

AI Applications

Computer Vision & First Person Vision

Yin Li

`yin.li@wisc.edu`

University of Wisconsin, Madison

The vision story begins

“What does it mean, to see? The plain man's answer (and Aristotle's, too) would be, **to know what is where by looking.**”

-- David Marr, *Vision* (1982)

Computer vision deals with how computers can understand the contents in visual data, including images and videos.

Scene Recognition

Living Room



Slide Credit: Abhinav Gupta

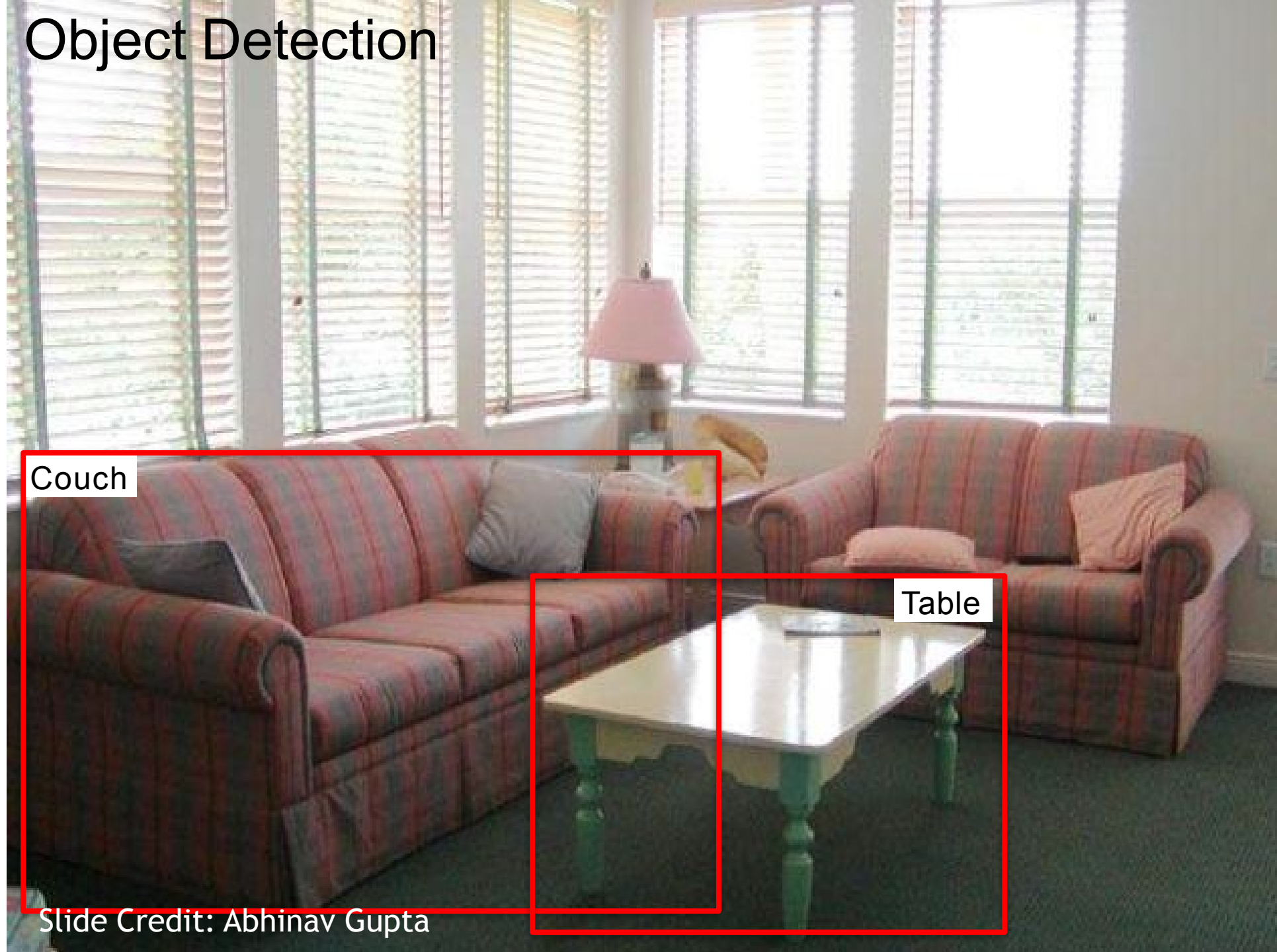
Object Recognition

Couch, Table, ...



Slide Credit: Abhinav Gupta

Object Detection

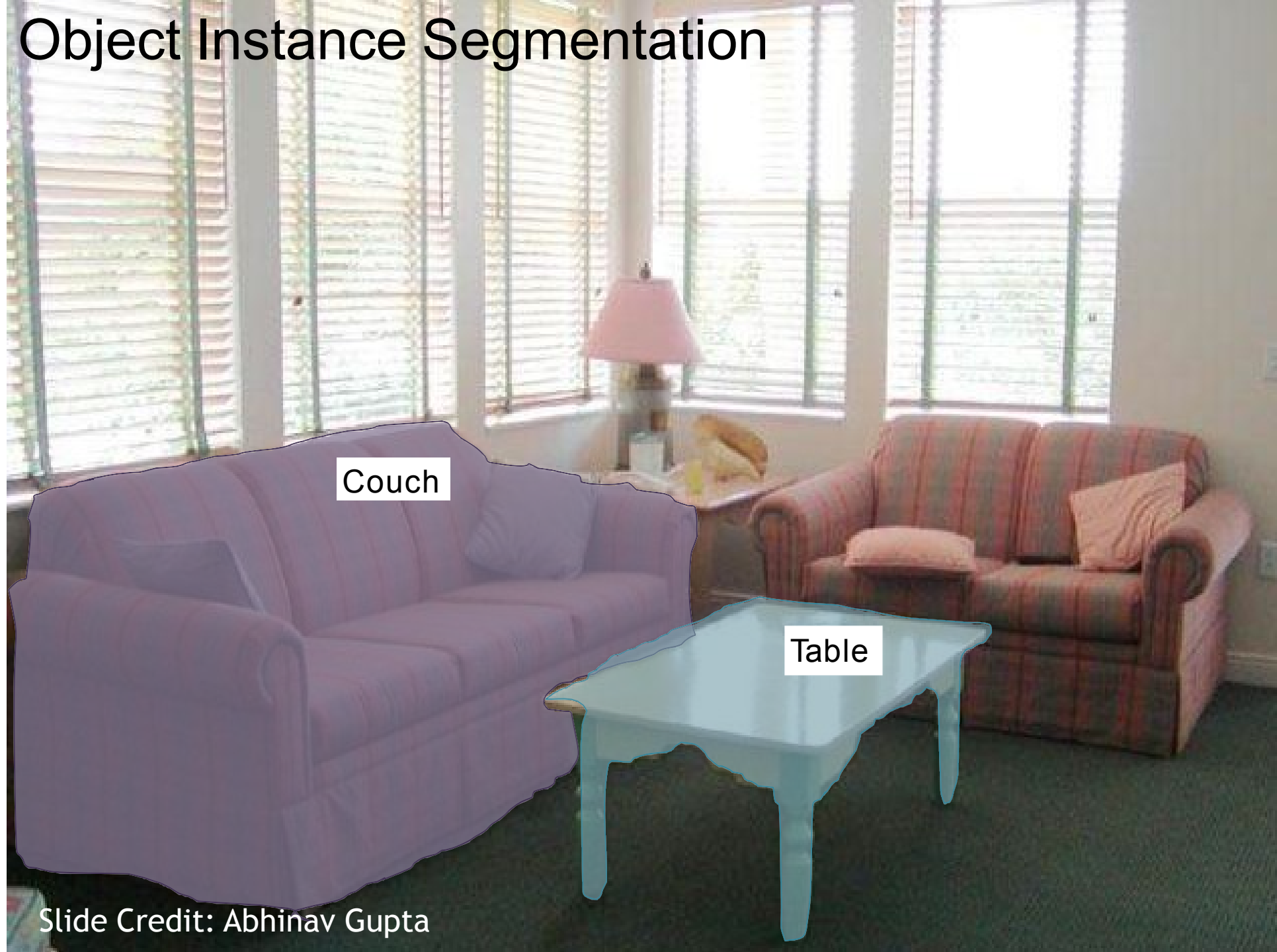


Couch

Table

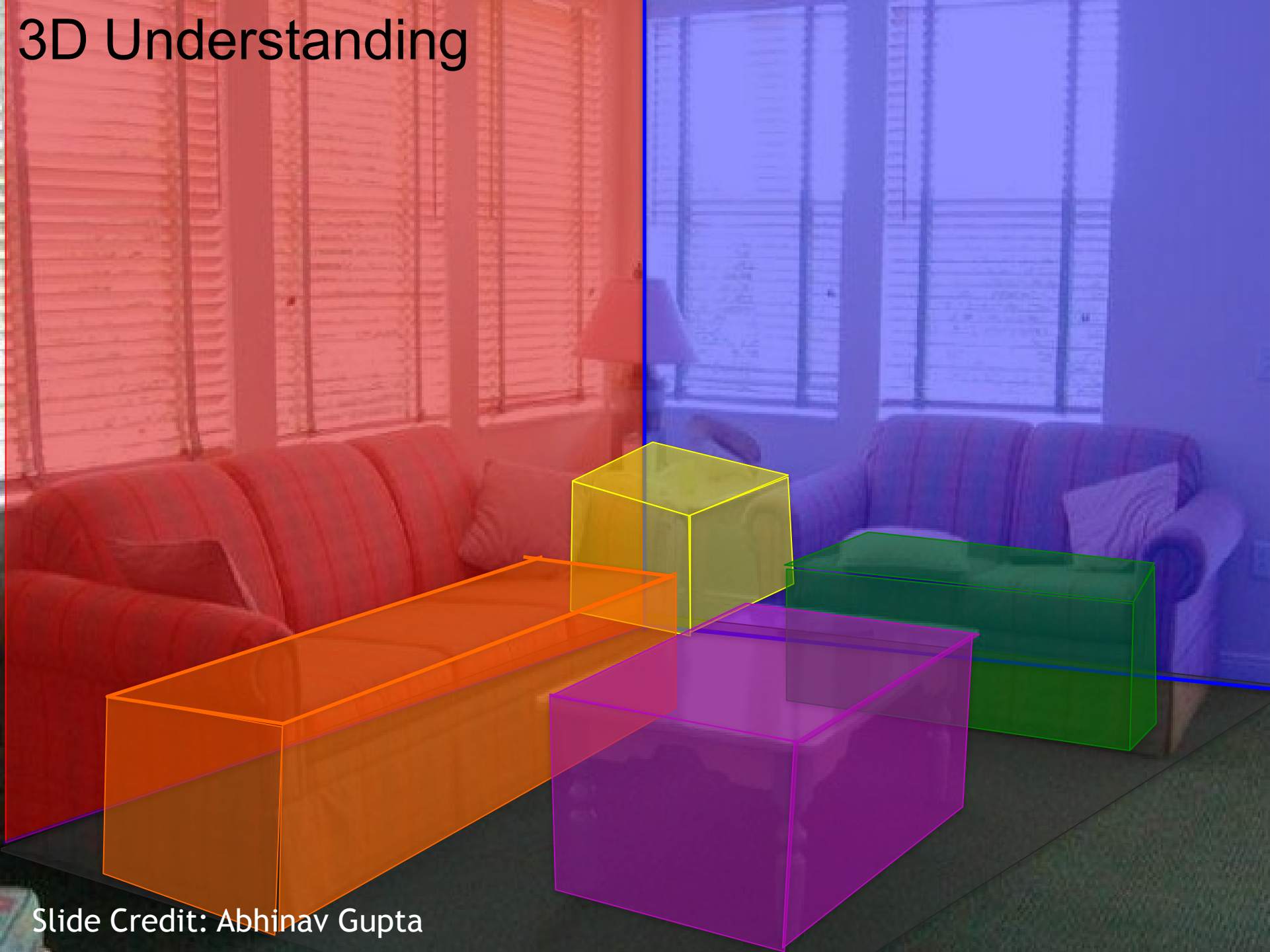
Slide Credit: Abhinav Gupta

Object Instance Segmentation



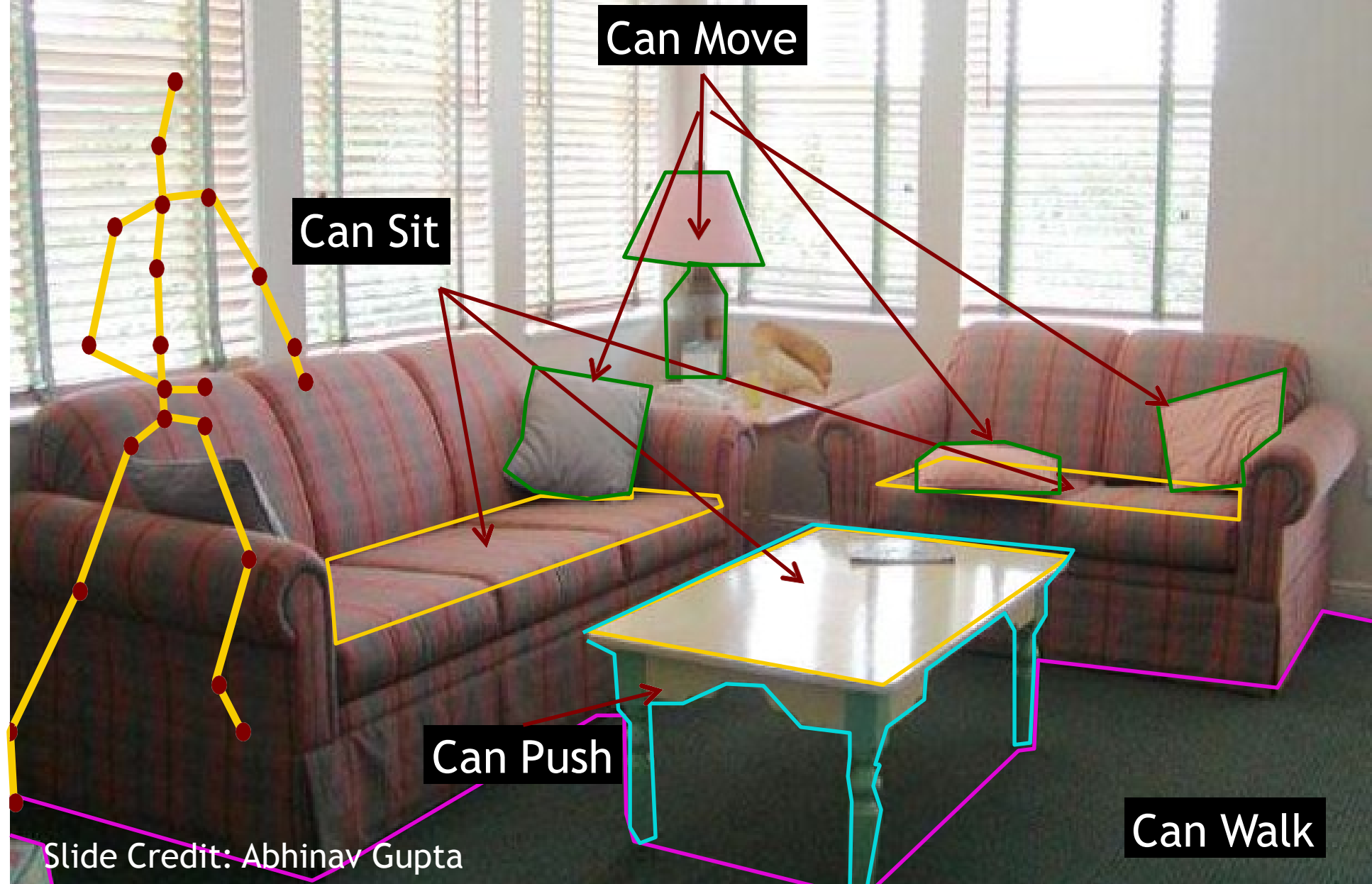
Slide Credit: Abhinav Gupta

3D Understanding



Slide Credit: Abhinav Gupta

Functional Understanding



Can Move

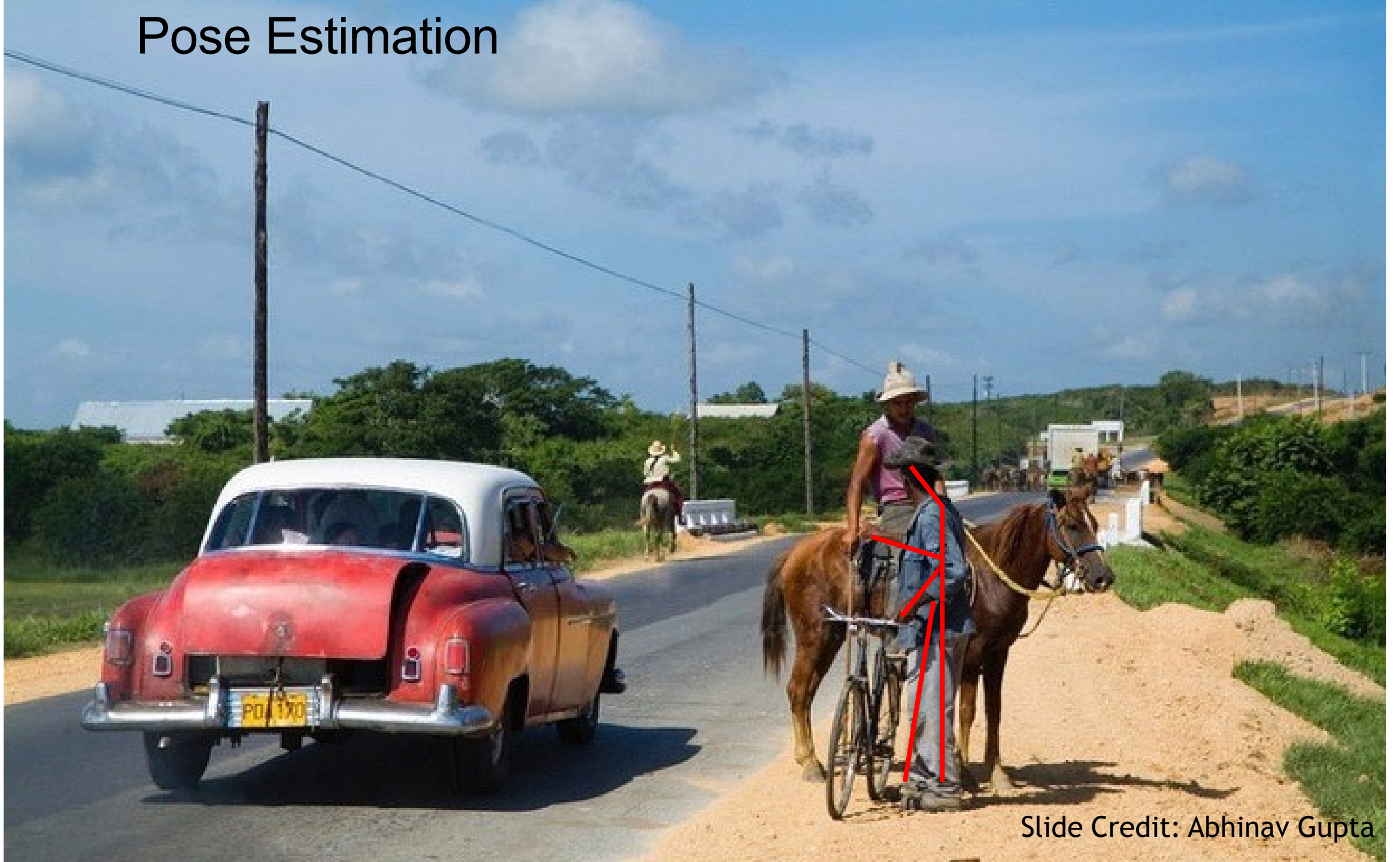
Can Sit

Can Push

Can Walk

Slide Credit: Abhinav Gupta

Pose Estimation

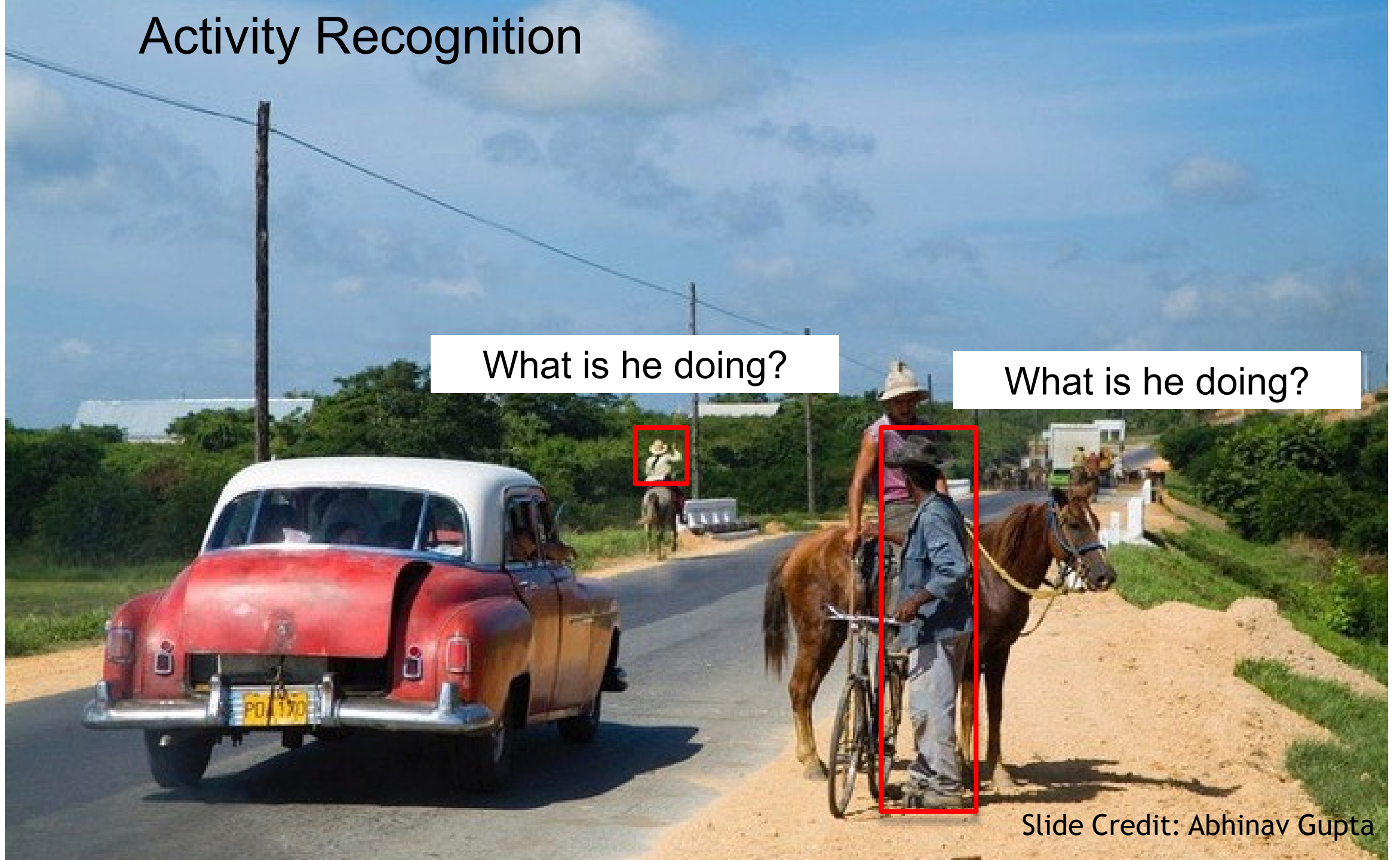


Slide Credit: Abhinav Gupta

Activity Recognition

What is he doing?

What is he doing?



Interested in computer vision?

- CS 534: Computational Photography
- CS 766: Computer Vision
- BMI 767 / CS 767: Computational Methods for Medical Image Analysis
- BMI 826 / CS 838: Learning Based Methods in Computer Vision

First Person Vision (**FPV**): *Automatic Analysis of First Person Videos*



SMI



Google Glass



HoloLens



Tobii

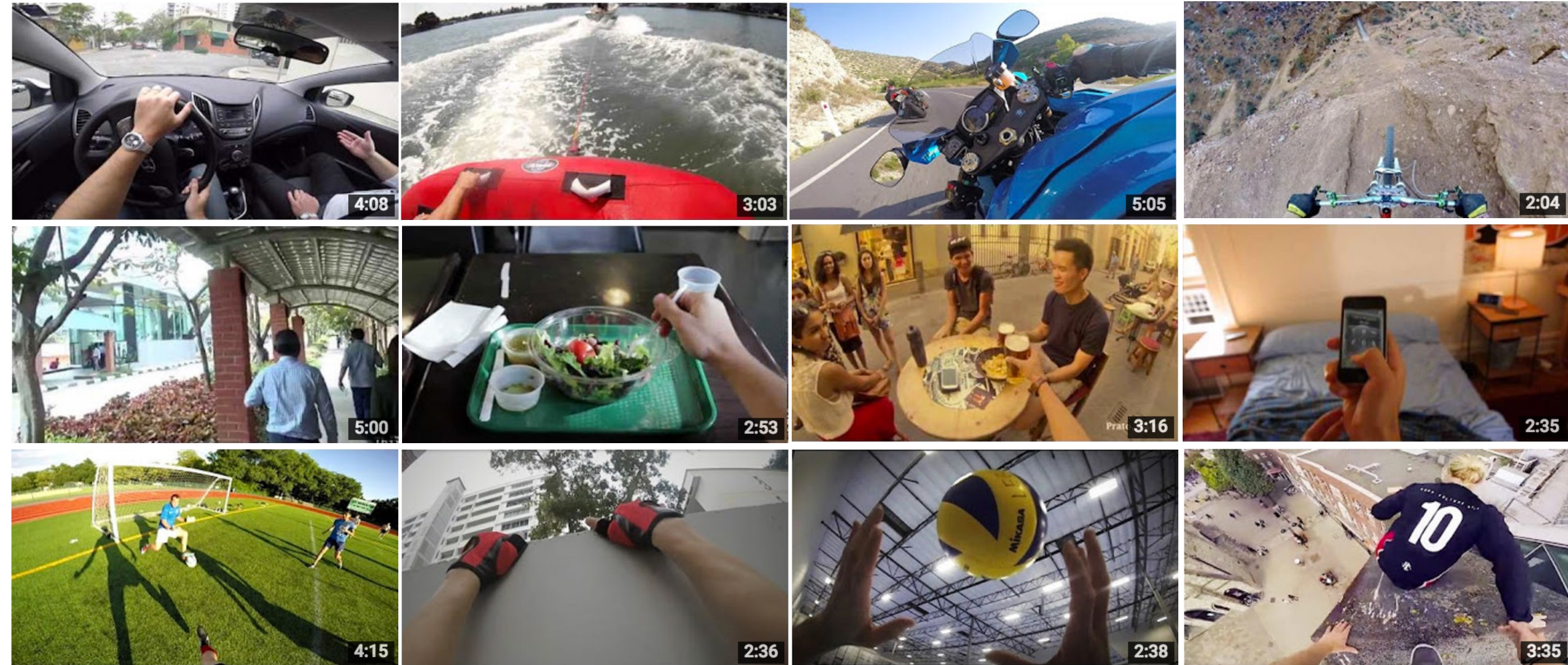


GoPro



Vuzix Blade

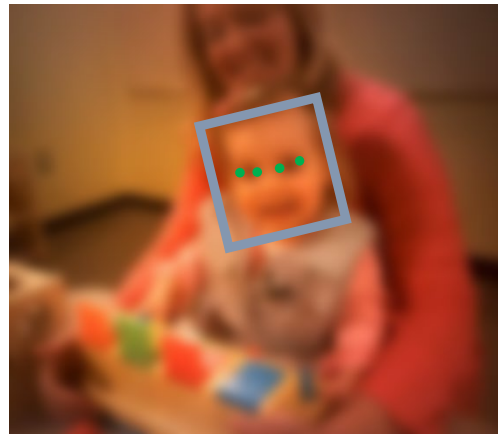
12% of Videos on Social Media are FPV



The Open World of Micro-Videos, Nguyen, Rogez, Fowlkes and Ramanan, ArXiv 2016

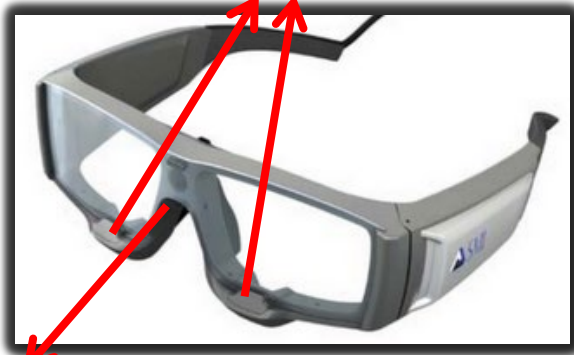
FPV: sensing our daily behavior

- What is the participant doing?
- How is the participant doing so?
- Most importantly, the health implications ...



Wearable Eye Tracking

Eye tracking cameras

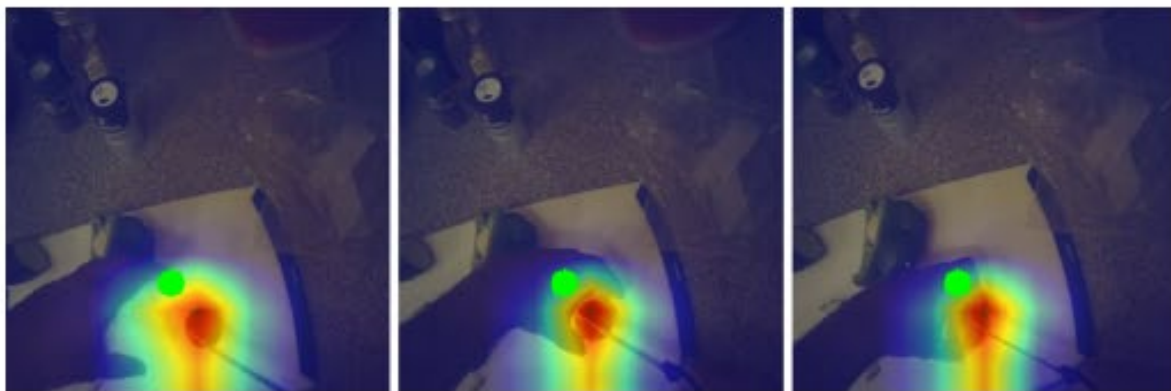


Scene camera SMI

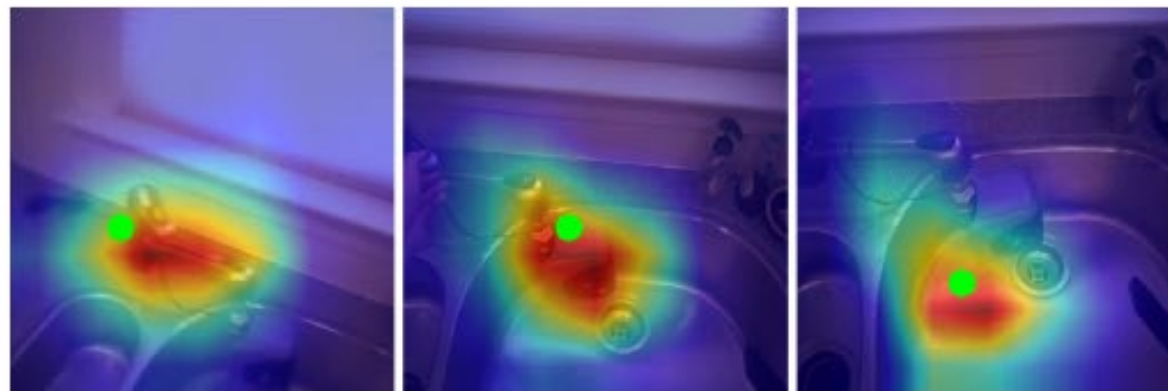


Action recognition & gaze estimation in FPV

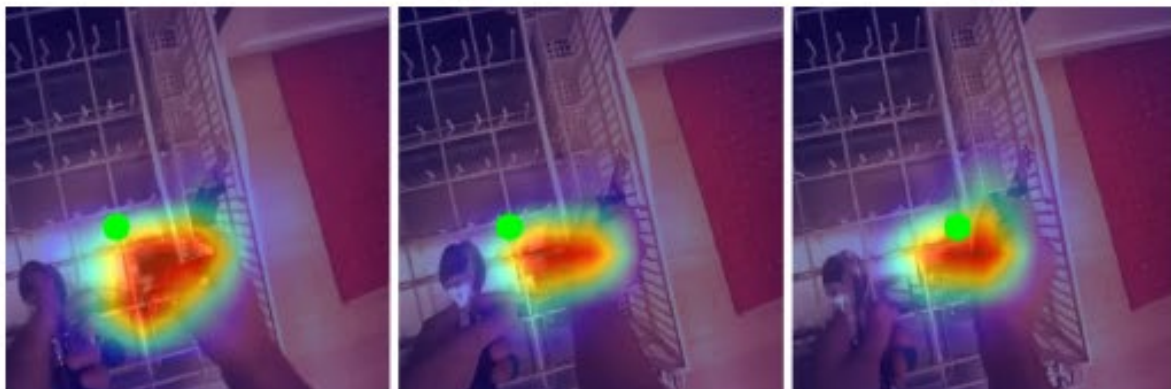
Predicted: Cut tomato GT: Cut tomato



Predicted: Turn on faucet GT: Turn on faucet



Predicted: Take eating_utensil GT: Take eating_utensil



Predicted: Inspect/Read recipe GT: Inspect/Read recipe



Action forecasting in FPV

Pred: Operate Stove GT: Operate Stove



Pred: Operate Microwave GT: Operate Microwave



Pred: Wash Coffee Cup GT: Wash Coffee Cup



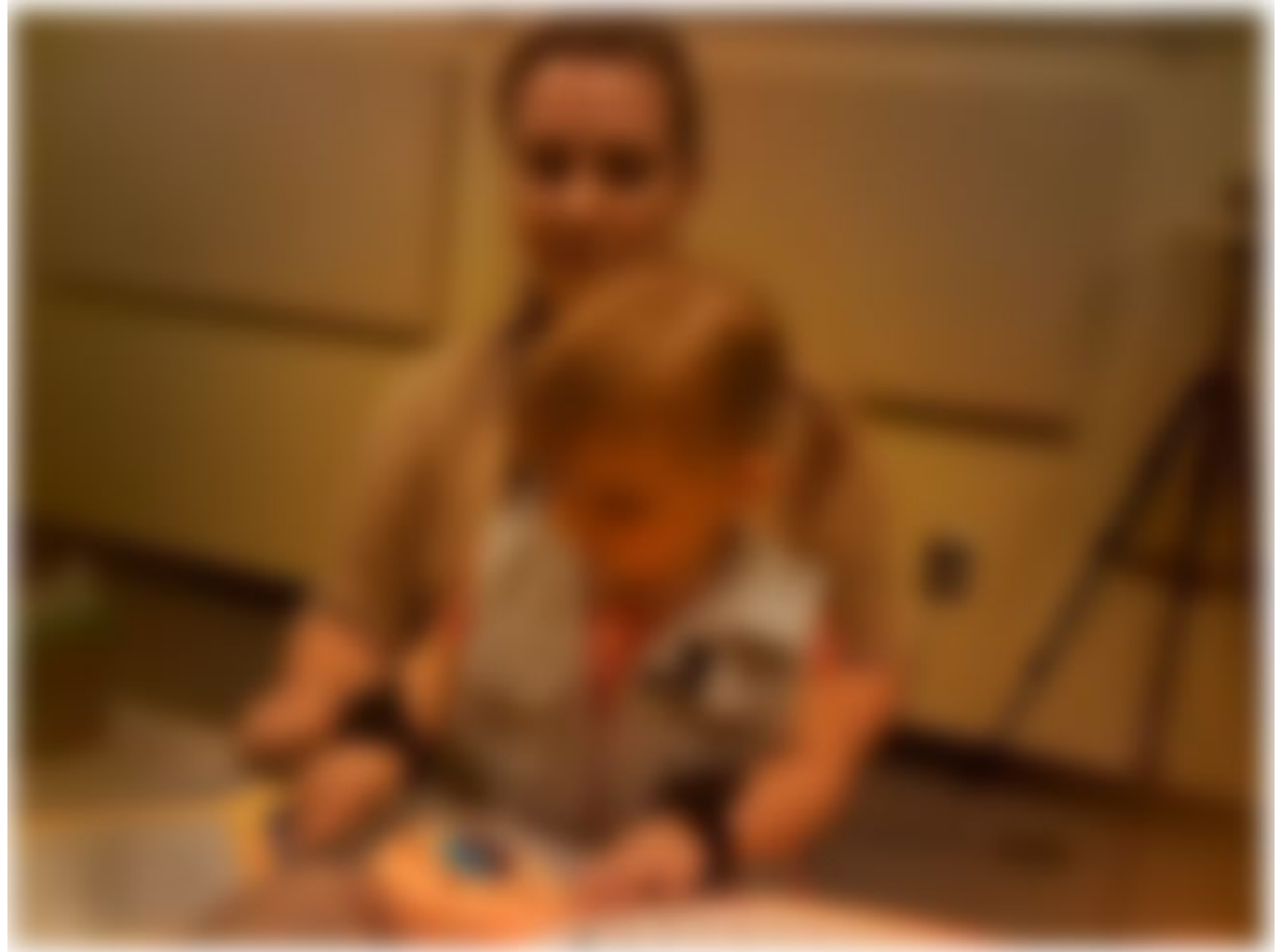
Pred: Put Plate GT: Put Plate



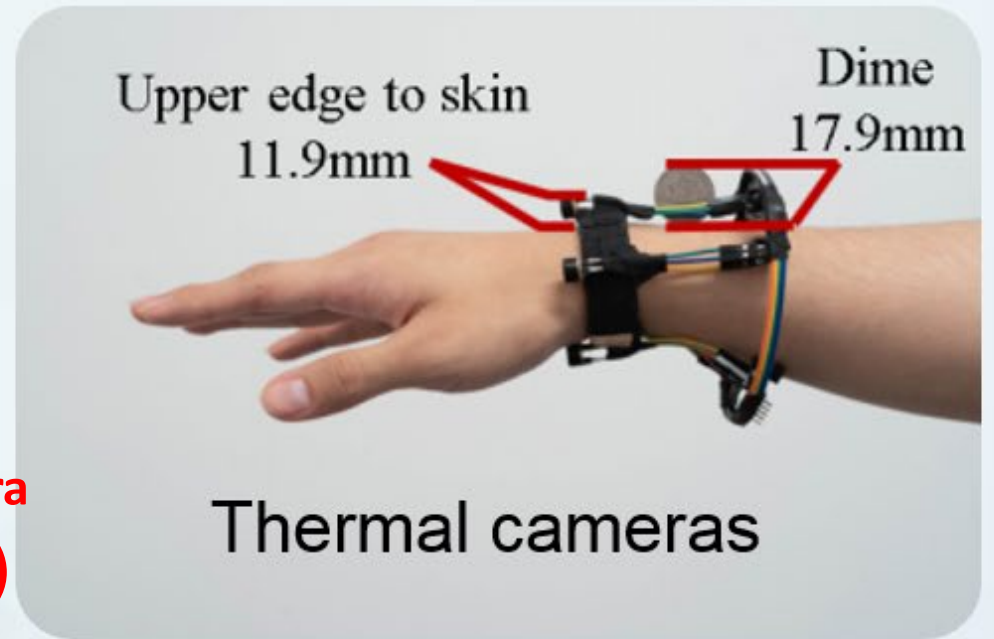
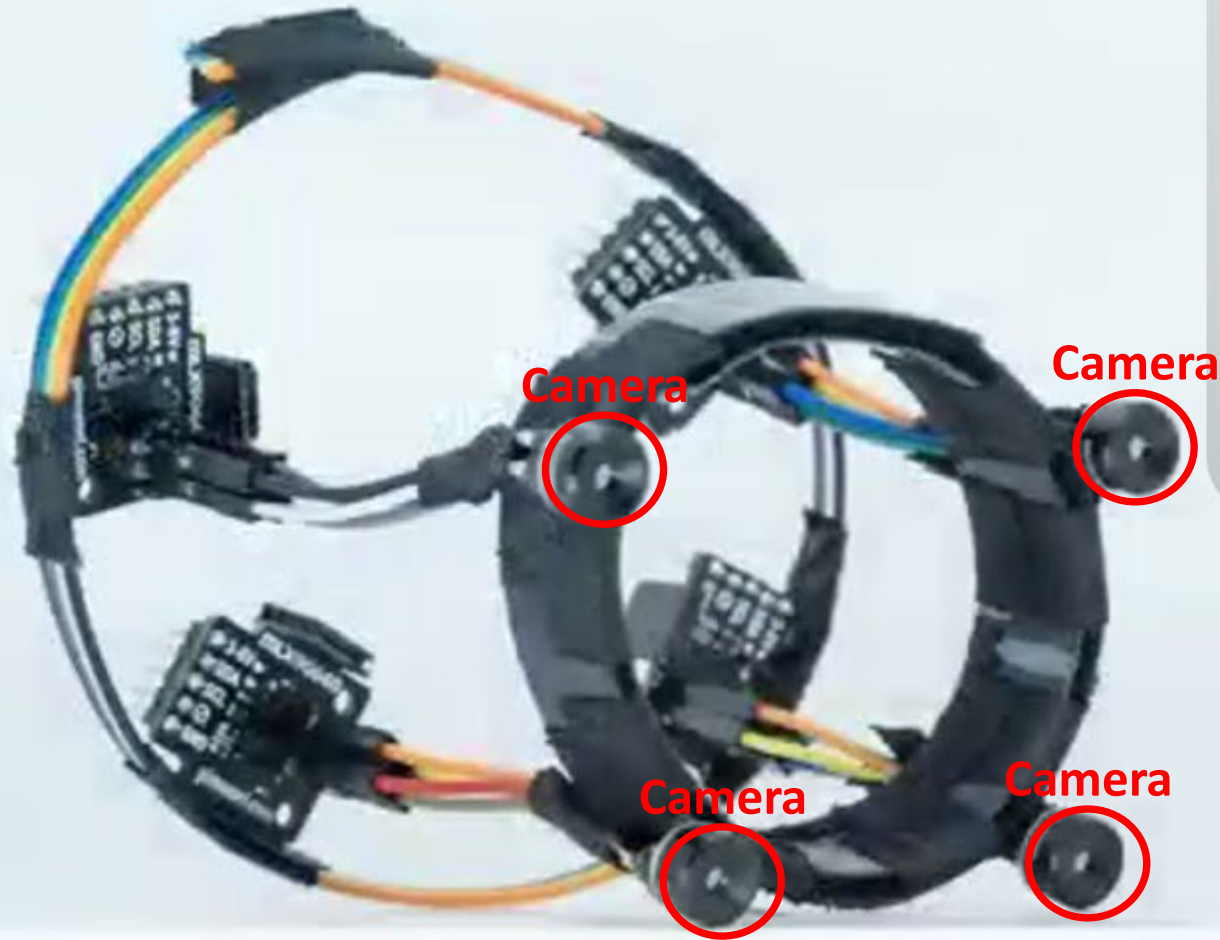
Eye contact detection in FPV



- Eye contact:
important signal
for developmental
disorders

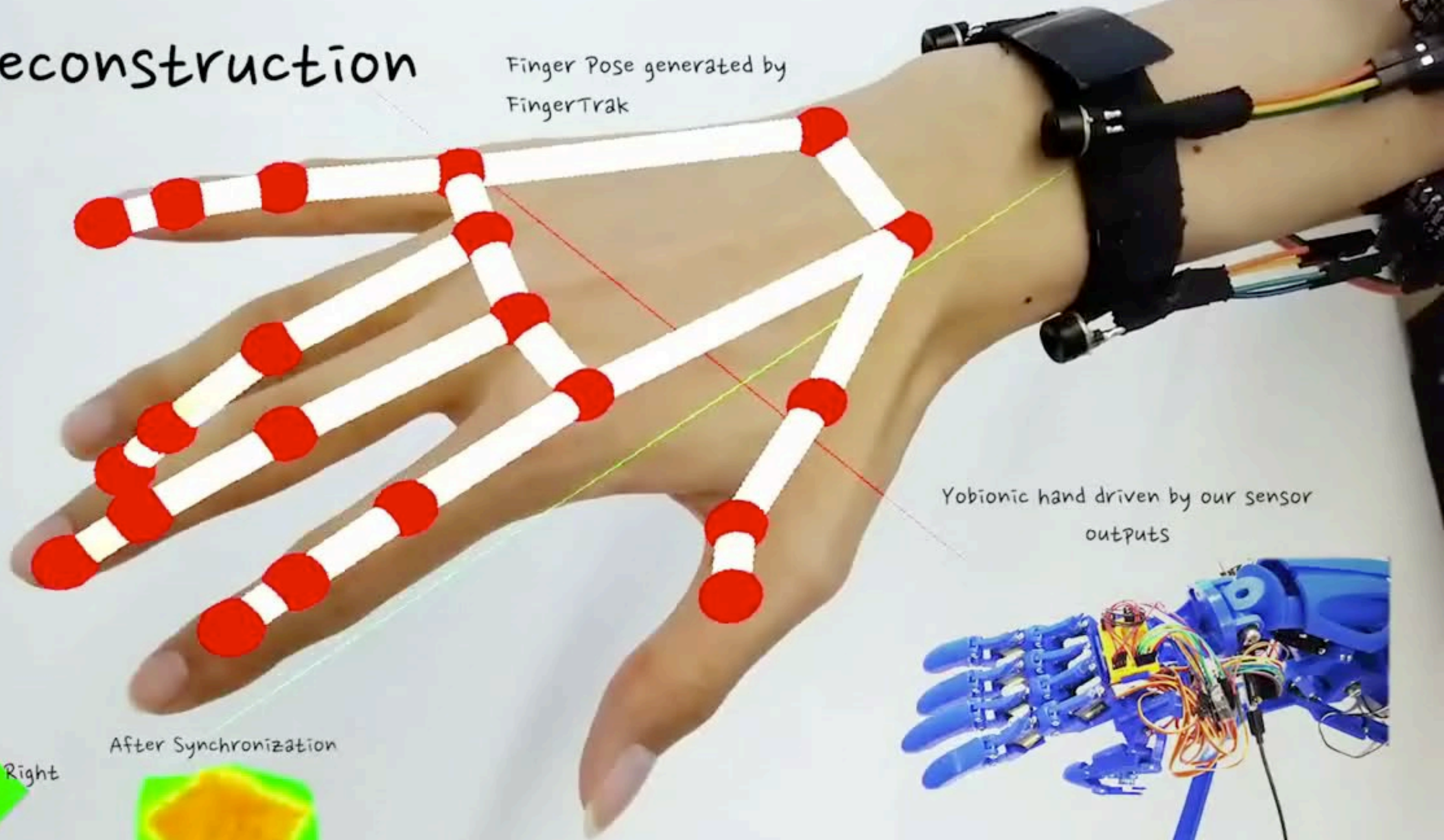


Finger tracking in FPV

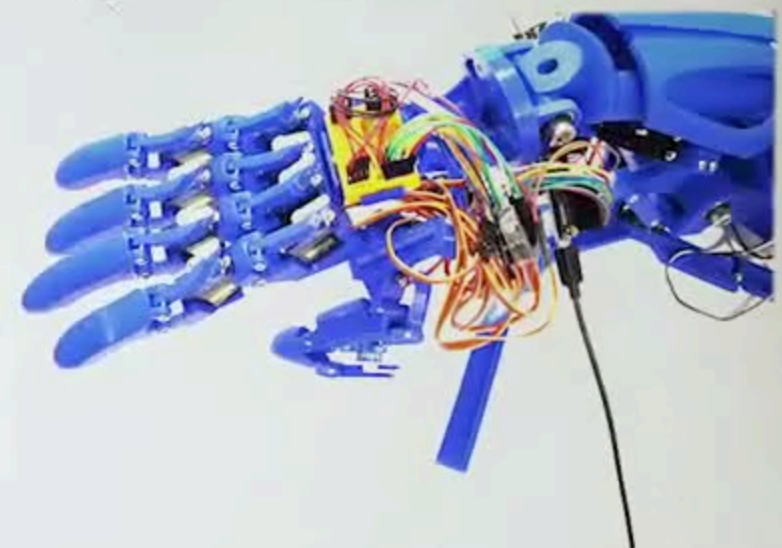


Reconstruction

Finger Pose generated by FingerTrak



Yobionic hand driven by our sensor outputs



Thermal cameras



After Synchronization

