Research Statement

Min Lee
School of Computing and Information Systems,
Singapore Management University
Email: mhlee@smu.edu.sg
December, 2024

Recent advances in artificial intelligence (AI) have made it increasingly applicable to address challenges in every facet of our lives. Yet, deploying AI systems remains challenging and has received concerns about their harms.

Human-AI/Robotic Collaborative Systems

For an effective human-AI/robot collaboration, we have iteratively engaged with therapists and post-stroke patients to design, develop, and evaluate two human-AI collaborative systems to improve the practices of physical stroke rehabilitation: 1) an AI-based decision support system for therapists [1] and 2) a robotic exercise coach for post-stroke patients [2].

Interactive Hybrid Intelligence Approach

For transparent and personalized interactions with users, we have developed a hybrid intelligence approach [3]. Given an annotated dataset on a task (e.g. assessing patient's quality of motion), this approach automatically identifies salient features of a task using reinforcement learning [3]. With identified salient features, this approach supports transparent interactions with users (e.g. summarized patient-specific analysis for therapists [1] and corrective feedback for patients [2]). In addition, this approach integrates a machine learning model with a rule-based model as a hybrid model to accommodate user inputs (e.g. feature relevance, exercise motions) and tune a system for personalized interactions [1,2].

Human-AI Collaborative Decision Making

The idea of decision support systems to provide insights into patient diagnosis and treatment has attracted for decades. However, the adoption of these systems has failed in practice, and the gains of these systems remain unclear.

We have created a human-AI collaborative decision-making (HACDM) system that leverages a hybrid intelligence approach to generate summarized analysis on patient's quality of motion and accommodate therapist's feedback (e.g. feature relevance and relabelling) for personalized rehabilitation assessment [2,4].

To leverage the analytic strengths of AI/ML and the expert knowledge of therapists, the HAICDM integrates an AI/ML model with a rule-based model into a hybrid model (HM) for assessing the patient's quality of motion [3]. The system also automatically identifies salient frames [8] and salient features of assessment [3] to

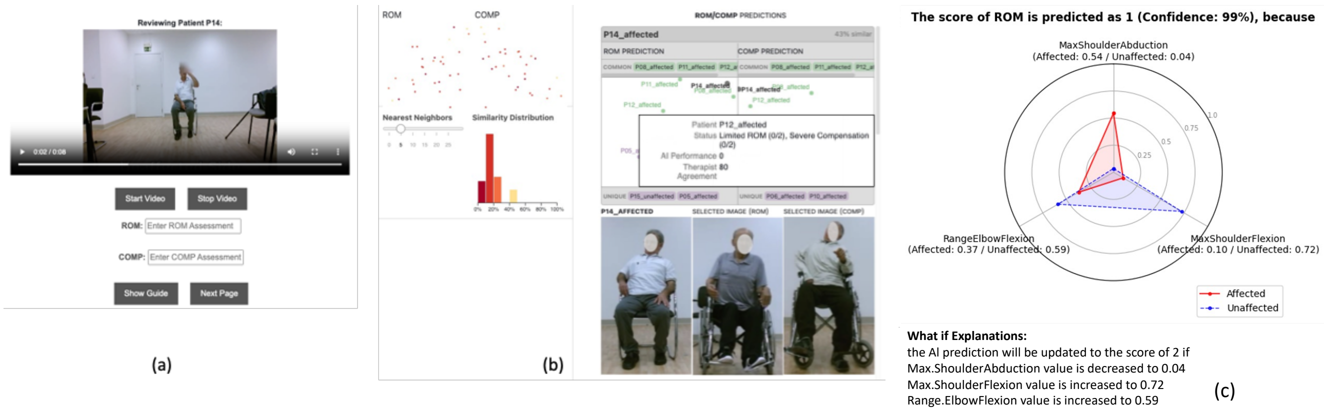

Figure 1. Interface of *HAICDM* that presents (a) the video of the patient's exercises, (b) interactive example explanations to assist the user's onboarding with AI, and (c) patient-specific analysis with the three most important features between unaffected and affected sides

generate a patient-specific analysis along with counterfactual explanations for rehabilitation assessment (Figure 1c) [5]

To improve an health professionals' onboarding process with AI, we leveraged the existing guidelines of human-AI interaction to create onboarding materials for the system and conducted interviews with 12 therapists and 4 students in medicine and health to collect their feedback on an onboarding process with AI [9]. Drawing on the insights from the interviews that therapists desire to quickly review examples along with the AI performance, we developed interactive example-based explanations of the HAICDM (Figure 1b) for onboarding with AI [9].

The evaluation study with 16 domain experts and 15 novices showed that our interactive example-based explanations assisted participants to have a 9.3% higher ratio of changing their decisions into `right' than those without interactive example-based explanations [9]. Interactive example-based explanations could assist users in building mental models on AI and onboard with it before algorithmic decision support.

In addition, studies with therapists show that the HACDM system empowers therapists to have a richer understanding of patient's performance and achieve significantly higher agreement on assessment (0.71 F1-score) than the traditional system (0.66 F1-score, $p < 0.01$) [2]. After reviewing the analysis of the HACDM, therapists can understand the capabilities of the HACDM and provide feedback to significantly improve its performance to replicate the therapist's assessment from 0.83 to 0.91 F1-score ($p < 0.01$) [2].

These results show that both domain experts and an AI system can learn from each other's strengths over interactions and generate hybrid intelligence on a complex decision-making task with improved accuracy.

Human-Robot Collaborative Physical Therapy

My research also explores how an interactive robotic coach can collaborate with post-stroke patients to improve their engagement in physical rehabilitation therapy [2]. This system also applies our hybrid intelligence approach to automatically monitor and guide rehabilitation exercises of patients through social interactions (e.g. verbal, visual, and gesture-based feedback in Figure 2). In contrast to prior work that utilizes pre-defined, generic feedback, this system tunes with a new patient's motions and generates transparent, personalized corrective feedback.

Figure 2: An interactive robot coach monitors a patient's exercise and provides personalized feedback

A study showed that our system can be tuned with patient's exercises to significantly improve its performance to provide personalized assessment and feedback from 0.74 to 0.82 F1-score ($p < 0.01$) [2]. In addition, a real-world evaluation with 10 participants showed that our system can adapt to new participants and achieved a 0.81 F1-score to assess their exercises, which is comparable to experts' performance and better than fine-tuning a feed-forward neural network (0.74 F1-score) [2].

Future Work

I seek to design and develop approaches that enable people to responsibly collaborate with AI/robotic systems and evaluate their impact on a task of diverse application domains (e.g. healthcare [5,6] and public services [7]).

- Explore an approach to make AI more explainable [8], fair, and interactive
- Examine factors that affect the uptake of AI and its effects on stakeholders

Selected Publications and Outputs

1. **Min Hun Lee**, Daniel P. Siewiorek, Asim Smailagic, Alexandre Bernardino, and Sergi Bermúdez Badia. 2021. "A Human-AI Collaborative Approach for Clinical Decision Making on Rehabilitation Assessment". In Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI '21).
2. **Min Hun Lee**, Daniel P. Siewiorek, Asim Smailagic, Alexandre Bernardino, and Sergi Bermúdez Badia. 2020. "Design, development, and evaluation of an interactive personalized social robot to monitor and coach post-stroke rehabilitation exercises." User Modeling and User-Adapted Interaction 33.2 (2023): 545-569.
3. **Min Hun Lee**, Daniel P. Siewiorek, Asim Smailagic, Alexandre Bernardino, and Sergi Bermúdez Badia. 2020. "Interactive hybrid approach to combine machine and human intelligence for personalized rehabilitation assessment." Proceedings of the ACM Conference on Health, Inference, and Learning. 2020.
4. **Min Hun Lee**, Daniel P. Siewiorek, Asim Smailagic, Alexandre Bernardino, and Sergi Bermúdez Badia. 2022. "Towards Efficient Annotations for a Human-AI Collaborative, Clinical Decision Support System: A Case Study on Physical Stroke

Rehabilitation Assessment”. In 27th International Conference on Intelligent User Interfaces (IUI '22).
5. **Min Hun Lee** and Chong Jun Chew. 2023. "Understanding the Effect of Counterfactual Explanations on Trust and Reliance on AI for Human-AI Collaborative Clinical Decision Making." Proceedings of the ACM on Human-Computer Interaction 7.CSCW2 (2023).
6. **Min Hun Lee**, Daniel P. Siewiorek, and Alexandre Bernardino. 2023. "Designing a Human-Centered Intelligent System to Monitor & Explain Abnormal Patterns of Older Adults." Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility. 2023.
7. Logan Stapleton, **Min Hun Lee**, Diana Qing, Marya Wright, Alexandra Chouldechova, Ken Holstein, Zhiwei Steven Wu, and Haiyi Zhu. 2022. “Imagining new futures beyond predictive systems in child welfare: A qualitative study with impacted stakeholders.” In 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22)
8. **Min Hun Lee** and Yi Jing Choy. 2023. “Exploring a Gradient-based Explainable AI Technique for Time-Series Data: A Case Study of Assessing Stroke Rehabilitation Exercises". In ICLR 2023 Workshop on Time Series Representation Learning for Health.
9. **Min Hun Lee** R.B.X. Ng, S.X.Y. Choo, and S.D. Thilarajah. 2024. “Interactive Example-based Explanations to Improve Health Professionals’ Onboarding with AI for Human-AI Collaborative Decision Making". In 27th European Conference on Artificial Intelligence (ECAI 2024), 2024.