Detecting Exercise and Calories Burned Using Human Pose Estimation

Shivesh S. Jadon, Dr. Vangelis Metsis
Department of Computer Science, College of Science and Engineering, Texas State University

Introduction

Background
With countless devices present nowadays, detecting body vitals, like calories burned during a workout session, is not that difficult. Many smart watches automatically detect exercise and calories based on body movement sensors like gyroscope and heart rate sensors. Auxiliary sensors like GPS, infrared, visible-light LED, and pedometer are often needed to track your data in the background.

Motivation
- Do we need these many devices and sensors to track calories burned?
- How accurate are smart watches' vital calculations?
- Can we design a system that solely works based on a mobile phone's camera to detect calories burned in an activity?
- Can we automatically detect and classify which exercise user is doing based on just camera inputs?

Goals
To develop a mobile application that can independently track:
- human body joints,
- posture,
- exercises and workouts,
- calories burned

without the use of any external devices other than a mobile phone.

Materials

Human Pose Estimation is a Computer Vision technique to track localized human body keypoints (joints) like head, elbows, shoulders, hip, knee etc. with the help of input images/video. The sole input in our application is the mobile's camera feedback, opposed to popular multi-sensor setup.

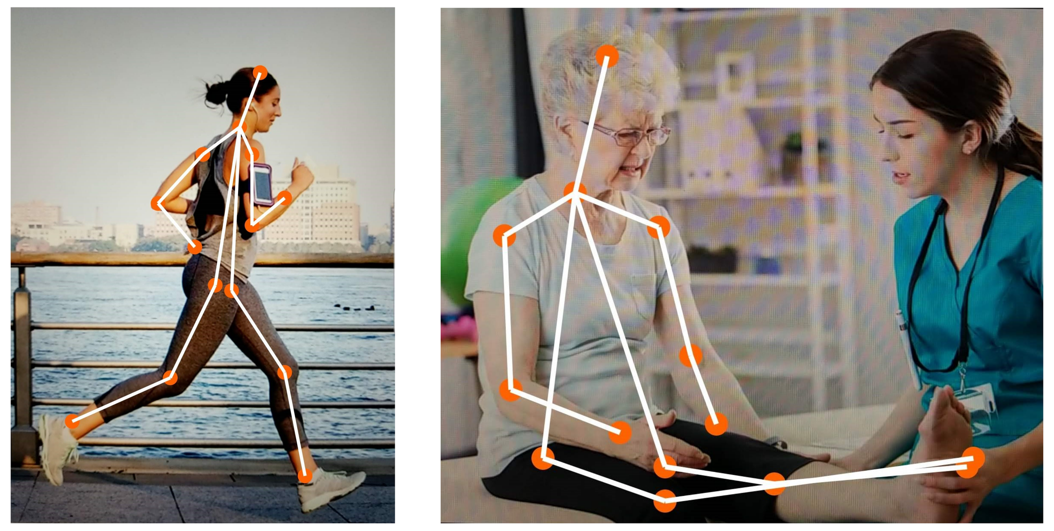


Fig 1,2: Body joints as visualized through Human Pose Estimation technique with our application running on Android mobile phone

Android Studio | TensorFlow | Fitbit | Unity3D | Google ARCore

Methodology

Data Collection
6 activities are tracked on the Fitbit smartwatch and our mobile application. The data is composed of 14 joints that are tracked using Pose Estimation. The vector data is then cleaned and smoothed out using locally weighted smoothing (LOESS)

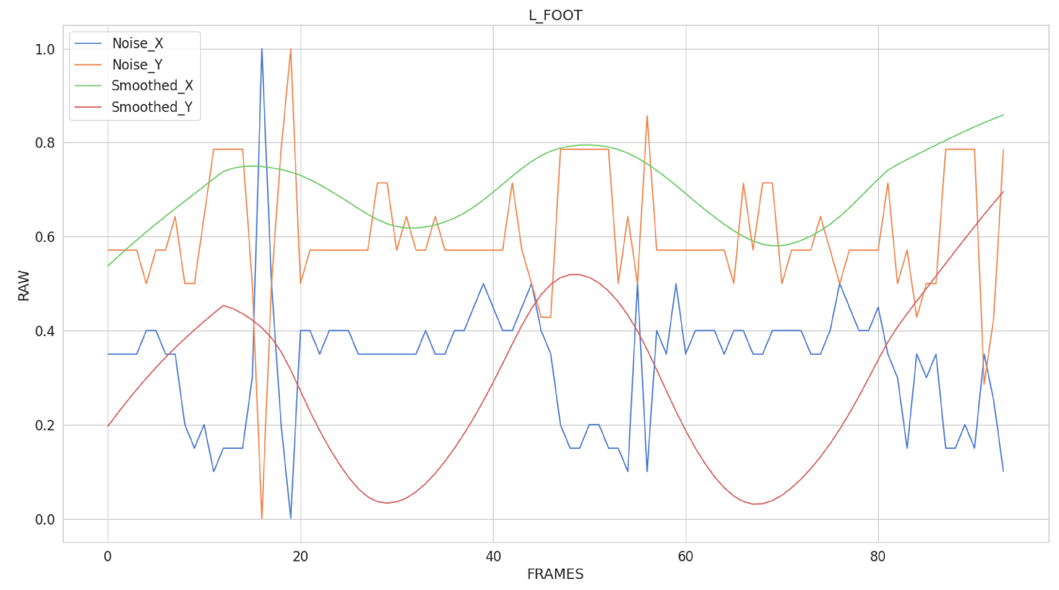


Fig 3: Raw and smoothed out data showing X and Y time-series data for left foot movement as tracked by pose estimation model

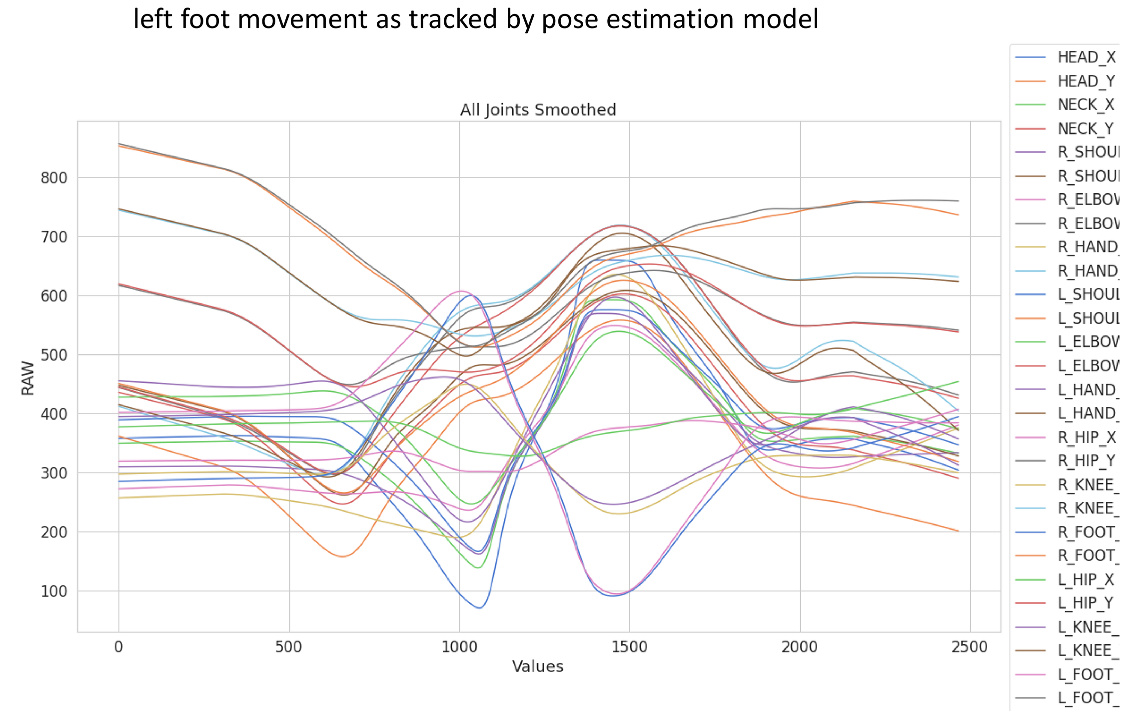


Fig 4: LOESS smoothed-out data showing X and Y time-series data for all the 14 joints with respect to time as tracked by pose estimation

Activity Classification
We detect workout patterns using Dynamic Time Warping (DTW) and training Long Short Term Memory (LSTM) recurrent neural network to automatically classify exercises based on Joints pattern and movement.

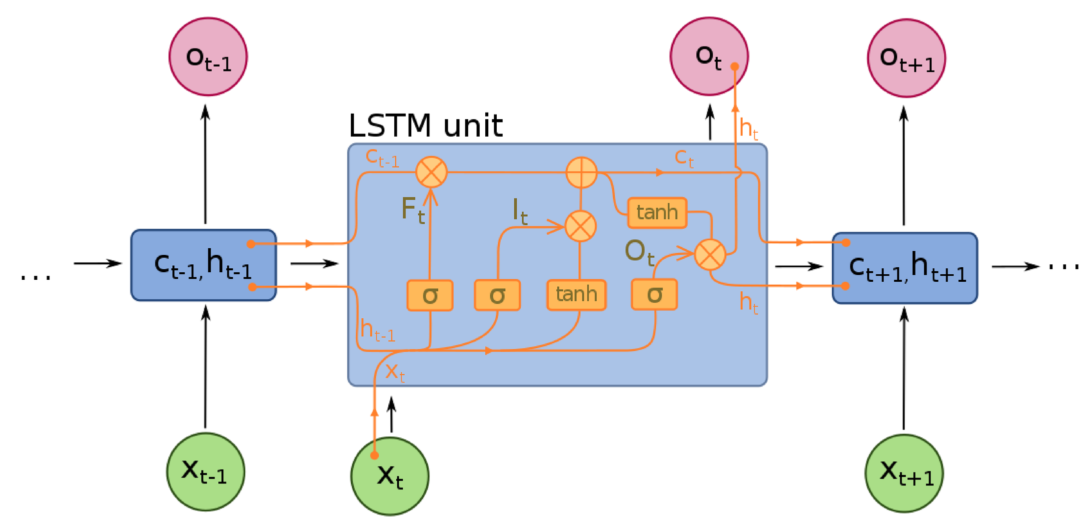


Fig 5: Long Short Term Memory (LSTM) recurrent neural network architecture design

After classifying exercise, we use Metabolic equivalent to Calculate the calories burned at each time-series.

Energy expenditure (calories/minute) = .0175 x MET x weight

Expected Results

Exercise Classification
The model can actively classify between trained activities using dynamic time warping with 14 joints on a mobile device. We expect an increase in classification accuracy and better results with the use of LSTM recurrent neural network.

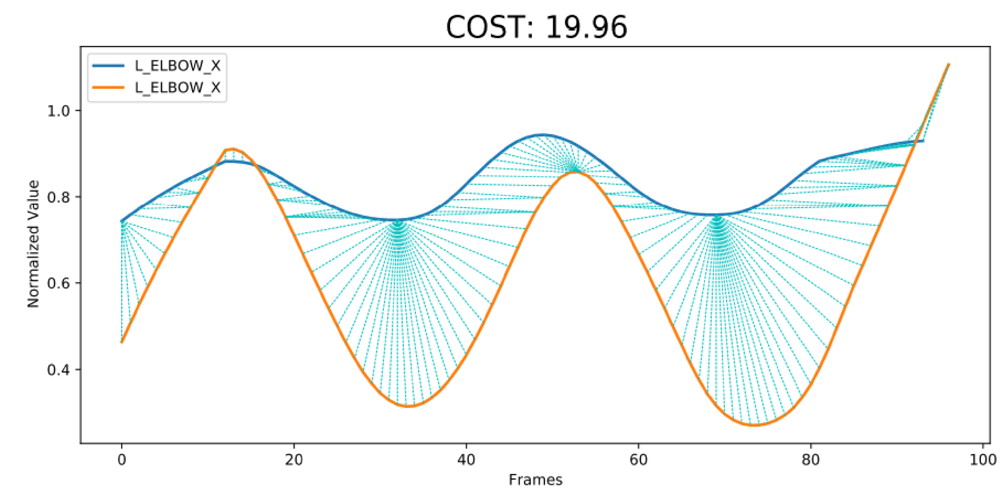


Fig 6: Left Elbow X time-series comparison with Dynamic Time Warping.

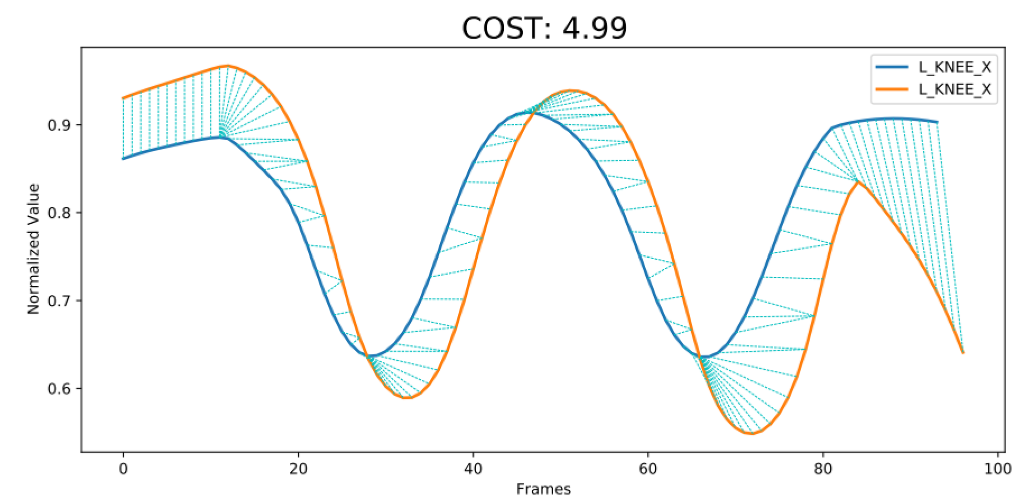
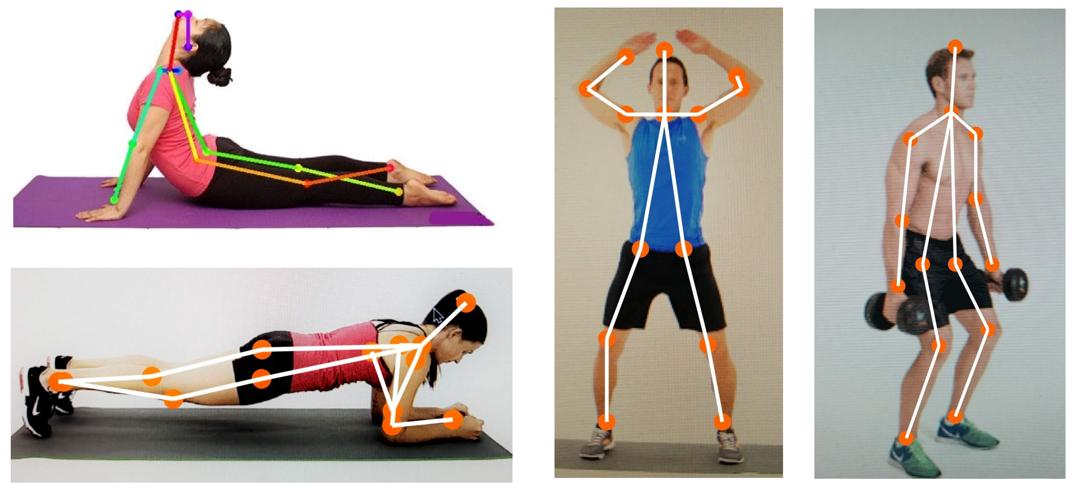


Fig 7: Left Knee X time-series comparison with Dynamic Time Warping.

Calories Burned
The designed system can calculate:
- Calories burned in real-time
- Future/expected calorie burn with on-going performance
- Expected time to reach calorie goal
- Real-time Augmented Reality based visualizations
- Statistics about body vitals with processed data

With this technology, users can simply place their phone in front of them and start exercising or workout. The phone's camera will immediately recognize the type of activity user is performing and calculate the calories burned to achieve that task in real-time, every frame. The system can track up to 14 joints and their movements simultaneously without a need for a depth camera. The application classifies and calculates calories without any external devices or utilities.



We plan on comparing the accuracy of our model to that of various smartwatches, test new systems for calorie calculations, and further improve the accuracy of our activity classification.

Future Studies

The research will directly effect the future development in mobile phone's capabilities in health and medical fields like:
- Tele-rehabilitation therapy
- Joint therapy,
- Correcting exercise posture,
- Fall detection for seniors,
- Coaching of elderly population
- Body movement coordination in people with Parkinson's disease
- Physical therapy for Cerebral palsy and Ataxia

Commercial Applications

Commercial ready applications that can be used by millions of people who cannot afford expensive tracking devices for in-home training and workouts as well as specific cases like movement or posture correction, joint therapy in seniors, etc.

With further research in independent muscle and joint based calorie calculations, the commercial aspect of tracking human vitals can be revolutionized.

Commercial centric application specially for:
- Senior citizens,
- People with movement-posture accessibilities
- People lacking advanced medical care and equipment
- Population in isolated and rural area

References

1. Newell, Alejandro, Kaiyu Yang, and Jia Deng. "Stacked hourglass networks for human pose estimation." European conference on computer vision. Springer, Cham, 2016.
2. Wei, S. E., et al. "Convolutional pose machines. arXiv 2016." arXiv preprint arXiv:1602.00134.
3. Sandler, Mark, et al. "Mobilenetv2: Inverted residuals and linear bottlenecks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

TEXAS STATE UNIVERSITY®
The rising STAR of Texas
MEMBER THE TEXAS STATE UNIVERSITY SYSTEM