Tool Action Recognition Using CNN-LSTM model and Explainable AI

Monika Chavan, Shashank, Apoorva Vaidya, Taimoor Hussain

1. Task Overview

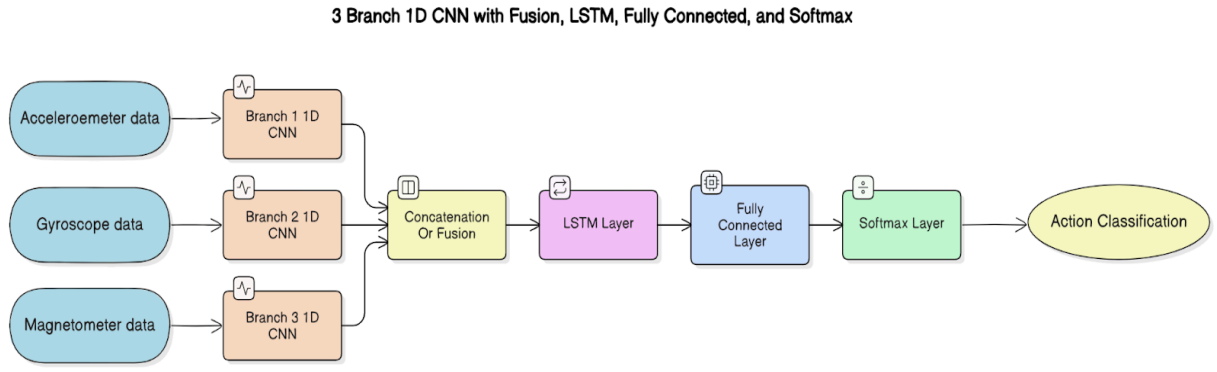
We aim to perform **Tool Action Recognition** using the Tool Tracking dataset, classifying different actions performed by the three tools based on their temporal behavior.

2. Baseline DL Model

We will use a **Multi-branch 1D CNN with Late Fusion LSTM** as our baseline deep learning model.

- The CNN captures local temporal features.
- The LSTM models long-term dependencies.

We'll begin with a simpler LSTM-only model, then improve performance using the multi-branch CNN for better feature representation.



3. Advanced Method

We plan to implement **Explainable AI (XAI)** techniques—specifically **SHAP** and **LIME**—to interpret model decisions and provide real-time feedback.

Alternative Plan: Depending on progress, we may instead explore **semi-supervised learning** or **time-aware modeling** to further improve accuracy.

4. Task Distribution

Here is a tentative distribution of subtasks among the team members:

- **Baseline LSTM**: Shashank and Monika Radhakisan Chavan
- **Multi-branch CNN + LSTM**: Apporva Vaidya
- **Explainable AI (SHAP & LIME)**: Taimoor Hussain