary depending on the session ID and child ID, as not every child

participated in all activities, leading to differences in the action space.

To enable goal-directed behavior in Active Inference, it is essential to incorporate a representation of a desired state or goal within the generative model. While reinforcement learning achieves this through reward functions, Active Inference uses a prior distribution over expected observations, often referred to as 'prior preferences' or a 'goal distribution'. These prior preferences influence the inference over control states, guiding the agent to select actions that are likely to lead to states where preferred observations are expected [4]. In our environment, however, we opted to use rewards to guide the agent's behavior, combining elements of both approaches.

The reward calculation for each step incorporates a thorough evaluation of both static and dynamic parameters present in the current state. A key dynamic parameter in this assessment is the engagement level, which is evaluated by a human expert. For our preliminary experiments, we used binary engagement levels, where the engagement level ranges from 0 to 1 and corresponds to reward values -1 to 0, respectively. The agent achieves its target state when the cumulative reward for a step reaches zero, indicating a successful and desirable action. This reward calculation determines the target state of the environment, which is reached when the robot performs all actions with a reward value of 0. This means the agent has executed all necessary actions to achieve optimal engagement from the child.

IV. Results

The training of the dAIF agent was performed on the environment presented in III-B with four different variations of a state space. All four versions included all child attributes, but differed on subsets of incorporated pose landmarks. Only binary engagement levels were used for reward calculations. Total reward value, i.e. accumulated reward at each time step, was used as an evaluation metric. In the first variant (Fig. 1) we used all pose landmarks (body, left hand, right hand and face). The second variant uses only face landmarks (Fig. 2). Face and both left and right hand landmarks were used in the third variant (Fig. 3), and face and body were used in the last one (Fig. 4). All four figures (1-4) show the total reward value (returns) over 5000 training episodes.

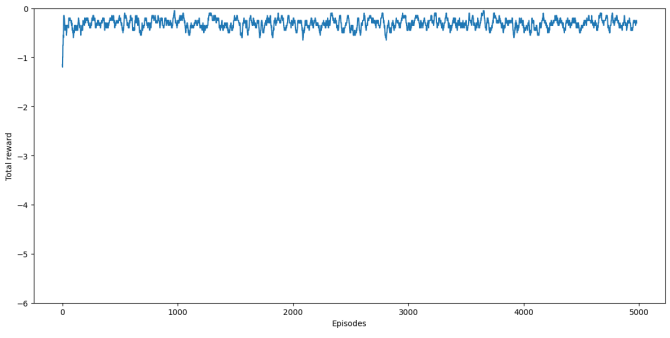

Fig. 1. Total reward value with all body pose landmarks

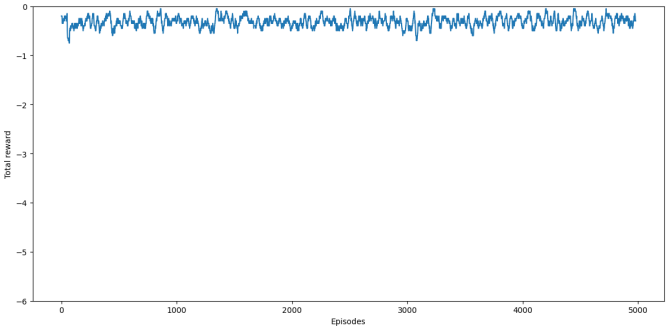

Fig. 2. Total reward value with face landmarks

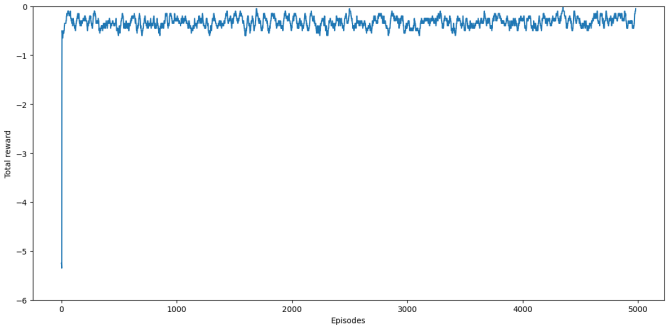

Fig. 3. Total reward value with face and hands landmarks

Among the four variations of the model, two (all and face+hands) experienced an initial drop in returns before quickly recovering and stabilizing close to zero. The other two models (face and face+body), however, started off with a stable performance, with returns remaining close to zero from the very beginning. Despite these differences in the early stages, all four models ultimately demonstrated similar long-term behavior, maintaining stable returns near zero with only minor fluctuations. This consistency across the models suggests that overall the dAIF agents are capable to reach a comparable level of stability.

The face+hands version of the model experienced a notably sharp initial drop, where the returns value plunged close to -6 at the very beginning. This decline was more pronounced

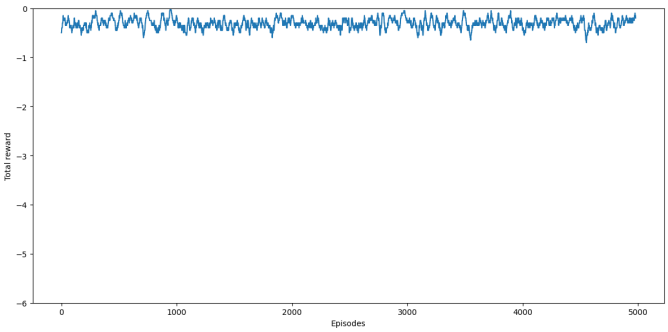

Fig. 4. Total reward value with face and body landmarks

compared to the version with all pose landmarks. This might suggest that using hand landmarks may not be ideal for training models to recognize engagement. Although it eventually recovered, this initial difficulty implies that hand landmarks may introduce noise or complexity that the model struggles to process effectively.

In contrast, the model that incorporated all possible landmarks experienced a smaller initial drop, implying a slight improvement in stability but still showing some signs of initial difficulty. On the other hand, the models that utilized only face or face and body landmarks performed more consistently.

These observations indicate that hand landmarks might not be as reliable or informative for recognizing engagement, potentially complicating the model's learning process. Focusing on face and body landmarks alone appears to provide a more stable and effective foundation for the model, suggesting that hand landmarks may not contribute positively to engagement recognition and could even hinder model performance.

V. Conclusion and Future Work

Despite the differences in how each variation of the model handled the initial phases of training, all four versions ultimately demonstrated promising results during the training process for estimating child engagement during therapy. While the models that included hand landmarks faced some initial instability - especially the one that combined face and hand landmarks, which experienced a sharp drop in returns - they all eventually stabilized and delivered consistent results. The models using only face or face and body landmarks started off strong and maintained stable performance throughout.

Overall, these results indicate that while the choice of landmarks can influence the early stages of model training, our models were able to adapt and achieve reliable performance in the long run. This suggests that, regardless of the specific combination of landmarks used, our proposed approach to recognizing human engagement has high potential for development of intelligent assistive tools for RAAT.

As a next work, we will enhance our model by transitioning from binary classification to a five-point scale for engagement levels. This approach would allow us to capture a broader range of engagement, providing a more comprehensive understanding of user behavior. Additionally, we plan to compare our model with classical deep reinforcement learning techniques to assess its performance relative to other approaches. This comparison will offer valuable insights into the strengths and potential areas for improvement in our approach, helping to further refine our active inference model's ability to recognize human engagement.

References

[1] J.-J. Cabibihan, H. Javed, M. Ang, and S. M. Aljunied, "Why robots? A survey on the roles and benefits of social robots in the therapy of children with autism," *International journal of social robotics*, vol. 5, pp. 593–618, 2013.

[2] B. Scassellati, H. Admoni, and M. Matarić, "Robots for Use in Autism Research," *Annual review of biomedical engineering*, vol. 14, no. 1, pp. 275–294, 2012.

[3] N. Rakhymbayeva, Z. Balgabekova, M. Nurmukhamed, K. Burunchina, W. Johal, and A. Sandygulova, "To Transfer or Not to Transfer: Engagement Recognition within Robot-Assisted Autism Therapy," in *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2022, pp. 1002–1006.

[4] K. J. Friston, J. Daunizeau, J. Kilner, and S. J. Kiebel, "Action and Behavior: a Free-Energy Formulation," *Biological cybernetics*, vol. 102, pp. 227–260, 2010.

[5] B. Millidge, "Deep Active Inference as Variational Policy Gradients," *Journal of Mathematical Psychology*, vol. 96, p. 102348, 2020.

[6] O. van der Himst and P. Lanillos, "Deep Active Inference for Partially Observable MDPs," in *Active Inference*. Springer International Publishing, 2020, pp. 61–71.

[7] A. Zhanatkyzy, Z. Telisheva, A. Amirova, N. Rakhymbayeva, and A. Sandygulova, "Multi-purposeful activities for Robot-Assisted Autism Therapy: What Works Best for Children's Social Outcomes?" in *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*, 2023, pp. 34–43.

[8] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime multi-person 2D pose estimation using part affinity fields," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7291–7299.

[9] G. Brockman, V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, and W. Zaremba, "Openai gym," 2016.