**🧩 Project: Hand Action Recognition — Next Steps**

**🔹 Phase 1: Organize Raw Data**

1. **Create a directory** named raw\_videos/ and place all videos received from Hamza inside.
2. **Rename videos** to clearly indicate the operation they represent (e.g., screw1.mp4, click1.mp4, etc.).

**🔹 Phase 2: Segment Long Videos into Short Clips**

1. **Split each 2–3 minute video** into multiple short clips of **3 to 7 seconds** each.
   * Tools:
     + **Recommended**: [LosslessCut](https://github.com/mifi/lossless-cut) (GUI-based tool).
     + **Alternative**: Use a Python script with moviepy or ffmpeg (ask for help if needed).
2. **Save the short clips** in a new folder named video\_segments/.

**🔹 Phase 3: Preprocess Videos**

1. **Remove audio** from each short video clip:
   * Use a Python script (provided upon request).
   * Output should be saved in a folder named clean\_videos/.

**🔹 Phase 4: Frame Extraction (V\_50)**

1. **Create a Python function** to:
   * Load each video from clean\_videos/
   * Extract **50 equally spaced frames**
   * Save frames as images in a folder structure like:  
     V\_50/operation\_name/video\_name/frame\_01.jpg ... frame\_50.jpg
2. Name this dataset **V\_50**.

**🔹 Phase 5: Tool Image Collection**

1. **Collect tool images** used in the videos:
   * Either extract representative frames showing the tool clearly
   * Or download relevant images from the web
2. Save them in a folder named tool\_images/.

**🔹 Phase 6: Research (Model Selection)**

1. **Search for the top 5 state-of-the-art models** for egocentric action recognition. For each model, write:

* Model name
* Paper title + year
* Key features
* A short summary of how it works
* Link to the GitHub repo (if available)

**🔹 Phase 7: Final Report**

1. Write the final report summarizing:

* Dataset creation process (including segmentation and V\_50)
* Preprocessing steps
* Tool image collection
* Description of the 5 models
* Any challenges faced or next steps