




## Summary

- We introduce a new benchmark dataset (IPN Hand) for **hand gesture recognition** (HGR).
- IPN hand has sufficient size and challenge to evaluate complex deep learning models for **continuous and isolated HGR**.
- We evaluate our dataset for real-time HGR with **multimodal inputs**, such as RGB-seg (segmentation) and RGB-flow (optical flow).

## IPN Hand Challenges

- Continuous gestures without transition states:
 
- Natural behaviors of users' hands (non-gestures):
 
- Different real-world backgrounds:
 

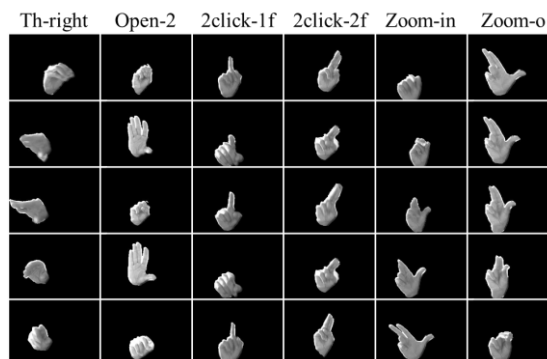
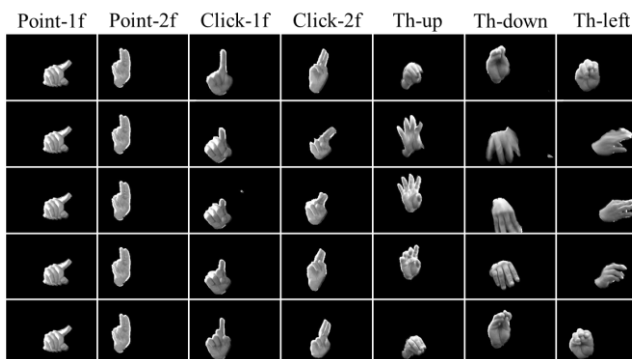
## Data collection

- 50** different subjects
- 4,218** gesture instances
- 200** long videos (RGB)
- 21.1** inst. per video
- 640x480** res. at **30** fps



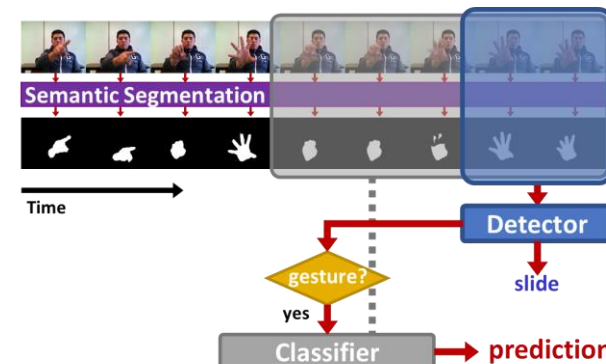
## Static and Dynamic Hand Gestures

- 13 hand gestures** for interaction with touchless screens (pointer and actions).



## Continuous Hand Gesture Recognition

- Based on a **two hierarchical 3D-CNNs** approach:



## Benchmark Evaluation

- Levenshtein accuracy** is used as evaluation metric for continuous recognition.
- Accuracy is based on the **predicted labels** in the **correct temporal order** of the detected gestures.

Model	Modality	Accuracy	Size	Time
ResNeXt-101	RGB	25.34	370 MB	30.1 ms
ResNeXt-101	RGB-Flow	42.47	375 MB	53.7 ms
ResNeXt-101	RGB-seg	39.01	386 MB	39.9 ms
Resnet-50	RGB	19.78	360 MB	20.4 ms
Resnet-50	RGB-Flow	39.47	365 MB	43.1 ms
Resnet-50	RGB-seg	33.27	376 MB	29.2 ms

## Publicly Available

- Code and models available at:**  
<https://github.com/GibranBenitez/IPN-hand>

