YOGASOFT

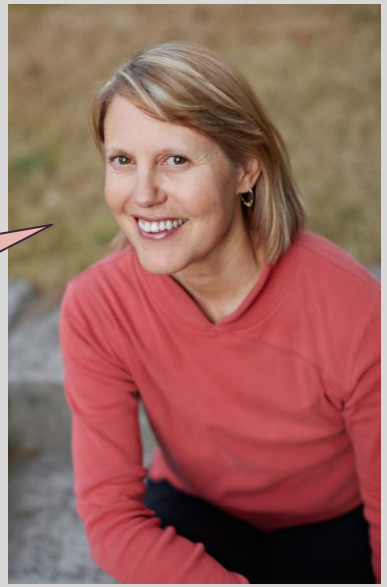
Alexandra Savelieva
Pritesh Patel
Bryan Morgan
<we are open for one more team member, esp. from yoga fans>

MIDS w251 Final Project

Problem

- With COVID-19 pandemics, gyms and private yoga practices around the world have closed for “in-person” instruction; instructors are trying to switch to online classes. It comes with challenges:
- Some students don’t have good enough network connection to allow for video streaming of their practice
- When it’s more than 3 students, it’s very hard for the instructors to monitor the class
- Some students are not willing to share their videos for privacy reasons, but want feedback

Depending on what online platform I am using to teach on, either I can’t see the students AT ALL (Facebook Livestream) or with Teams and Zoom, if I am going to record the sessions, then I have to pin my own video, which makes the students so tiny that I couldn’t possibly help correct their poses!



Lynn Jensen, E-RYT, RPYT, C-IAYT, MBA

Proposal

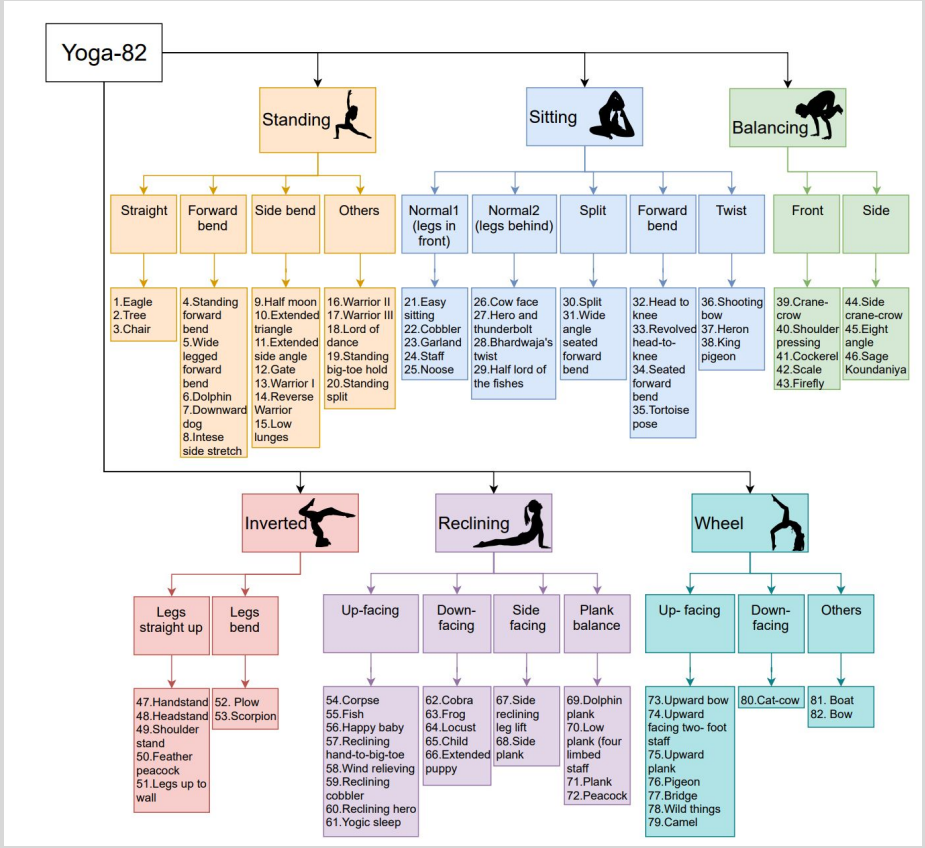
- Develop an IoT solution that facilitates guided yoga or pilates that facilitates guided and tracked practice
- Recognize poses on the device and send them to the cloud, without sending the full video stream
- Provide feedback to the participants
  - Real-time (correction of poses)
  - Trends (FitBit-like approach, except that we track how many asanas and how well you did, progress, day challenge, etc.)
- Provide feedback to the instructor on how good individual students and the group is doing
  - Real- time (are people following or out of sync)
  - Trends (how many participants follow, do they improve week to week, etc.)

Data set

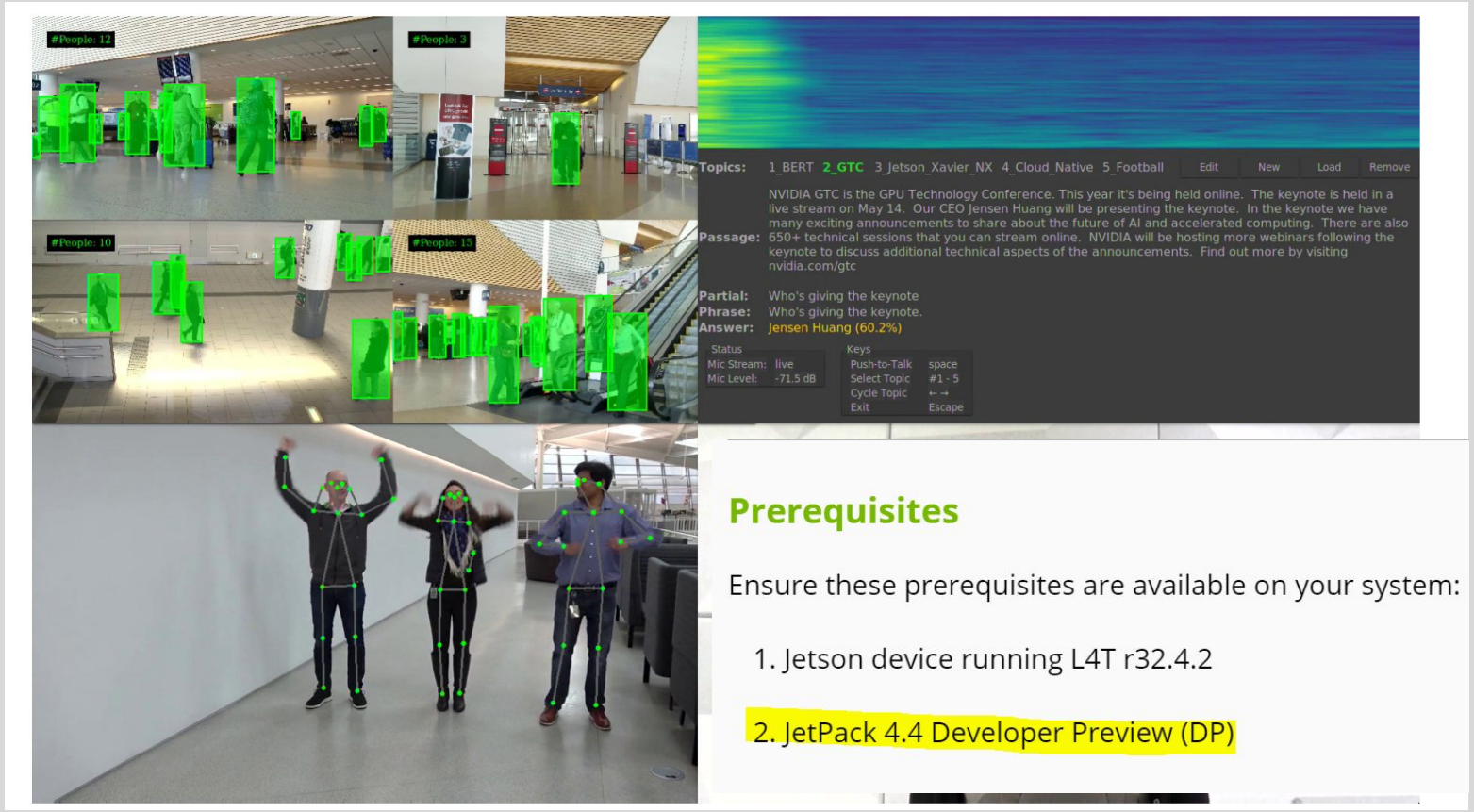
- Verma et.al. Yoga-82: **A New Dataset for Fine-grained Classification of Human Poses** (Powered by Bing pictures): https://arxiv.org/pdf/2004.10362v1.pdf [22 Apr 2020]
- Other existing data sets on pose analysis:

Datasets	#Train	#Test	Source	Target poses
MPII [4]	25,000	-	YouTube	Diverse
LSP [17]	1,000	1,000	Flickr	Sports
LSP-Ext. [18]	1,0000	-	Flickr	Sports
FLIC [23]	6,543	1,016	Movies	Diverse
SHPD [6]	18,334	5,000	Surveillance	Pedestrian
Yoga-82	21,009	7,469	Bing	Yoga

- Self-curated labeled data with the support of Certified yoga trainer **Lynn Jensen:** https://yogaforfertility.net/
  -



Evaluation of Feasibility with Jetson



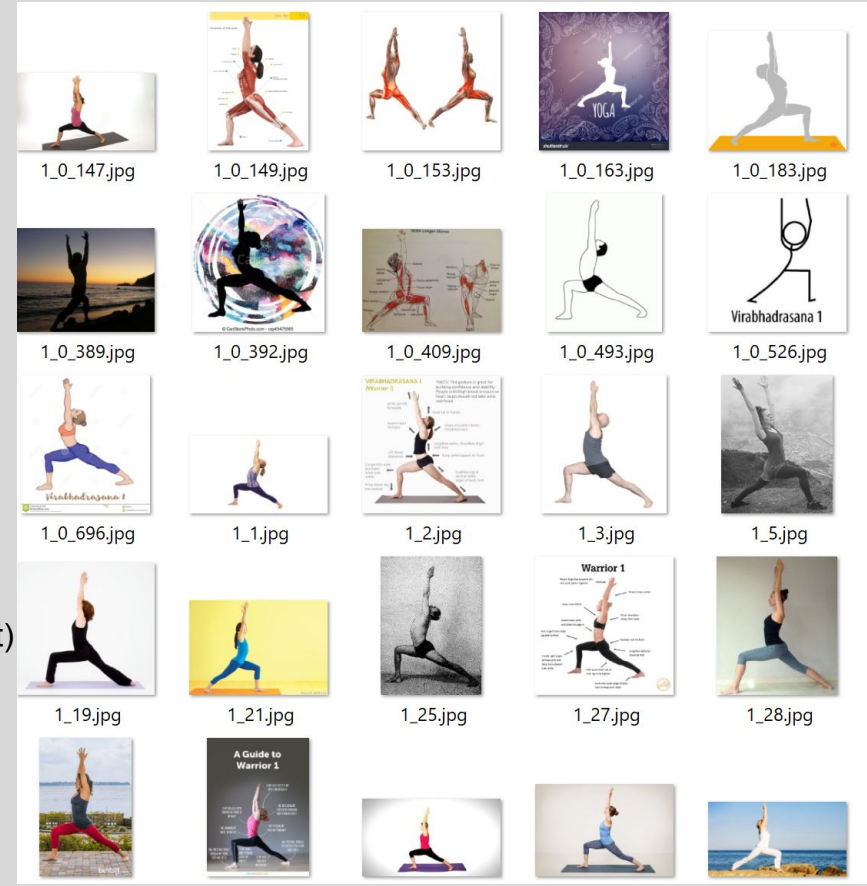
- https://github.com/NVIDIA-AI-IOT/jetson-cloudnative-demo

Proof of concept: Raspberry Pi “YogAI”

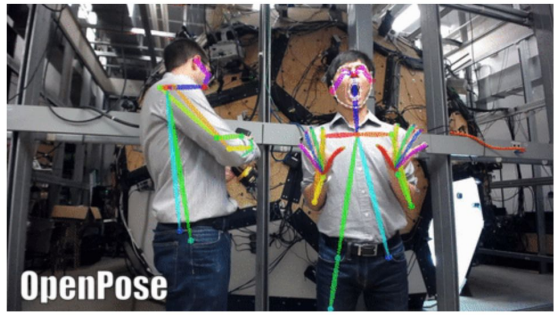
- https://www.hackster.io/yogai/yogai-smart-personal-trainer-f53744

Exploratory data analysis of Yoga-82 dataset

- Tried out Cobra, Dolphin, Warrior 1 buckets (1121, 113, 293 samples)
- Data quality:
  - ~7.5% URLs are unavailable
  - ~11% images are corrupted
  - ~20% images are drawings, not photos
  - ~7% are repeats (in different zoom / cropped /etc.)
  - <.5% are unrelated junk
  - Different variations of same pose are represented
  - Different angles (front, side, left, right) are captured
  - [>1] are wrongly labeled (e.g. upward facing dog or sphinx instead of cobra)
  - [>1] are poorly executed (e.g. the pose is close to correct, but bad neck alignment)
  - [>1] have text over / side-by-side with the pose
  - ~4% have more than one person doing the pose
  - ~75% are women
- Down to [at best] 915, 75, 242 samples for each pose respectively (<80%)



Pose estimation: PoseNet and OpenPose



Authors Gines Hidalgo (left) and Hanbyul Joo (right) in front of the CMU Panoptic Studio

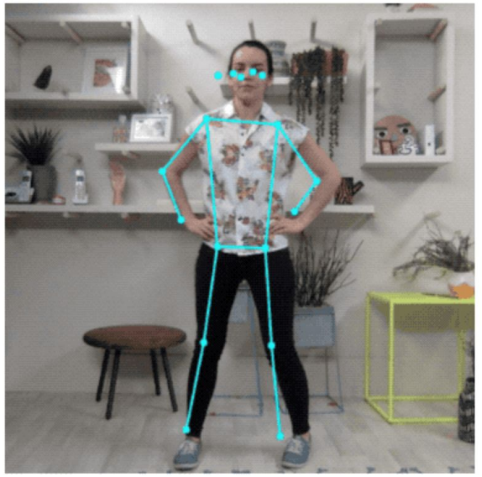
Features

- **Functionality**:
  - **2D real-time multi-person keypoint detection**:
    - 15 or 18 or **25-keypoint body/foot keypoint estimation**. **Running time invariant to number of detected people**.
    - **6-keypoint foot keypoint estimation**. Integrated together with the 25-keypoint body/foot keypoint detector.
    - **2x21-keypoint hand keypoint estimation**. Currently, **running time depends** on **number of detected people**.
    - **70-keypoint face keypoint estimation**. Currently, **running time depends** on **number of detected people**.
  - **3D real-time single-person keypoint detection**:
    - 3-D triangulation from multiple single views.
    - Synchronization of Flir cameras handled.
    - Compatible with Flir/Point Grey cameras, but provided C++ demos to add your custom input.
  - **Calibration toolbox**:
    - Easy estimation of distortion, intrinsic, and extrinsic camera parameters.
  - **Single-person tracking** for further speed up or visual smoothing.
- **Input**: Image, video, webcam, Flir/Point Grey and IP camera. Included C++ demos to add your custom input.
- **Output**: Basic image + keypoint display/saving (PNG, JPG, AVI, ...), keypoint saving (JSON, XML, YML, ...), and/or keypoints as array class.
- **OS**: Ubuntu (14, 16), Windows (8, 10), Mac OSX, Nvidia TX2.

Note: We've just released Version 2.0 with a new ResNet model and API. Check out the new documentation below.

This package contains a standalone model called PoseNet, as well as some demos, for running real-time pose estimation in the browser using TensorFlow.js.

Try the demo here!



PoseNet can be used to estimate either a single pose or multiple poses, meaning there is a version of the algorithm that can detect only one person in an image/video and one version that can detect multiple persons in an image/video.

Refer to this blog post for a high-level description of PoseNet running on Tensorflow.js.

To keep track of issues we use the tensorflow/tfjs Github repo.

Open questions

- How many pictures per pose will we need for training?
- Is it okay to use drawings together with photos?
- What is negligible level of “poorly” executed poses?
- Does text over picture interfere with training?
- Does it matter if the poses are by women or men?
- 2d or 3d pose estimations for yoga?
- Brainstorm actual implementation
- Are Springer articles available to Berkeley for free?

Next steps

- Finish data download and cleaning process
- Labeling (good yoga pose, bad yoga pose, not a yoga pose)
- Augmenting the dataset with albimentations (zoom, shift, rotate, flip)
- Choosing & training DL algorithms

References

Data samples:

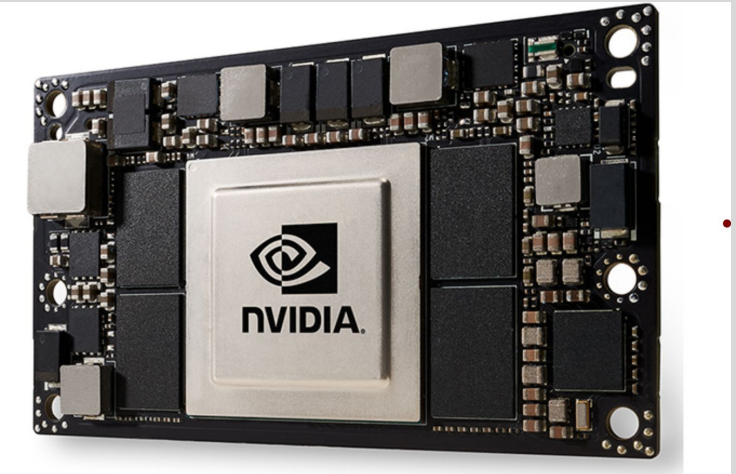
- Chair - https://drive.google.com/open?id=1_puK9zwIsMJTVV5C9k50CK9Ow7oxA6xq
- Dolphin - https://drive.google.com/open?id=15uEspxXFp0RW95JZorjbQQFD0_N0vymN
- Warrior 1 - https://drive.google.com/open?id=1dlLrfOjdEZaCL8MR68ENdosgAILPvKnY
- Cobra - https://drive.google.com/open?id=1UBLWellUV5R6nSgVt92QlFpaOCnBWTj

Relevant code:

- Posenet: https://github.com/tensorflow/tfjs-models/tree/master/posenet
- OpenPose: https://github.com/CMU-Perceptual-Computing-Lab/openpose
- Albumentations: https://pypi.org/project/albumentations/

Relevant papers:

- Yadav, Santosh & Singh, Amitojdeep & Gupta, Abhishek & Raheja, Jagdish. (2019). Real-time Yoga recognition using deep learning. Neural Computing and Applications. https://link.springer.com/article/10.1007/s00521-019. 10.1007/s00521-019-04232-7.
- Verma, Manisha & Kumawat, Sudhakar & Nakashima, Yuta & Raman, Shanmuganathan. (2020). Yoga-82: A New Dataset for Fine-grained Classification of Human Poses.



Questions?