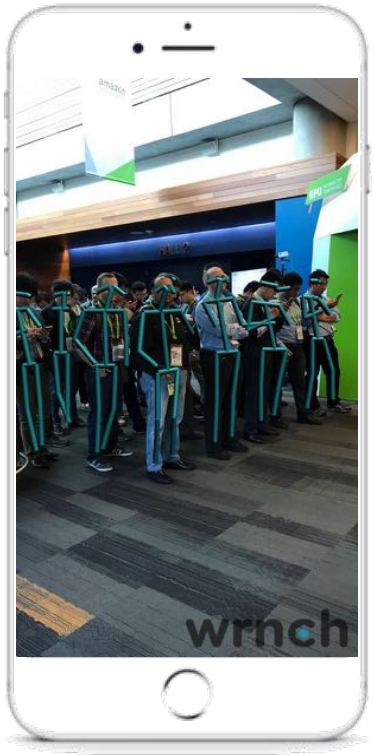
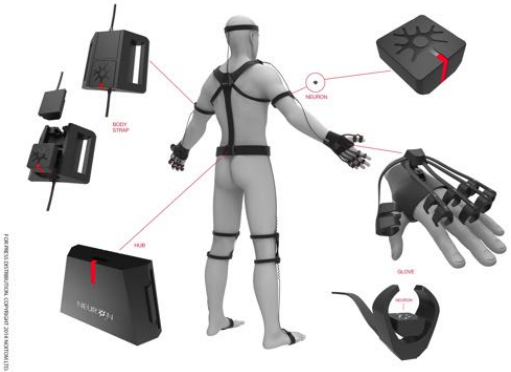
Using AI to Read Human Body Language In Real-Time from Standard Video



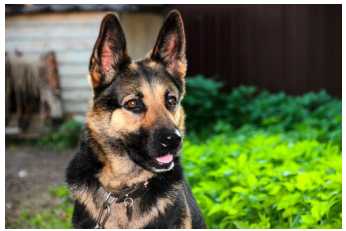
Friction: Humans are Physical / Computers are Digital

- Current Solutions
  - Keyboards
  - Voice
  - Specialty Sensors / Cameras
- Shortcomings
  - Expensive to own / operate
  - Time consuming to set up / use
  - Require specialty sensors / suits
- How do we enable computers to communicate with us in a natural way



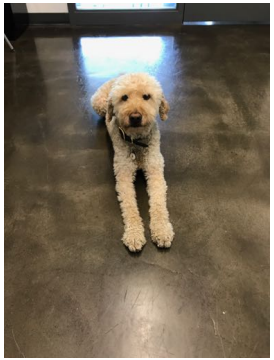
Inspired by Nature

- Humans are Visual Creatures
  - 80% of language is non-verbal
  - 80% of that is body language
- Dogs have co-evolved with humans to read human body language



Security

Work



Companionship

Entertainment

Navigation

Solution: Teach Computers to Read Human Body Language

- How do we make the digital version of “Man’s Best Friend”?
- Solution:
  - Ordinary cameras + GPUs + deep learning
- Value Proposition
  - Natural Interactivity
  - Consumer Hardware keeps getting cheaper and better
  - Deep Learning Brains keep getting smarter

Why Now? Peace Dividends of the Game Console Wars

- Kinect paved the way
- GPUs enable deep learning
- Game Engines enable synthetic data and thousands of new interactive applications

wrnch™: Human Pose Estimation Engine

User Application

In:
Video

Out:
Humans

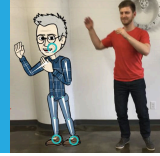
wrnch™

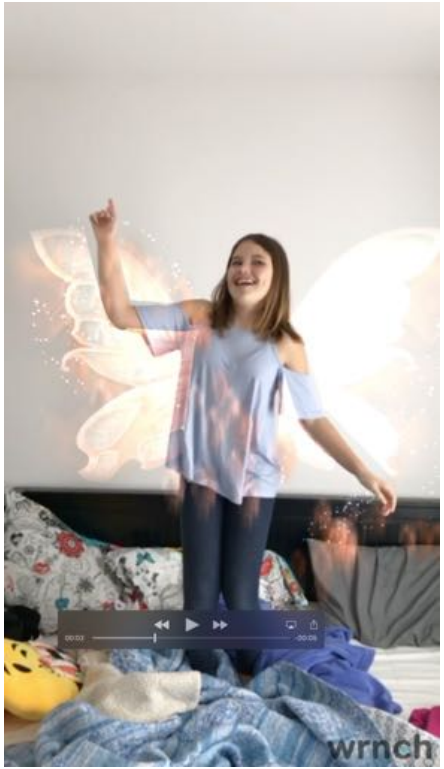
Wrnch™ Engine: Unique Features

- Deep Learning
  - Accurate
  - Robust
  - Keeps getting smarter
- Real-time
  - Enables interactivity
- No Specialty Hardware
  - Consumer grade cameras & GPUs
  - Take as input: any video feed from anywhere

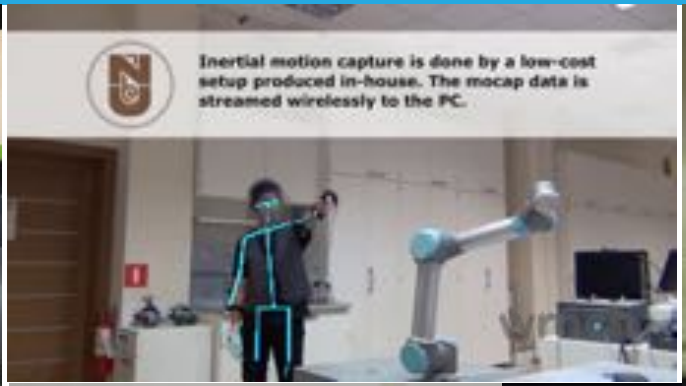
Live Demos

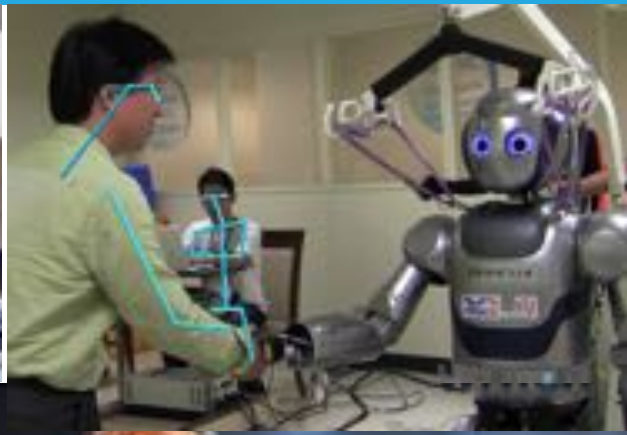
Applications: AR / VR





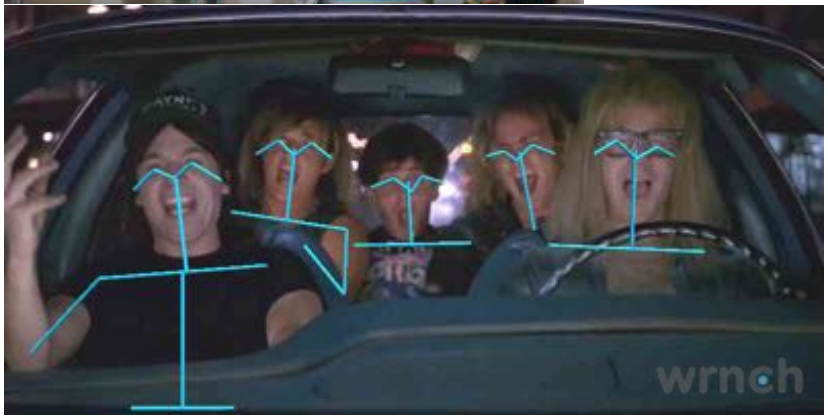
Applications: Robot Interaction



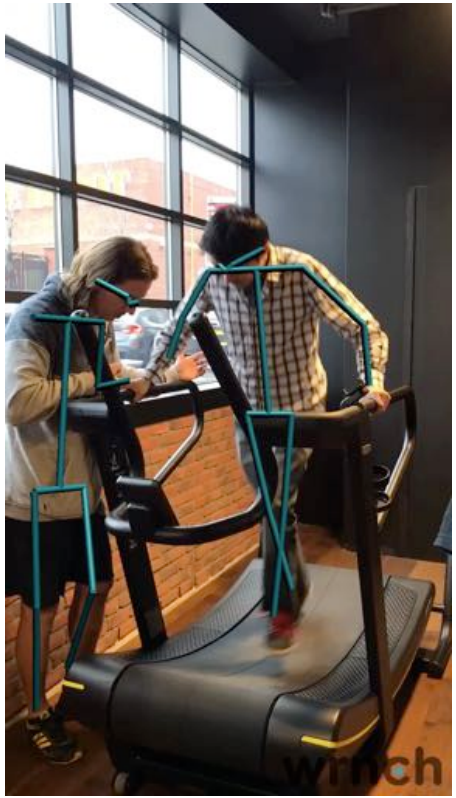


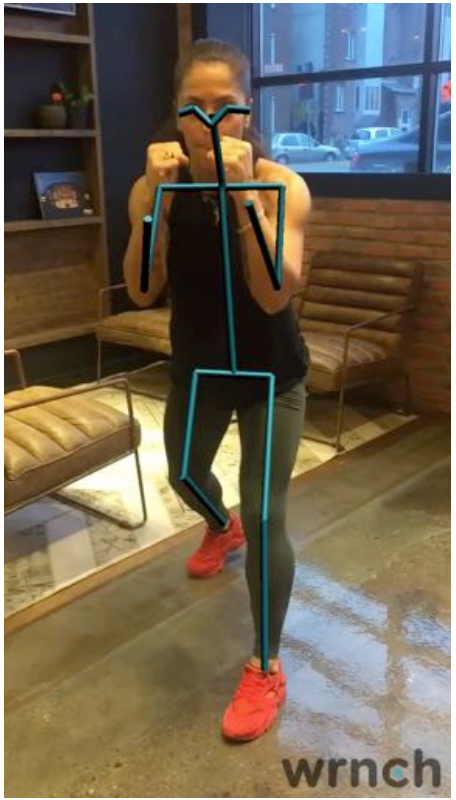
Applications: Security / Human Monitoring

- In City for safety
- In Home for elder care, etc.
- In Store for retail
- In Vehicle for awareness
- In Factories for accidents, etc.

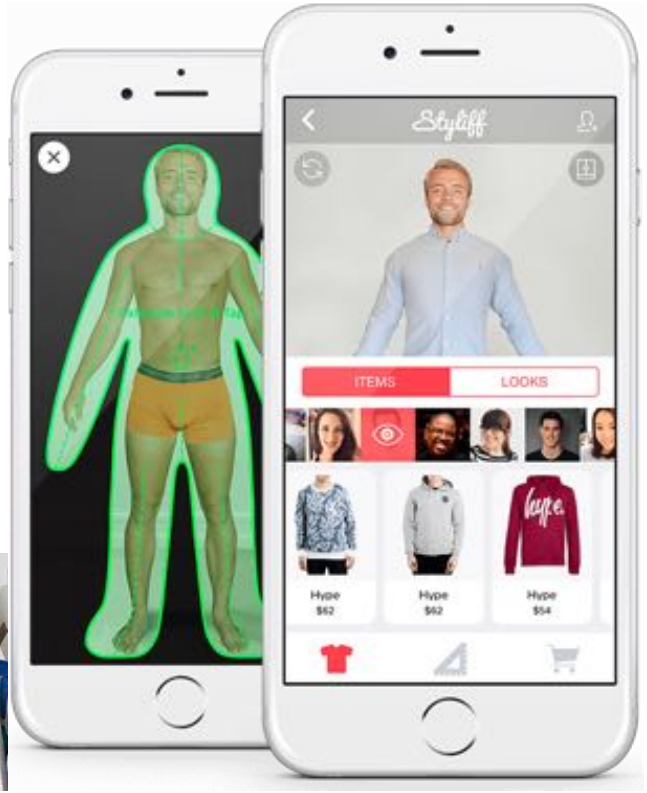


Applications:
Health & Wellness

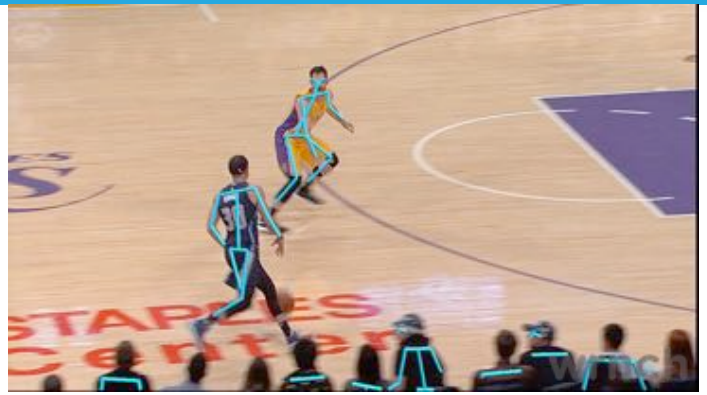




Applications: Eyes for Virtual Assistants

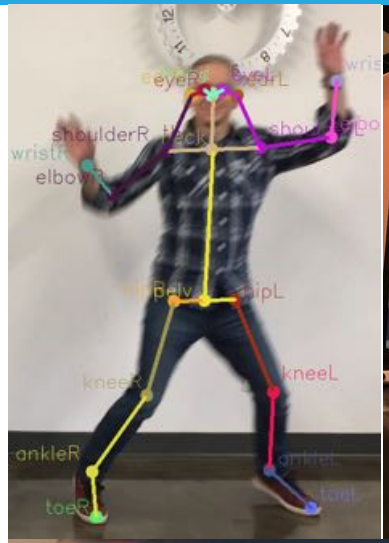


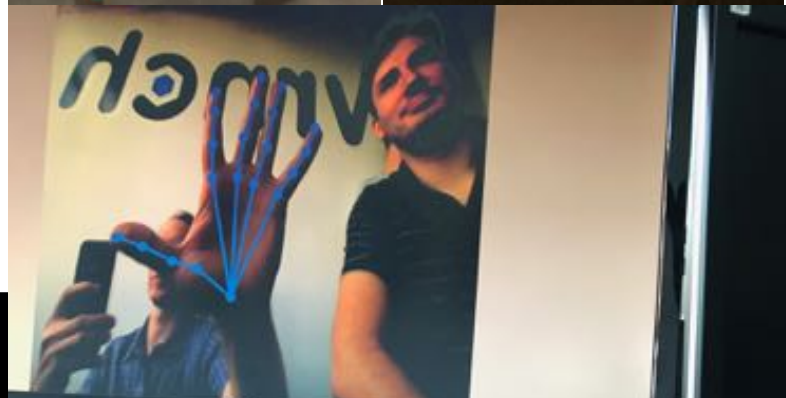
Applications: Sport Analytics



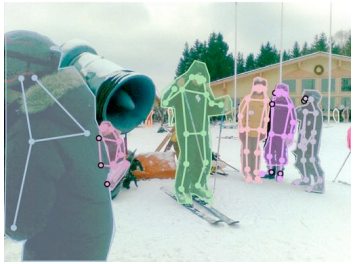
BodySLAM™: Unique Features

- Fast allowing real-time interactivity
- Accurate tracking of 63 body parts per person including fingers
- Robust across large numbers of people in crowded conditions

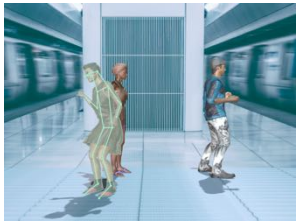




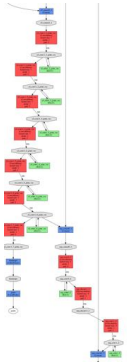
Wrnch Deep Learning Training Pipeline



Real Data



Synthetic Data

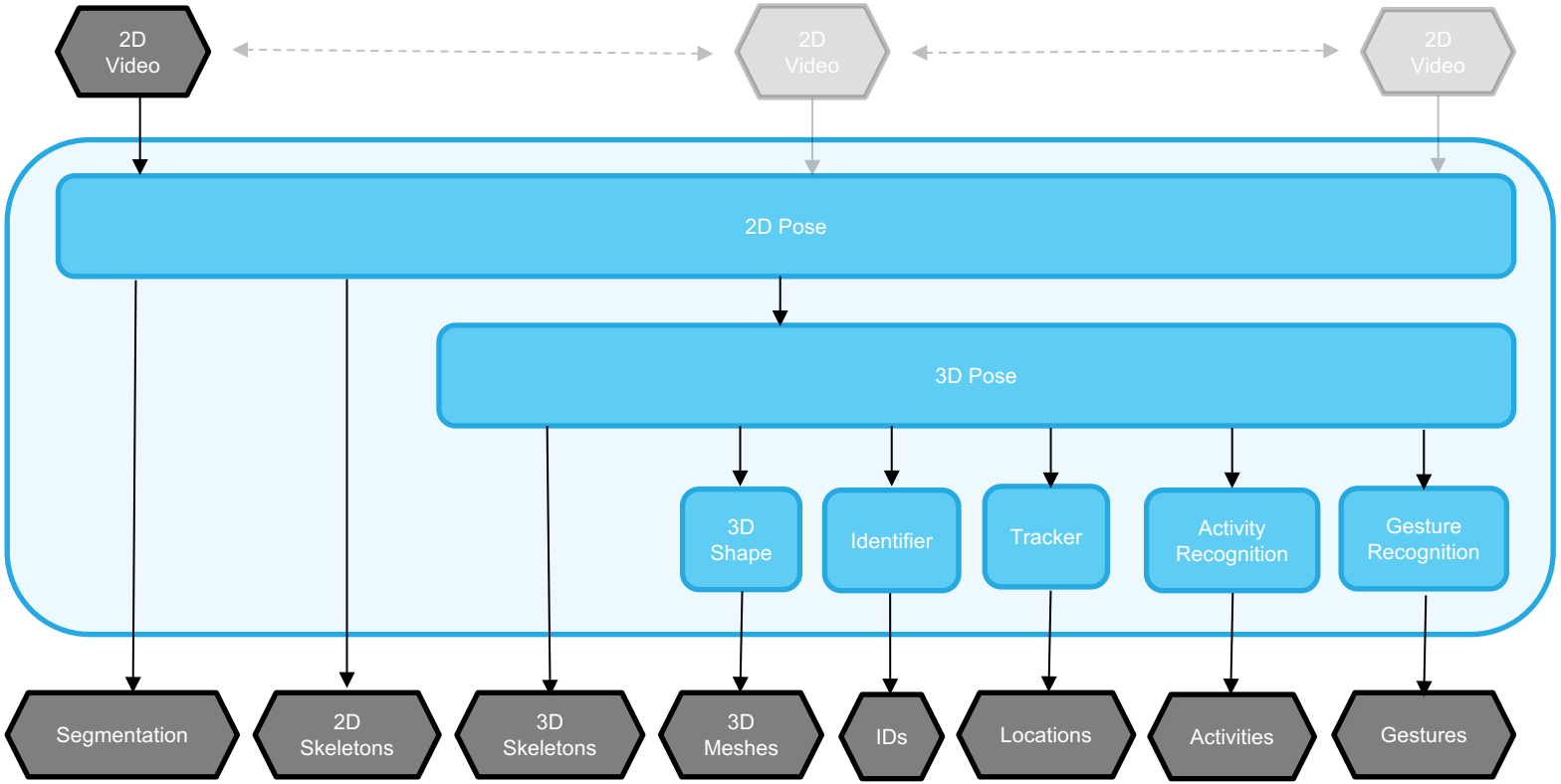


DL Models

Wrnch DL Training Framework

wrBrain

Wrnch Deep Learning Inferencing Architecture



TensorRT

Runtime Performance

GPU	CPU	OS	Total (mSec)	2D CNN	3D CNN	Misc
Jetsen TX2	n/a	Linux	**297**	200	84	13
K80	Intel Xeon E5-2686v4 2.3GHz	Linux	**131**	90	36	5
1080	Intel i7-5930K 3.5GHz	Linux	**37**	22	12	3
1080	Intel i7-5930K 3.5GHz	Windows	**41**	24	13	4
Titan XP	Intel i7-5930K 3.5GHz	Linux	**30**	16	11	3
Titan XP	Intel i7-5930K 3.5GHz	Windows	**34**	18	12	4

Roadmap

- Gesture Recognition
  - Point At
  - Thumbs up
  - Fist
- Activity Recognition
  - Fall detection
  - Pick up item
  - Put down item
- Multiple Camera Support
  - Triangulation
  - Persistent tracking
  - 3D

Questions?