

Keywords: 3D Pose Detection, Video Synchronization, Physical Rehabilitation, GPT-4, Large Language Models

Resume:

In my thesis, "LLM-Guided Rehabilitation: Detecting and Correcting Rehabilitation Errors with LLMs", I explore how Large Language Models (LLMs) and Vision Language Models (VLMs) can assist in physical rehabilitation by identifying and correcting posture errors during exercises. This work addresses key challenges such as the increasing demand for rehabilitation, the shortage of professionals, and the limitations of traditional therapy, especially for patients needing at-home solutions.

I divided the process into four main components:

1. 3D Pose Detection: I used advanced models like MeTRAbs and HybrIK-X to estimate joint positions and create 3D SMPL skeleton representations of patients and therapists.
2. Video Synchronization: Using Dynamic Time Warping (DTW), I aligned exercise videos of therapists and patients to identify frames with the most significant posture deviations.
3. Prompt Engineering for GPT-4: I crafted detailed prompts to provide GPT-4 with contextual information about exercises and skeleton data, enabling it to describe errors and suggest corrections. I experimented with various input formats, including images, skeleton data, and multi-camera perspectives.
4. Evaluation: I assessed GPT-4's ability to detect errors and provide accurate feedback, finding that its performance improved significantly when given precise inputs and focus-of-attention mechanisms.

The results showed that GPT-4 achieved up to 60% accuracy in error detection on the dataset. I found that multi-camera setups and well-defined prompts were crucial for enhancing its effectiveness. While these systems cannot replace human therapists, they have the potential to assist in making therapy more accessible, especially for at-home use. Through this work, I aim to demonstrate how AI tools can complement human expertise in rehabilitation and pave the way for innovative solutions in the field.