Demographic Inference from Smartphone Gait Acceleometry

Applying Deep Convolutional Networks to the MyHeartCounts Six-Minute Walk Test

Daniel Wu[1], Anna Shcherbina[2], Steve Hershman[2], Euan Ashley[2]

Department of Computer Science[1] and Medicine[2], Stanford University

Background

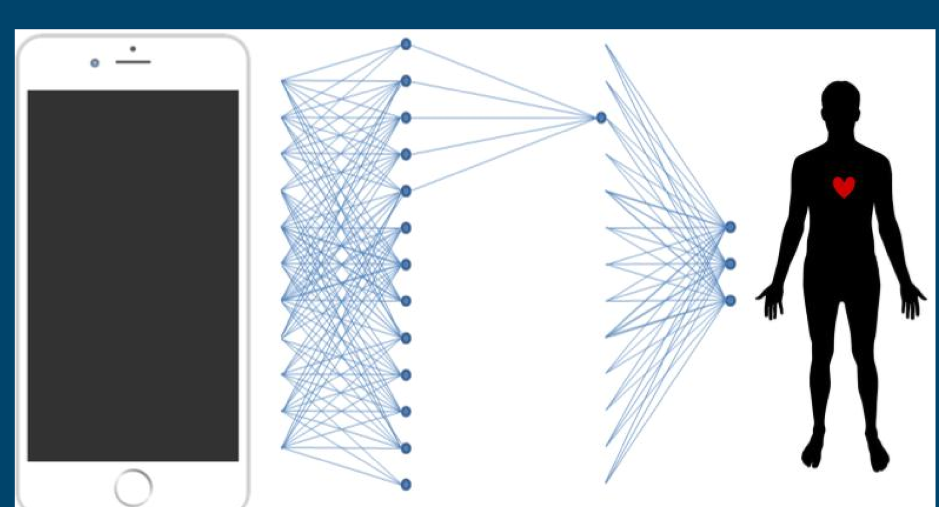
Modern smartphones are ubiquitous, and outfitted with a rich sensor suite, making them promising platforms for distributed health monitoring. The MyHeartCounts app was a flagship cardiovascular health study deployed to over 50,000 iOS users.

Figure: My Heart Counts iOS App

Participants downloaded the MyHeartCounts app, and were directed to an eConsent process. In addition to taking a six minute walk test, users disclosed gender, age, weight, height, smoking history, sleep habits, and cardiovascular illness.

Objectives

We aim to develop deep learning gait models which can infer, from a six minute walk test, a patient's age, gender, BMI, and cardiovascular risk.



Methods

Six minute walk test data from smartphone acceleometer, gyroscope, and acceleometer and gyroscope were augmented with a "magnitude" channel, which provides a rotation-invariant representation of the data.

We explored various preprocessing steps, including bandpass and lowpass Butterworth filtering, dynamic time warping, and gait cycle alignment with sigmoidal fitting, but ultimately decided on noise removal with no filtering and simple data augmentation.

As a benchmark, we used random forest and gradient boosted decision trees on both the raw data and simple statistical features extracted from the raw data.

We then trained various deep learning models, from simple single layer perceptrons to VGG-esque architectures.

We experimented with transfer learning, residual connections, optimizers, loss functions, and then conducted hyperparameter tuning to maximize accuracy at the gait cycle level.

Finally, we ensemble together model predictions on all of the gait cycles of a user in order to create a user prediction.

Conv 1-1
Conv 1-2
Pooing

Conv 2-1
Conv 2-2
Pooing

Conv 3-1
Conv 3-2
Conv 3-3
Pooing

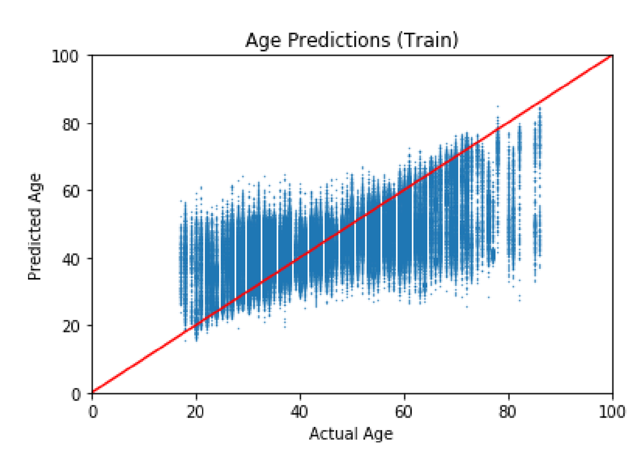
Conv 4-1
Conv 4-2
Conv 4-3
Pooing

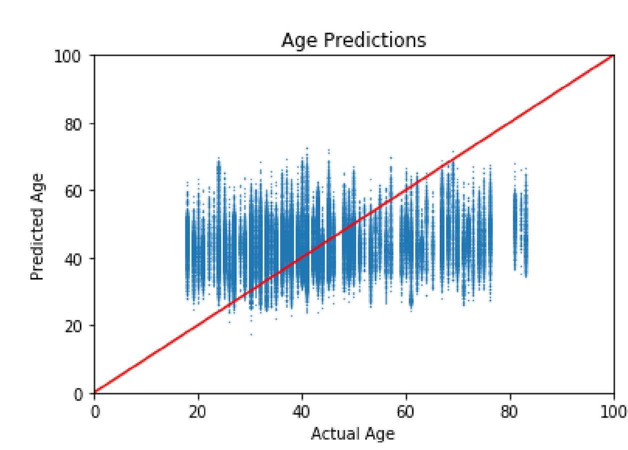
Conv 5-1
Conv 5-2
Conv 5-3
Pooing

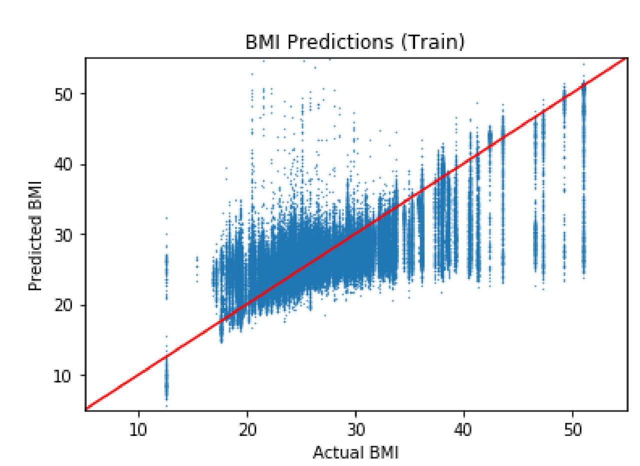
Dense
Dense
Dense

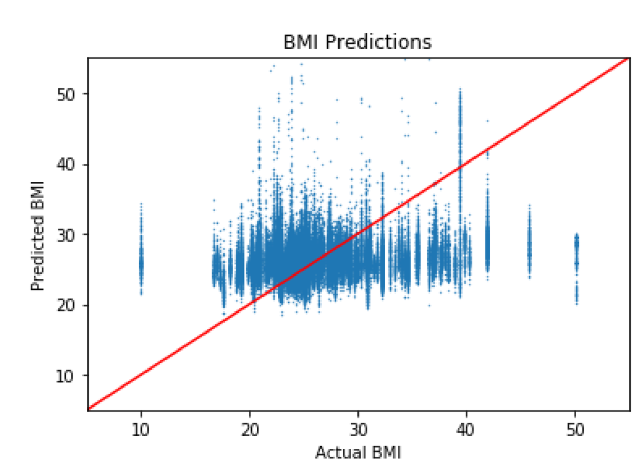
Results

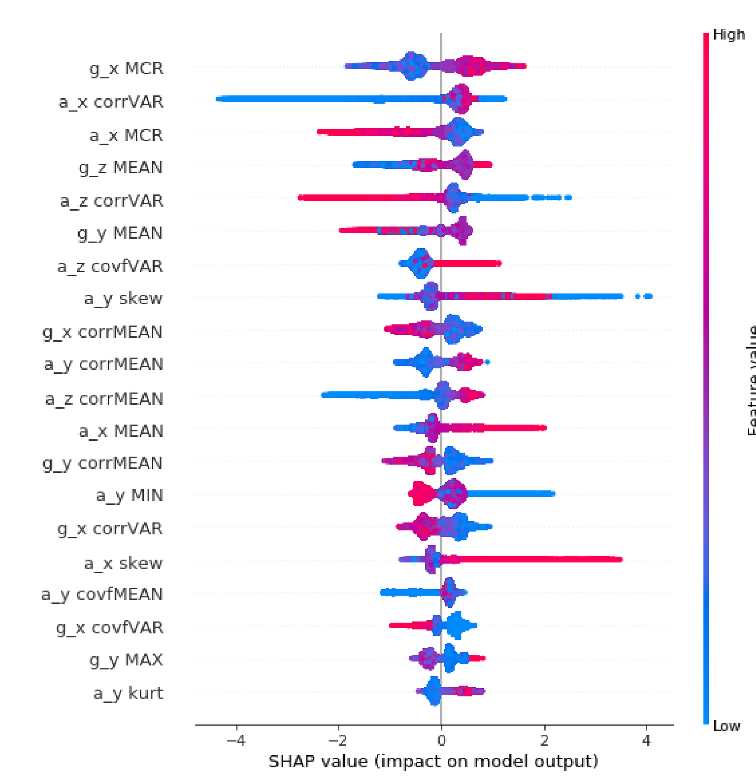
Models are able to achieve better than random performance on all tasks, but are still far from reasonable performance, due to data quality issues and poor model generalization. Deep learning models, while having higher performance, lack the explainability of classical machine learning algorithms.

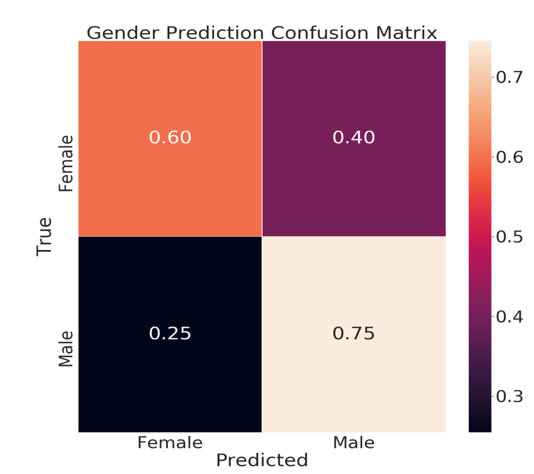






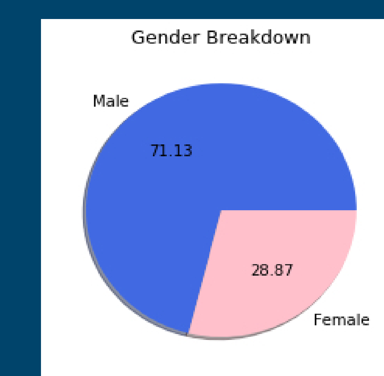


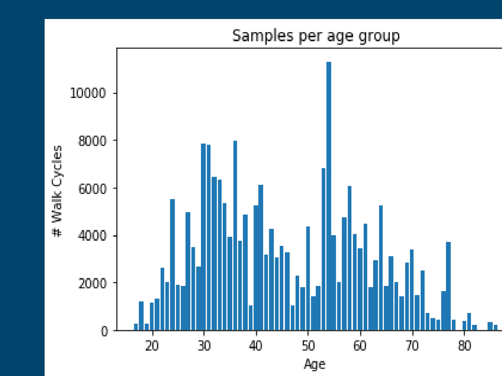


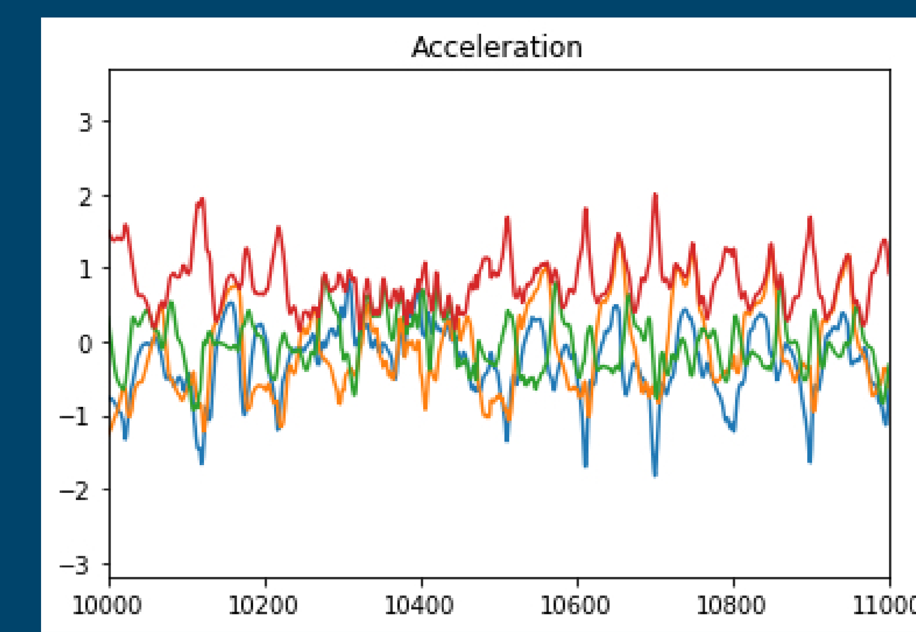


Dataset

- We curated a novel dataset of six minute walk tests collected from 8329 participants in the MyHeart Counts Study.
- Sensor readings were queried from the iPhone's onboard IMU, and reported triaxial acceleometer and gyroscope values, sampled at 100 Hz.
- The data is highly multimodal, with a wide variety of mobile devices, placed at various locations (hip, hand, bag, etc.) on the user, walking on a variety of terrains.







Sample acceleometer readings (10 seconds)

Conclusion

- Distributed mobile diagnostics are a medium with high translational potential, and efficient scalability.
- Large scale introduces data quality concessions, but enables research methodologies which take advantage of massive data quantity.
- Deep Neural Networks can effectively infer demographic information from gait acceleometry.