My View is the Best View: Procedure Learning from Egocentric Videos

Siddhant Bansal, Chetan Arora, C.V. Jawahar

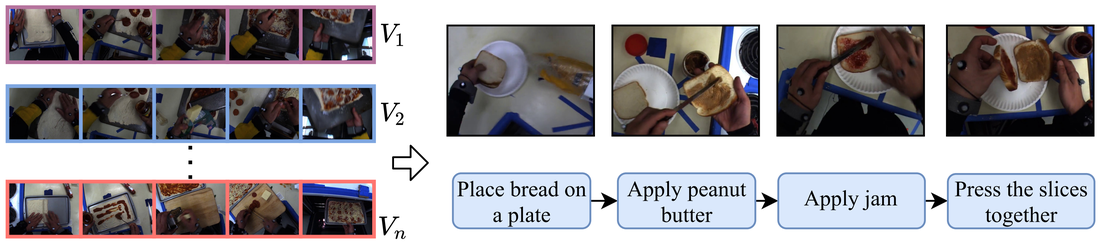
INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY HYDERABAD

CVIT

INDIAN INSTITUTE OF TECHNOLOGY DELHI

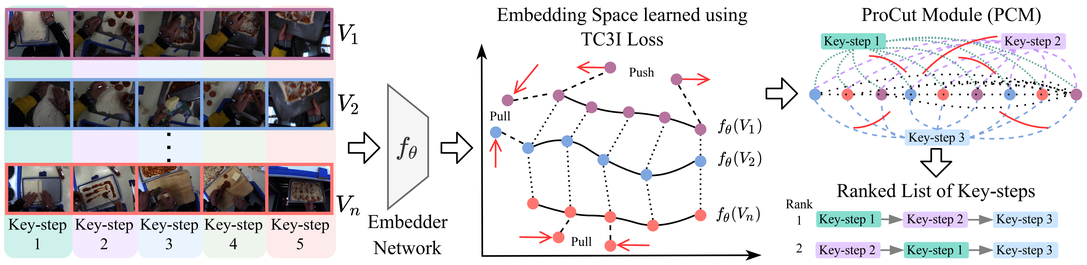
ECCV TEL AVIV 2022

What is Procedure Learning?



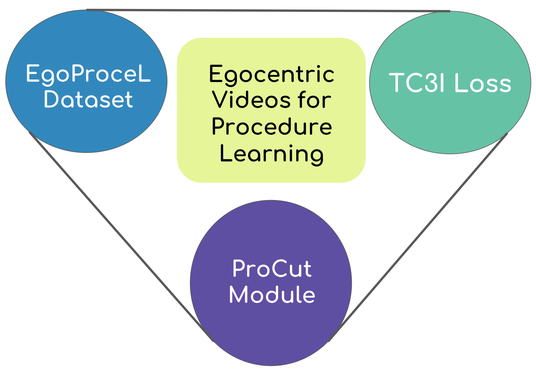
Given a few task's videos, the aim is to identify the key-steps required to perform the task.

Correspond and Cut (CnC) framework for Procedure Learning



- **TC3I loss** is used to train the Network.
- Similar embeddings are learned for the **corresponding key-steps** across multiple videos and **temporally close frames**.
- **PCM** localizes the key-steps by converting the clustering problem to a **multi-label graph cut problem**.
- Output is the key-steps and their ordering.

Major Contributions



Download EgoProceL, Code, and Models
sid2697.github.io/egoprocel

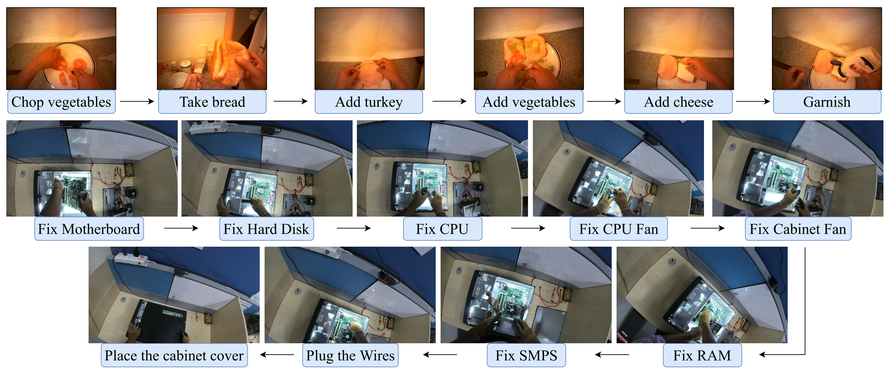
Potential Applications

Monitoring Procedures. Enabling identifying missing or wrong steps.

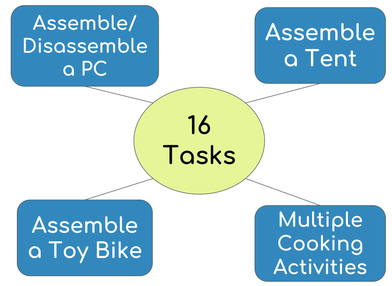
Guidance Systems. Identify the current step and show the next step to perform the task.

Automated Systems. Learn autonomously the key-steps for performing a task by observing the task being performed.

EgoProceL Dataset: Egocentric Videos for Procedure Learning

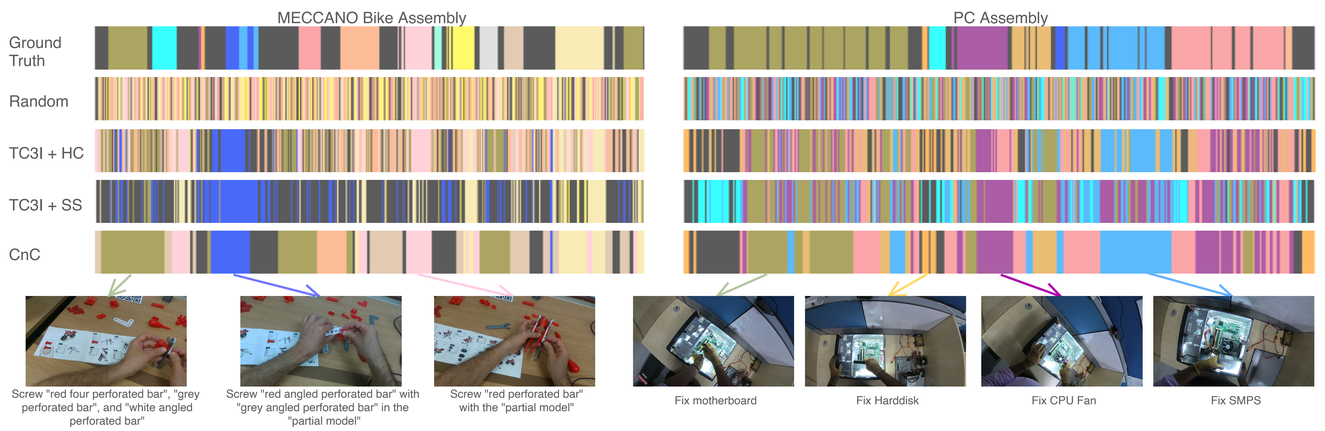


Key-step annotations in EgoProceL for making turkey sandwich and assembling a PC.



- 62 hours of videos
- 130 subjects
- 17 maximum key-steps
- 0.38 avg. foreground ratio

Results and Analysis



Qualitative Results for MECCANO and PC Assembly

	EgoProceL											
	CMU-MMAC		EGTEA G.		MECCANO		EPIC-Tents		PC Assembly		PC Disas.	
	F1	IoU	F1	IoU	F1	IoU	F1	IoU	F1	IoU	F1	IoU
Random	15.7	5.9	15.3	4.6	13.4	5.3	14.1	6.5	15.1	7.2	15.3	7.1
TC3I + HC	19.2	9.0	20.8	7.9	16.6	**8.0**	15.4	7.8	21.7	11.0	24.9	14.1
TC3I + SS	19.7	8.9	20.4	7.9	16.3	7.8	15.9	7.8	24.8	11.9	23.6	14.0
CnC	**22.7**	**11.1**	**21.7**	**9.5**	**18.1**	7.8	**17.2**	**8.3**	**25.1**	**12.8**	**27.0**	**14.8**

Qualitative Results on EgoProceL. Here CnC has the best results this highlights the effectiveness of the proposed **TC3I loss** and **PCM**.