2024 International Conference on Advanced Robotics and Intelligent Systems (ARIS 2024)
National Taiwan University of Science and Technology, Taipei 10607, Taiwan, August 22- August 24, 2024

Improving Human-Robot Collaboration Efficiency: AE-LSTM Model for Continuous Gesture Recognition in Precision Assembly

Ashiq hussain Teeli, Chin-Te Lin

Department of Mechanical Engineering, National Central University

***Abstract*— This research paper aims to solve the continuous gesture recognition issues when human and robot work in a shared workspace. Previous work proposed a setup integrating a ZED 2 camera for action recognition and a CoolSo bracelet for gesture recognition, aiming to enable collaborative work between humans and robotic arms. While the deep learning model developed for action recognition performed well, the continuous gesture recognition accuracy was suboptimal. The proposed system reduced preprocessing time, and the deep learning model is evaluated on a dataset of continuous gestures from the precision assembly tasks, demonstrating significant improvements in continuous gesture recognition accuracy compared to the previous deep learning model. The results demonstrate the feasibility and effectiveness of the proposed approach in identifying continuous human hand behaviors in real time. The approach can improve the performance of human and robot collaboration.**

I. Introduction

In this study, we propose a deep learning model to increase the accuracy of continuous gesture recognition in real-time for human-robot collaboration in a shared workspace. This system combines CNN and RNN models for continuous action recognition. The proposed approach is expected to improve the performance and safety of human-robot collaboration in various industrial and manufacturing settings.

II. Methods and Results

We developed a deep learning model using an Auto Encoder and Long Short-Term Memory Network (AE-LSTM) architecture to increase the real-time accuracy of continuous human action recognition. This system leverages Auto Encoders for dimensionality reduction and feature extraction, coupled with Long Short-Term Memory networks to predict and classify actions. The AE-LSTM model was evaluated on a dataset collected from the CoolSo bracelet, which recorded continuous gestures from precision assembly tasks. The model demonstrates significant improvements in continuous gesture recognition accuracy compared to the previous model, while also reducing preprocessing time, thus enhancing efficiency. These improvements in detecting human actions aim to facilitate better human-robot collaboration.

To evaluate the system's performance, we created a scenario where datasets from three individuals were collected: two technical users and one non-technical user. Seven actions were recorded during the assembly of a Lego car. The actions included noise, feeding, point left, point right, installing beam, installing bearing, and installing layers. 300 CSV files were collected for each gesture, resulting in a comprehensive dataset for model training and testing. To test the efficacy of the model, we used the non-technical user's dataset as the test set and the technical users' datasets as the training set. This approach allowed us to evaluate the model's accuracy in continuous action recognition across different user skill levels. The model demonstrates significant improvements in continuous gesture recognition accuracy compared to the previous model, achieving a best result of 75%. Additionally, while increasing the sequence length led to slight improvements in accuracy, it also increased the prediction time, impacting real-time performance. Therefore, a sequence length of 30 was determined to be optimal for balancing accuracy and real-time responsiveness, enhancing overall efficienc.

This research was funded by National Science and Technology Council [MOST 111-2218-E-008-007, 112-2218 -E-008 -009 -]

Figures

Figure 1. *Proposed Model Archtecture*

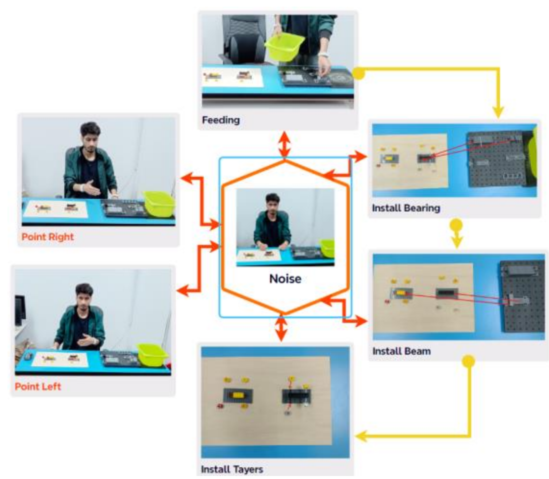

Figure 2. Lego car assembly flow chart

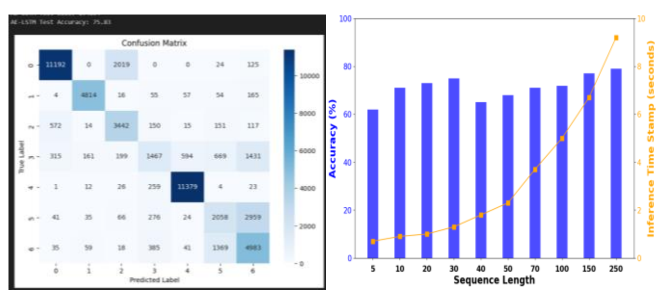

Figure 3. Results

References

[1] Bo zhang, L. Z., Mojun Wu, and Yan Wang (2021). "Dynamic Gesture Recognition Based on RF sensor and AE-LSTM Neural Network."

[2] Harshala Gammulle, D. A.-A., Simon Denman, Lachlan Tychsen-Smith, Lars Petersson, Clinton Fookes (2022). "Continuous Human Action Recognition for Human-Machine Interaction: A Review."