Hand Action Recognition

# Overview

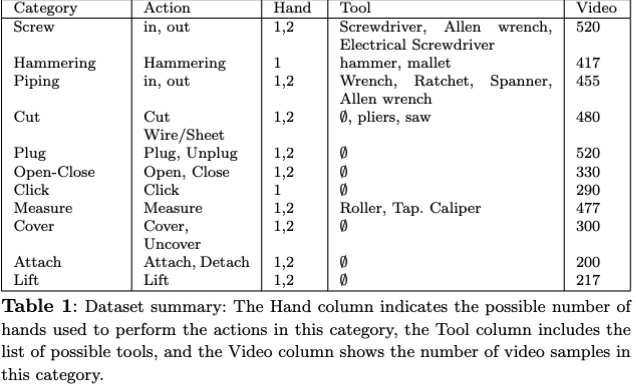
Hand action recognition aims to identify hand gestures and It has numerous practical applications, such as sign language recognition, virtual reality interactions, and assistive technology for people with disabilities. Deep learning has revolutionized computer vision by significantly improving performance in various tasks such as image classification, object detection, segmentation, and image generation. Its success is mainly due to its ability to automatically learn spatial hierarchies and complex patterns from images without needing manual feature extraction.

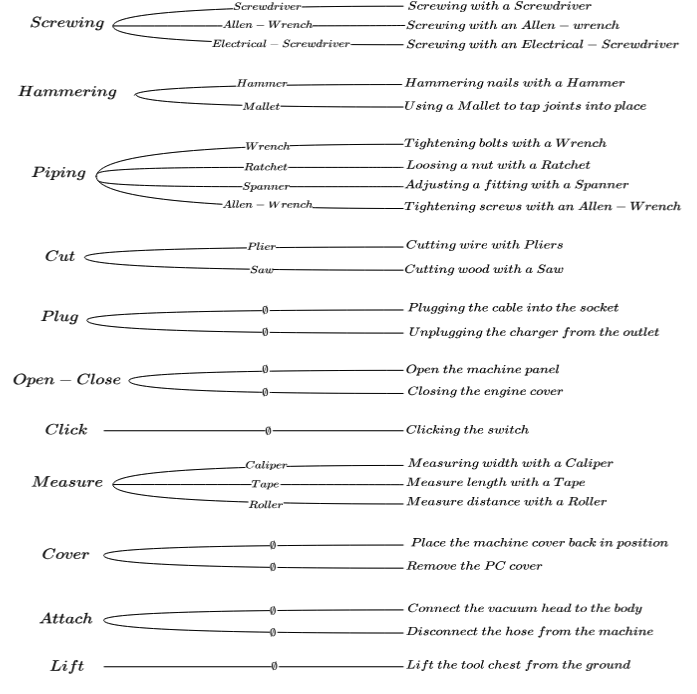
This project explores utilizing deep learning to recognize manual operations from an egocentric view. Training datasets are the core of any supervised deep-learning approach. Therefore, you will build the Maintenance Datasets, which include video segments of manual operations.

# Dataset

* use the following camera settings:

1. Lens: Wide
2. Resolution FULL HD(1920x1080)
3. Hyper-smooth On

* Dataset Summary:
* Category-to-multiple operation mapping. Each general action category (e.g., 'Screwing,' 'Cutting') encompasses multiple operations defined by the tools used. This one-to-many relationship emphasizes the flexibility within each category, as different tools enable various approaches to achieve the same overall task.



For this project, a custom dataset was created by recording various manual

operations using a GoPro camera set to a resolution of (1920X1080) pixels with a

widescreen aspect ratio. The operations recorded were chosen to cover a range of

common manual tasks that are typically performed in industrial or workshop settings.

The following operations were included in the dataset:

Attach & Detach

**Attach**, this operation involved joining two or more parts together, such as

connecting the head of a vacuum to the vacuum’s body. The recordings

demonstrated various techniques for aligning and connecting components.

**Detach**, this operation involved separating two or more connected parts, such as

removing the head from the vacuum body. The recordings demonstrated various

techniques for disassembling components.

**Click**,this operation involved pressing different types of buttons or switches (e.g.,

mouse click, remote controller button, etc…). The recordings included various button

types to help the model learn to recognize small, precise movements

Cover & Uncover

**Cover,** this operation involved placing or laying a cover over an object to enclose or

protect it. The recordings focused on different techniques for positioning and

securing the cover to ensure it adequately covers the object.

**Uncover,** this operation involved removing a cover from an object to reveal or

access its contents. The recordings demonstrated various methods for lifting or

taking off the cover to expose what was underneath.**Cutting,** this operation involved using the scissors as a tool to slice through a

material or object to separate it into parts or remove a portion.

**Hammering**, this operation involved using a hammer to strike an object with force to

change its shape, move it. The recordings captured different hammering techniques

to show variations in force, speed, and angles

**Lift**, this operation involved raising objects of various weights and sizes from a lower

to a higher position.

**Measuring**, this operation involved using a tape measure to determine the length,

width, or height of various objects. The recordings highlighted the process of

extending the tape measure, positioning it accurately, and reading measurements to

ensure precise results.

Open & Close

**Open**, this operation involved moving an object or mechanism to create an

accessible or exposed state, such as opening a door or container.

**Close**, this operation involved moving an object or mechanism to secure or cover it,

such as closing a door or container.

Tug; plug & unplug

**Plugging**, this operation involved connecting a plug into an electrical socket or port.

The recordings focused on the techniques used to securely insert the plug, ensuring

a proper connection.

**Unplugging**, this operation involved removing a plug from an electrical socket or

port. The recordings captured the methods used to safely disconnect the plug.

Push & Pull

**Screw; Screw-in & Screw-outScrew In**,this operation involved using a screwdriver to turn screws clockwise to

drive them into various materials. The recordings captured different types of screws

and materials, highlighting variations in the speed required for this clockwise motion.

**Screw Out**, this operation involved using a screwdriver to turn screws

counterclockwise to remove them from various materials.

# TASKs

* Create a dataset which contains video segments (3sec-7sec) depicting an operation (The bigger the data the more score you get).
* Collect images of different tools (might be frames from the video dataset you recorder or images from the web)
* Preprocessing Step
  + Clean the dataset by removing irrelevant videos
  + Remove the audio from each video of the dataset
  + Create a frame selection function that takes the most relevant 50 frames from each video. call this mini data as V\_50
* Search for top five state-of-art models for action recognition from egocentric view
* Write the final report based on these tasks