Emotion-based Robotic Action Optimization System for Human-Robot Collaboration

Jordan Murphy, Jesse Parron, Weitian Wang, and Rui Li*, *Member, IEEE*

Abstract—**Although collaborative robots aim to boost productivity in manufacturing, misalignment between robot's actions and the human's intentions of the collaboration can cause discomfort or frustration, potentially discouraging future collaborations. Inspired by human-to-human interactions, this paper aims to help solve this problem by enabling a collaborative robot to adjust how it moves and acts based on human emotions to improve the overall collaboration process. To achieve this goal, an emotion-based robotic action optimization system was developed and integrated into a collaborative robot. The system utilizes hierarchical reinforcement learning (HRL) to train and guide the robot to adjust its actions according to detected human emotions. Specifically, this paper introduces (1) a HRL model that leverages a vision-audio-based emotion recognition model to determine and adjust robot actions (movement speed, drop-off distance, reaction time, and rate of success) according to human emotions. The goal of this model is to avoid negative emotions of the human user that are triggered by the robot actions. (2) A robot motion control method driven by recognized human intentions and actions from the HRL model, guiding the robot arm and gripper to adjust movements and deliver parts as desired. (3) objective and subjective evaluation experiments to evaluate the effectiveness of the developed system. The results and analysis of the experiments demonstrated the effectiveness of our developed system in a human-robot collaboration setting.**

I. INTRODUCTION

In the manufacturing scene, human-robot collaboration is set in stone in the way the robots interact with their human collaborators. The main purpose of these collaborations is to improve productivity and efficiency, relying on repetitive and predictable movements to do so [1]–[4]. This method of interactions can be effective, but the preprogrammed motions could potentially deter any future collaborations or even leave the human collaborators feeling uncomfortable or irritated by the robot's actions if they do not align with the human's expectations. As such, a challenge in enhancing human-robot collaboration would be to take a human's emotional state into consideration when a collaborative robot decides its next sequence of actions. To enable a collaborative robot to adapt its actions based on the emotions of its human collaborator and draw inspiration from human-to-human interactions, this paper develops an emotion-based robotic action optimization system. The system employs a hierarchical reinforcement learning approach to train the robot to modify its actions while considering the emotions of its human collaborator. This optimization system has been developed and integrated into a Franka Emika Panda robot arm [5], utilizing a human's recognized emotions [6], [7] to adjust its various actions with changes in its movement speed, drop-off distance, reaction time, and rate of success. Specifically, our work includes the following three parts. First, the training of a hierarchical reinforcement learning model that uses the recognized visual-audio information-based emotions to determine the desired robot actions that need to be optimized. Second, recognizing the human collaborator's intentions via vocal commands to enable the robot to call upon a specific order of saved joint state positions to assist in co-assembly tasks. Third, validation experiments and subjective user evaluations to evaluate the effectiveness of the robot's changing actions through real-world assembly tasks. The results and analysis of the experiment help demonstrate the effectiveness of our developed system in a human-robot collaboration setting.

This paper's contributions are: (1) developing an emotion-based robotic action optimization system that adjusts robot actions to human emotions for a friendly human-robot collaboration; (2) creating a hierarchical reinforcement learning system that updates user preferences online and in real-time, enabling flexible, easy training; and (3) implementing a lightweight method that considers personalized emotional factors, allowing the robot to dynamically respond to emotional changes.

II. RELATED WORKS

A. *Deep Learning in Manufacturing Robots*

Deep learning [8] has been applied to manufacturing robots for tasks like object detection. Chen et al. [9] used a faster R-CNN to train a UR5 robot to recognize objects such as tools and office supplies. The model, trained with various backbones (VGG-M, VGG-16) and different datasets, achieved 67.9% accuracy on a 10-class test set. Deep learning can also help predict human motion in human-robot collaboration. For example, Liu et al. [10] used a combined CNN-LSTM model to forecast human movements during a computer assembly task, achieving up to 83% accuracy. Another study done by Male et al. [11] used a CNN and wearable IMUs to help a UR3 robot recognize its environment and a human's actions during an assembly task. The system, tested in assembling a box and seat, accurately predicted action status in both offline and online trials. Wang et al [1] proposed a teaching-learning collaboration (TLC) model and deployed it on a UR10 robot arm. Based on maximum entropy inverse reinforcement learning, this TLC model (MaxEnt-IRL) allowed a human to program the collaborative robot using natural language instructions to demonstrate how to complete the assembly task. Another example, developed by Murphy et al [6], utilizes a lightweight facial recognition based on convolutional neural networks (CNN) to enable a Franka

Jordan Murphy, Jesse Parron, Weitian Wang, and Rui Li are with the School of Computing at Montclair State University, Montclair, NJ 07043 USA (Corresponding Author: Dr. Rui Li, E-mail: liru@montclair.edu).

Emika Panda collaborative robot to understand human emotions during co-assembly tasks.

B. Reinforcement Learning in Robotics Control

Reinforcement learning [12] enables robots to learn through trial and error, receiving rewards based on their actions. Implementation ranges from simple Q-Learning tables [13] to complex models combining reinforcement learning with deep learning [14], [15] for handling larger action spaces. Švaco et al. [16] developed a reinforcement learning algorithm for a UR3 robot to minimize actions when moving an object to a target. Their system achieved optimal performance with a learning rate of 0.4 and a discount factor of 0.6 as training progressed. Utilizing deep reinforcement learning, robots can learn to traverse complex paths to a designated target [17], [18]. For example, Iskandar et al. [19] used deep reinforcement learning to train swarm robots to navigate toward a target. Each simulated Webot learned individually while interacting with the swarm and environment. Over time, their total number of collisions decreased while gained rewards and the number of goals reached increased. Reinforcement learning is also used frequently in robots designed for manufacturing as well. One example of such was deployed by Pane et al. [20] onto a UR5 robot to train it to follow shaped paths. Using an actor-critic method, the robot's performance was measured by temporal difference rewards. Tested on square, circle, and custom paths, the reinforcement learning approach outperformed PD, MPC, and ILC controllers on the square and circular task. Another study by Oliff et al. [21] introduced a deep Q-learning (DQN) agent in a human-robot manufacturing simulation. Using shaped rewards to enforce assembly rules, their system reduced idle time and minimized performance disparities across operators and robots.

III. System Overview

This paper introduces an emotion-based robotic action optimization (ERAO) system (Fig. 1) designed for collaborative robots. The system focuses on optimizing four key factors that influence co-assembly tasks and human emotions: movement speed, drop-off distance, reaction time, and success rate. These factors are selected due to their significant impact on collaboration quality and task success, which in turn affect human emotional responses during interaction. To recognize human emotions, the system uses a combination of visual and audio data, analyzing both facial expressions and speech. Additionally, human intentions are identified using natural language processing. Based on the detected emotions and intentions, the robot employs a hierarchical reinforcement learning framework to optimize its actions which is aiming to enhance human satisfaction by promoting neutral or positive emotional states. The hierarchical reinforcement learning framework include two layers: a higher layer for selecting what type of action adjustment the robot must make and a lower layer for the sub-action choices for each adjustment type. Depending on a person's recognized emotions and the previously trained values in the sets of Q-tables, the system chooses first from the higher layer which type of adjustment is needed, then repeat the process on the lower layer with the specific action choice. Throughout the collaboration, the human can vocally command the collaborative robot to hand them the next necessary assembly part. Once the robot recognizes the intent of the command, it will go through set joint state positions for the target gesture to retrieve the assembly part and hand it to the human. Depending on real-time human emotions, the system will be rewarded or penalized to further improve its action selection process.

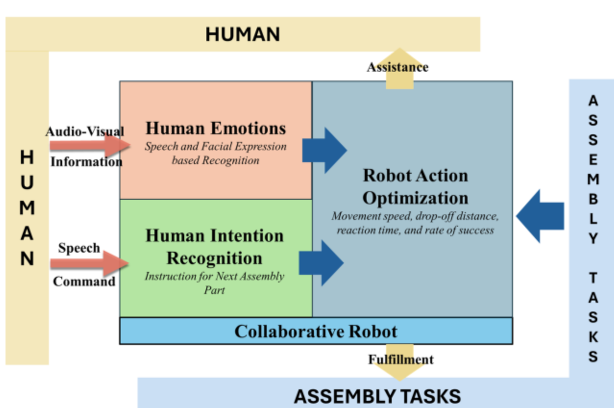

Figure 1. ERAO System Overview

IV. Methodology

A. Hierarchical Reinforcement Learning

Our system uses hierarchical reinforcement learning [12], [22], [23] to enable a collaborative robot to adjust its actions based on the user's analyzed emotions. The system iteratively learns over many trials to gain cumulative rewards for its actions. The robot has its current state s defined from the current human emotion. There is a set of potential states $S = \{s_1, s_2, \ldots s_\Psi\}$ for the total of Ψ human emotions. In the collaboration, the robot will select an action a to take from a set of potential actions $A = \{a_1, a_2, \ldots a_N\}$ that will reward or penalize the robot if the outcome of the action was desirable or not, where N is the total number of potential actions. The reward R at time t can be represented as:

$$R_t = \begin{cases} +R_{[s,a]} & Reward \\ 0 & Neutral \\ -R_{[s,a]} & Penalty \end{cases} \tag{1}$$

where $R_{[s,a]}$ is a reward value for a state and action pair$[s, a]$, which is positive when the system is rewarded for a preferred action and negative when penalized for an unpreferred one. When choosing sequential actions, the system measures the quality of the action using Q-Learning [13], finding the optimal decision-making policy based on the state and action pair. These preferences are saved in a table Q, where the rows are based on the potential states s and the columns are based on the potential actions a (Fig. 2). The table is updated after each action using the following equation:

$$Q[s,a] = Q[s,a] + \alpha(r + \gamma(Q[s',a']) - Q[s,a]) \tag{2}$$

where $Q[s,a]$ is the reward value in the table for state s and action a, α represents the learning rate, r represents the current reward for the state and action pair $[s,a]$, γ represents a discount factor for future rewards Q for the value in the table for the next for state s' and action a'. Using the table, the system can then choose its next action with the following function:

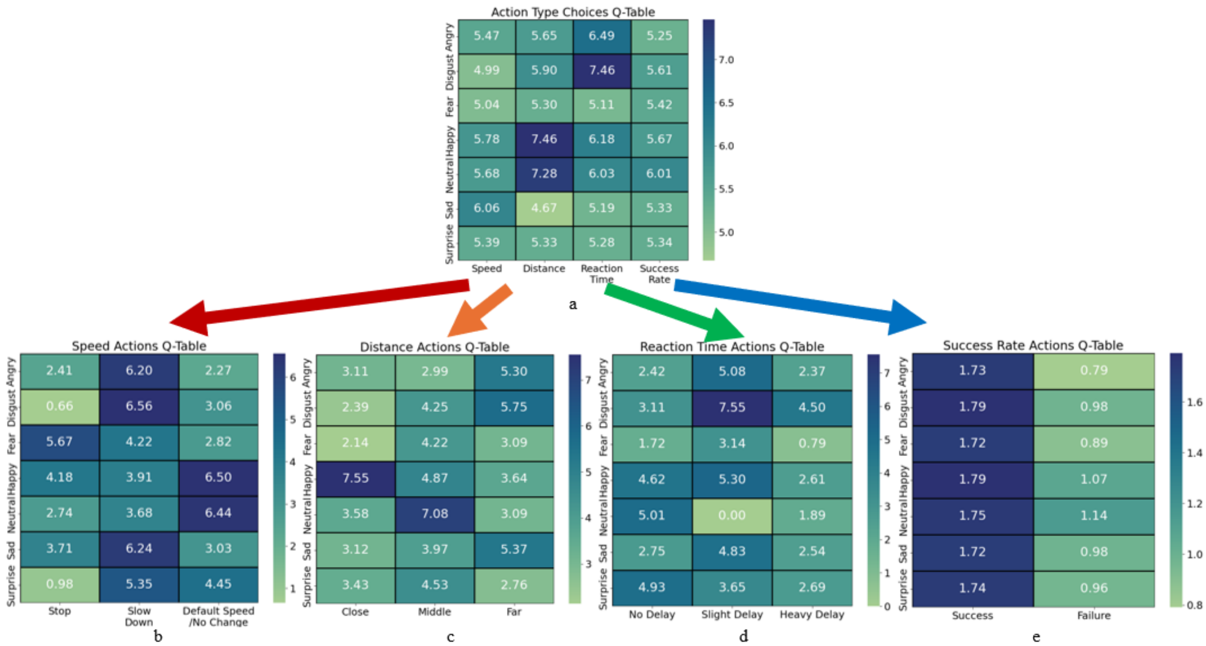

Figure 2. Structure of Hierarchical Reinforcement Learning

$$a(s,\varepsilon) = \begin{cases} random(A) & \varepsilon < x \\ argmax\ (Q[s,A]) & \varepsilon \geq x \end{cases} \tag{3}$$

where the selected action a based on a decaying epsilon value ε representing the probability of exploring other actions by comparing it to a randomly generated decimal value x between zero and one, $x \sim U(0,1)$. The function returns an action within the action set A, which would either be a random action choice in the set, or as the action with the maximum value in the reward table Q based on the current state s.

The system's decision-making has two layers, illustrated by the heatmaps in Fig. 2. Both layers use the human collaborator's emotional states as input but differ in their action sets. The higher layer selects the type of change: movement speed, drop-off distance, reaction time, or success rate. The lower layer then chooses a specific action for that type: stop, slow, or default speed for movement; close, middle, or far for drop-off; no, slight, or heavy delay for reaction time; and correct or incorrect tool pickup for success rate. Based on the proposed hierarchical reinforcement learning, the system determines its action type first, then calculates an optimized sub-action from the action type. The reward for that sub-action is then given to the respective lower layer table. The higher layer also receives the reward from the sub-action, but weighted to balance the choices so one is not preferred over the other. This can be represented as

$$R_t^{high} = R_t^{lower} * w^{choice}[s,a] \tag{4}$$

where the higher layer reward R_t^{high} is set by the lower layer reward R_t^{lower} multiplied by the weight of the action choice w^{choice} at the state and action pair $[s,a]$. The weight for the higher layer action choice is added to have the selection be balanced instead of preferring one action over another, and is done after each iteration of training the higher layer. It is calculated using the following equation:

$$w_i^{choice}[s,a] = 1 - \alpha \left(\frac{choices_{[s,a]}}{\sum_a choices_{[s,a]}} \right) \tag{5}$$

where $w_i^{choice}[s,a]$ denotes the action choice weight at training iteration i for each state and action pair $[s,a]$, α represents the learning rate, and $choices_{[s,a]}$ represents the count of selected action type choices stored in a separate table generated from the previous training iteration, whose values are based on the state and action pair $[s,a]$.

B. Human Intention Recognition

For the robot to assist its human collaborator in our experiments, it needs to be able to follow human instructions. This paper utilizes a microphone next to the robot to recognize the user's voice and recognize human verbal command using Google Speech-to-Text API [24].

C. Robot Action Control

The system will command the robot to plan its movement using ROS Moveit [25] to grab the next part. Once the part is secured in the gripper, the robot will return to its previous position above the part so there would be no collisions, then hands off the part to the user in the desired action type calculated by the hierarchical reinforcement learning method introduced in the previous Section IV-A. Each time the robot must grab a selected part; it references an array of set positions

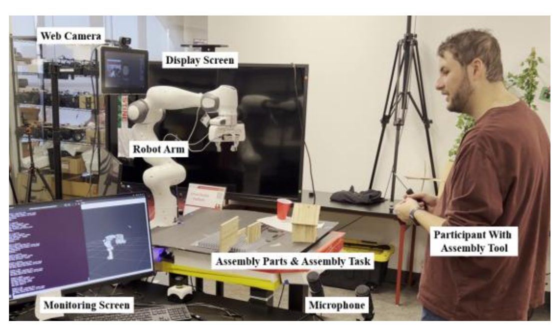

Figure 3. Experimental Setup

(a) (b) (c) (d) (e)

Figure 4. Human-Robot Co-assembly

with the index being of the part to grab. Each position for the robot can be represented as:

$$P^{\varphi} = \{(J_{\xi}^{\varphi}, V_{\xi}^{\varphi}) | J_{\xi}^{\varphi} \in [\alpha, \theta], V_{\xi}^{\varphi} \in [v_{min}, v_{max}], \xi \in [1, K]\} \tag{6}$$

where P^{φ} denotes the joint setting for arm gesture φ, $\varphi \in \{above,\ grab,\ drop\ off\}$ is sets of saved positions for the robot to call upon to properly grab a part for the user, J_{ξ}^{φ} denotes the rotation angle for joint ξ in gesture φ, V_{ξ}^{φ} denotes the speed for joint ξ in gesture φ. $J_{\xi}^{\varphi} \in [\alpha, \theta]$ and $V_{\xi}^{\varphi} \in [v_{min}, v_{max}]$ define the ranges for the rotation angle and the speed. $\xi \in [1, K]$ defines the total number of joints on the arm of the collaborative robot.

V. EXPERIMENT

A. Experimental Setup

Fig. 3 displays the experimental setup, where a collaborative robot will work alongside a human worker to complete the assembly. The robot used in this experiment is a Franka Emika Panda robot [5]. The setup also includes a web camera and microphone to analyze a person's facial expressions [6]-[7] and tone of voice to determine their overall dominant emotion during collaboration [6]. The resulting emotion output [6]-[7] is sent to the robot controller to decide how to change its actions accordingly. In the collaborative task, the robot assisted with assembling a birdhouse by handing over requested parts upon verbal cues from the user. As shown in Fig. 4, the user requests a part (a), the robot retrieves (b) and hands it over (c), and the user continues assembly (d), requesting the next part, which the robot prepares (e). User emotions were captured via microphone and webcam during the interaction. Based on detected user emotions [26] (anger, disgust, fear, happiness, neutral, sadness, and surprise), particularly negative ones, robots use hierarchical reinforcement learning to adjust actions (movement speed, drop off distance, reaction time, and/or success rate) to improve human satisfaction, measured by increased neutral and happy responses. Meanwhile, the Q-tables in the hierarchical reinforcement learning model are updated in real time to adapt to user preferences. Initially, Q-tables were pre-trained using randomly generated emotions. To evaluate the robot's adaptability in collaborative assembly, participants worked with the robot on repeated birdhouse assembly tasks. Trained Q-tables were recorded after approximately 15 and 30 minutes for later comparison.

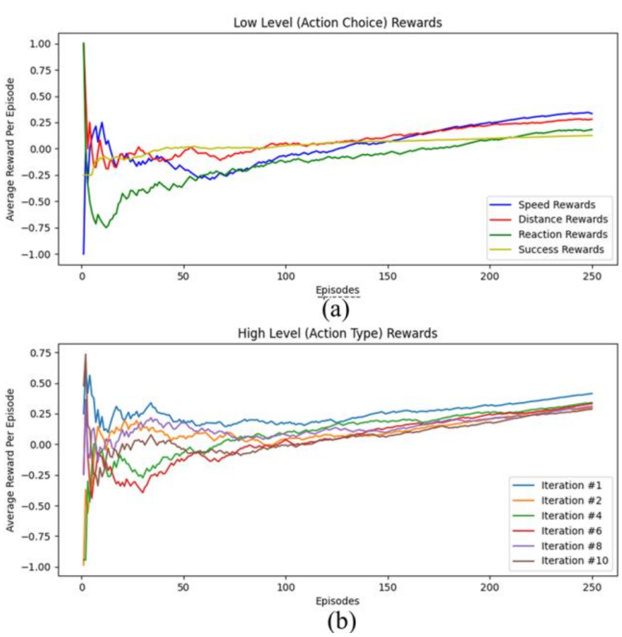

Figure 5. Average Reward Per Episode Graphs

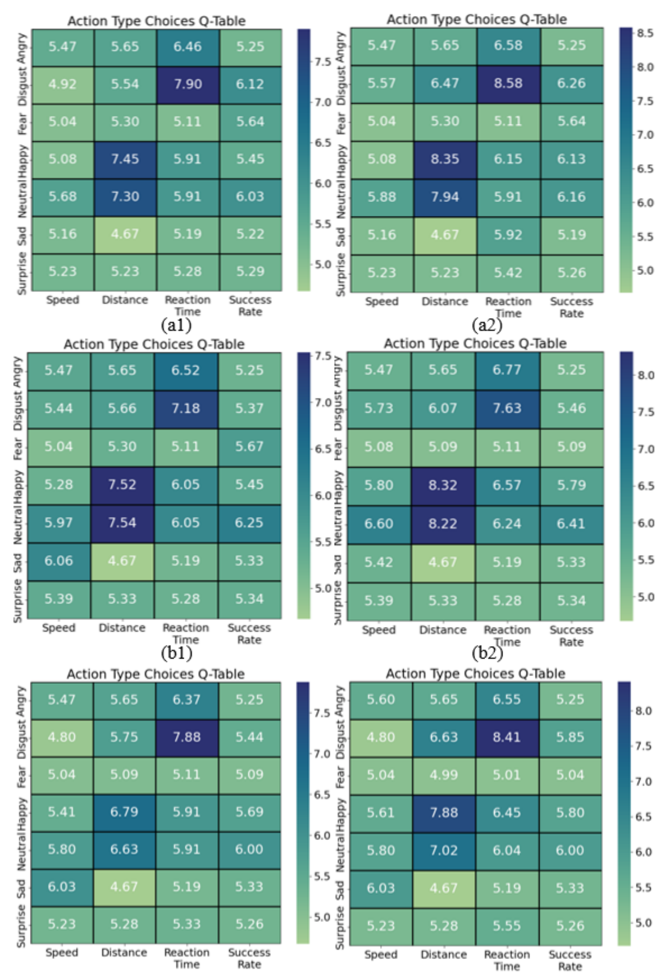

Figure 6. Heatmaps of Participants' Action Choice Q-Tables

B. Experimental Results

Before collaborating with a human, the robot requires initial training to avoid indecision from zeroed Q-values. To address this, it undergoes a brief simulated run using randomly selected emotions to establish a baseline for action-reward associations. As shown in Fig. 5, training begins with the lower layer of action choices, followed by the higher layer using prior results. Initial average rewards may fluctuate near extreme values but gradually stabilize between 0.3 and 0.4, indicating convergence and Q-table stabilization. In Fig. 5(b), average rewards for action types converge earlier, benefiting from prior lower-layer training. With this baseline, the system then trained with participants, generating two personalized Q-tables per person, one after ~15 minutes and another after ~30 minutes. These tables reflect each user's prevalent emotions and preferred robot actions. Fig. 6 shows heatmaps of action type Q-tables used by the system to choose changes in movement speed, drop-off distance, reaction time, or success rate. The left column (Fig. 6 (a1), Fig. 6 (b1), and Fig. 6 (c1)) displays Q-tables trained after one bird house assembly averaging at around 15 minutes of training, while the right column (Fig. 6 (a2), Fig. 6 (b2), and Fig. 6 (c2)) displays Q-tables after two bird house assemblies averaging at around 30 minutes of training. Each participant had a set of these Q-tables specifically for them, resulting in any preferences being visible when observing which values differed in one table compared to another. For example, in Fig. 6(a1) and Fig. 6(b1), participants had slightly higher values in drop-off distance and success rate choices when fear was detected, while in Fig. 6(c1), the participant had similar values across the board for that emotion. Between the first and second training sessions, different values in the Q-tables changed displaying either a reinforcement of someone's preferences if the value increased, or a correction if the value decreased. For example, the values for reaction time choices at anger and disgust as well as values for drop-off distance at happy and neutral had increased from Fig. 6(c1) to Fig. 6(c2), while the values for choices based on fear decreased and evened out from Fig. 6(b1) to Fig. 6(b2). Some preferences for participants were similar, seen in the higher values for distance choices from happy and neutral emotions. These preferences may be due to how participants reacted to the robot being common or could be any preferences from the reward values becoming known.

TABLE I. QUESTIONNAIRE FOR THE SUBJECTIVE EVALUATION EXPERIMENT

No.	Item
1	On a scale from 1-5, do you feel the developed system was **useful** for the assigned task?
2	On a scale from 1-5, do you feel the developed system was **easy to use**?
3	On a scale from 1-5, do you feel the developed system was **safe** to use?
4	On a scale from 1-5, do you feel **satisfied** with how the developed system operated?
5	On a scale from 1-5, do you feel people **would like to work with** a collaborative robot that adjusts its actions according to their emotions, like the one in the experiment?
6	On a scale from 1 to 5, how **important** do you feel it is to have a collaborative robot that adjusts its actions according to human emotions to assist with assigned assembly tasks?
7	On a scale from 1 to 5, do you feel the system's behaviors were **relevant** to the presented assembly task?
8	On a scale from 1-5, how would you rate the **quality** of the collaboration outcome?
9	On a scale from 1-5, how **comfortable** were you working with the collaborative robot that adjusts its actions according to your emotions?
10	On a scale from 1 to 5, do you think a robot system that adjusts its actions based on your emotions can **enhance** your interaction with it?
11	On a scale from 1 to 5, do you think the robot adjusting its actions helps you **feel better** when you are in negative emotions?
12	On a scale from 1 to 5, were you **aware** that the robot was attempting to make you feel happier or better via adjusting its actions, especially when you were experiencing negative emotions, even if its efforts did not immediately affect your emotion?

C. Subjective Evaluation

To evaluate system performance and user acceptance, 20 participants (11 male, 9 female) were recruited. Thirteen were aged 18–22, five were 24–30, and two were 60–70. Participants rated their experience on a 1–5 Likert scale through a questionnaire covering twelve factors, as shown in Table I. A rating of 5 indicates very good quality, high comfort, or strong agreement, while a rating of 1 reflects very poor quality, high discomfort, or strong disagreement. Table II displays the results of the subjective evaluation. The highest-rated factor was system safety (avg. 4.921), showing strong agreement that the system felt safe. Comfort followed closely (avg. 4.816, SD 0.380), indicating consistent feelings of ease during interaction. Perceived importance and preference for working with the system also scored high (avgs. 4.342 and 4.553, SDs 0.746 and 0.664), suggesting general approval with slight variation. Quality, usefulness, and ease of use had similar averages (4.158, 4.395, and 4.211) and moderate SDs, reflecting that most participants found the robot helpful and easy to work with. Relevance (avg. 4.132, SD 0.797) and satisfaction (avg. 4.184) were also rated well, though relevance showed more varied opinions. Participants agreed that emotion-based changes improved collaboration (avg. 4.421, SD 0.672). Emotional impact had the lowest average (3.737) and highest SD (1.110), showing mixed responses to this factor. Awareness of the robot's emotional adjustments was similarly mixed responses (avg. 3.763, SD 1.019), suggesting varied participant perception. Overall, the subjective evaluation confirmed the system's effectiveness and highlighted areas for improvement. These findings will inform future studies with a larger participant pool to further enhance user experience.

VI. CONCLUSION & FUTURE WORKS

This paper presented an emotion-based robotic action optimization system that adapts robot behavior in response to

TABLE II. SUBJECTIVE EVALUATION RESULTS

Factor	Total	Average	SD
Usefulness	83.5	4.395	0.542
Ease of use	80.0	4.211	0.769
Safety of the system	93.5	4.921	0.251
Satisfaction	79.5	4.184	0.606
Preference to work with	86.5	4.553	0.664
Importance	82.5	4.342	0.746
Relevance to the task	78.5	4.132	0.797
Quality	79.0	4.158	0.747
Comfort	91.5	4.816	0.380
Enhances Interaction	84.0	4.421	0.672
User Feels Better	71.0	3.737	1.110
User Awareness	71.5	3.763	1.019

human emotional changes using a hierarchical reinforcement learning approach. Co-assembly of a birdhouse was used for both objective and subjective evaluations, demonstrating the system's effectiveness and informing areas for improvement. Future work will involve analyzing participant feedback to enhance system quality and acceptance.

VII. ACKNOWLEDGMENT

This work is supported by the National Science Foundation under Grant CMMI-2301678, Grant CMMI-2338767 and CNS 2117308.

REFERENCES

[1] W. Wang, R. Li, Y. Chen, Z. M. Diekel, and Y. Jia, "Facilitating Human–Robot Collaborative Tasks by Teaching-Learning-Collaboration From Human Demonstrations," *IEEE Transactions on Automation Science and Engineering*, vol. 16, no. 2, pp. 640–653, 2019, doi: 10.1109/TASE.2018.2840345.

[2] W. Wang, R. Li, Z. M. Diekel, Y. Chen, Z. Zhang, and Y. Jia, "Controlling Object Hand-Over in Human–Robot Collaboration Via Natural Wearable Sensing," *IEEE Trans Hum Mach Syst*, vol. 49, no. 1, pp. 59–71, 2019, doi: 10.1109/THMS.2018.2883176.

[3] R. Li, and W. Wang, "Augmenting the Communication Naturalness via A 3D Audio-Visual Virtual Agent for Collaborative Robots," in *2021 IEEE International Conference on Big Data (Big Data)*, 2021, pp. 5944–5946. doi: 10.1109/BigData52589.2021.9671664.

[4] C. J. Conti, A. S. Varde, and W. Wang, "Human-Robot Collaboration With Commonsense Reasoning in Smart Manufacturing Contexts," *IEEE Transactions on Automation Science and Engineering*, vol. 19, no. 3, pp. 1784–1797, 2022, doi: 10.1109/TASE.2022.3159595.

[5] S. Haddadin, S. Parusel, L. Johannsmeier, S. Golz, S. Gabl, F. Walch, M. Sabaghian, C. Jähne, L. Hausperger, and S. Haddadin, "The Franka Emika Robot: A Reference Platform for Robotics Research and Education," *IEEE Robot Autom Mag*, vol. 29, no. 2, pp. 46–64, 2022, doi: 10.1109/MRA.2021.3138382.

[6] J. Murphy, J. Parron, M. Lyons, W. Wang, and R. Li, "A Visual Information-Based Bidirectional Emotion Interaction Interface for Friendly and Empathic Collaborative Robots," in *2024 International Conference on Networking, Sensing and Control (ICNSC)*, 2024, pp. 1–6. doi: 10.1109/ICNSC62968.2024.10760097.

[7] J. Loor, J. Murphy, and R. Li, "A Speech and Facial Information based Emotion Recognition System of Collaborative Robot for Empathic Human-Robot Collaboration," in *2024 33rd IEEE International Conference on Robot and Human Interactive Communication (ROMAN)*, 2024, pp. 2327–2332. doi: 10.1109/RO-MAN60168.2024.10731348.

[8] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.

[9] X. Chen, and J. Guhl, "Industrial Robot Control with Object Recognition based on Deep Learning," *Procedia CIRP*, vol. 76, pp. 149–154, 2018, doi: https://doi.org/10.1016/j.procir.2018.01.021.

[10] Z. Liu, Q. Liu, W. Xu, Z. Liu, Z. Zhou, and J. Chen, "Deep Learning-based Human Motion Prediction considering Context Awareness for Human-Robot Collaboration in Manufacturing," *Procedia CIRP*, vol. 83, pp. 272–278, 2019, doi: https://doi.org/10.1016/j.procir.2019.04.080.

[11] J. Male, and U. Martinez-Hernandez, "Deep learning based robot cognitive architecture for collaborative assembly tasks," *Robot Comput Integr Manuf*, vol. 83, p. 102572, 2023, doi: https://doi.org/10.1016/j.rcim.2023.102572.

[12] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement Learning: A Survey," *Journal of Artificial Intelligence Research*, vol. 4, pp. 237–285, May 1996, doi: 10.1613/jair.301.

[13] J. Clifton, and E. Laber, "Q-Learning: Theory and Applications," *Annu Rev Stat Appl*, vol. 7, no. Volume 7, 2020, pp. 279–301, 2020, doi: https://doi.org/10.1146/annurev-statistics-031219-041220.

[14] X. Wang, S. Wang, X. Liang, D. Zhao, J. Huang, X. Xu, B. Dai, and Q. Miao, "Deep Reinforcement Learning: A Survey," *IEEE Trans Neural Netw Learn Syst*, vol. 35, no. 4, pp. 5064–5078, 2024, doi: 10.1109/TNNLS.2022.3207346.

[15] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "Deep Reinforcement Learning: A Brief Survey," *IEEE Signal Process Mag*, vol. 34, no. 6, pp. 26–38, 2017, doi: 10.1109/MSP.2017.2743240.

[16] B. and P. M. and Š. F. Švaco Marko and Jerbić, "A Reinforcement Learning Based Algorithm for Robot Action Planning," in *Advances in Service and Industrial Robotics*, P. N. and M. V. C. Aspragathos Nikos A. and Koustoumpardis, Ed., Cham: Springer International Publishing, 2019, pp. 493–503. doi: https://doi.org/10.1007/978-3-030-00232-9_52.

[17] W. D. Smart, and L. Pack Kaelbling, "Effective reinforcement learning for mobile robots," in *Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No.02CH37292)*, 2002, pp. 3404–3410 vol.4. doi: 10.1109/ROBOT.2002.1014237.

[18] Z. Cui, and Y. Wang, "UAV Path Planning Based on Multi-Layer Reinforcement Learning Technique," *IEEE Access*, vol. 9, pp. 59486–59497, 2021, doi: 10.1109/ACCESS.2021.3073704.

[19] A. Iskandar, H. M. Rostum, and B. Kovács, "Using Deep Reinforcement Learning to Solve a Navigation Problem for a Swarm Robotics System," in *2023 24th International Carpathian Control Conference (ICCC)*, 2023, pp. 185–189. doi: 10.1109/ICCC57093.2023.10178888.

[20] Y. P. Pane, S. P. Nageshrao, J. Kober, and R. Babuška, "Reinforcement learning based compensation methods for robot manipulators," *Eng Appl Artif Intell*, vol. 78, pp. 236–247, 2019, doi: https://doi.org/10.1016/j.engappai.2018.11.006.

[21] H. Oliff, Y. Liu, M. Kumar, M. Williams, and M. Ryan, "Reinforcement learning for facilitating human-robot-interaction in manufacturing," *J Manuf Syst*, vol. 56, pp. 326–340, 2020, doi: https://doi.org/10.1016/j.jmsy.2020.06.018.

[22] S. Pateria, B. Subagdja, A. Tan, and C. Quek, "Hierarchical Reinforcement Learning: A Comprehensive Survey," *ACM Comput. Surv.*, vol. 54, no. 5, Jun. 2021, doi: 10.1145/3453160.

[23] M. M. Botvinick, "Hierarchical reinforcement learning and decision making," *Curr Opin Neurobiol*, vol. 22, no. 6, pp. 956–962, 2012, doi: https://doi.org/10.1016/j.conb.2012.05.008.

[24] B. IANCU, "Evaluating Google Speech-to-Text API's Performance for Romanian e-Learning Resources," *Informatica Economica*, vol. 23, no. 1/2019, pp. 17–25, Mar. 2019, doi: 10.12948/issn14531305/23.1.2019.02.

[25] S. Chitta, I. Sucan, and S. Cousins, "Moveit![ROS topics]," *IEEE Robotics & Automation Magazine - IEEE ROBOT AUTOMAT*, vol. 19, pp. 18–19, Dec. 2012, doi: 10.1109/MRA.2011.2181749.

[26] P. Ekman, "An argument for basic emotions," *Cogn Emot*, vol. 6, no. 3–4, pp. 169–200, 1992, doi: 10.1080/02699939208411068.