Deep Learning for Tennis Stroke Classification Using Tech Wearables
Ellen Roper, Tristan Goswalti
Stanford University

Abstract
Motion capture technology offers a myriad of applications in many different fields. TuringSense, a Santa Clara-based tech wearable company, created a product which uses motion capture to train athletes. Standard motion capture is often optical in nature, but TuringSense replaces this approach with digital technology, opting to use motion sensors instead. We used data collected with TuringSense’s sensors to create a deep learning model which can predict and classify a tennis player’s stroke. We trained a simple baseline model with one hidden layer, a deep network with four hidden layers, and a convolutional neural network on a multi-label stroke classification task, obtaining consistently high classification accuracy for five different tennis strokes on each, with the convolutional neural network achieving the highest accuracy.

Procedure
- Raw data converted to angle data by TuringSense’s algorithms.
- Extract individual strokes from metadata about stroke start/end.
- We pad each stroke with zeros to ensure all strokes are the same length, (determined by pre-analysis of our data).
- Sensor data at each timestamp is stacked vertically and appended into an 18 data point x 111 timestamps ‘image.’
- We run this image through a convolutional neural network to output a one-hot vector with the predicted classification. Our five stroke type classifications are:
  Class 1: Forehand Slice
  Class 2: Forehand Topspin
  Class 3: Backhand Slice
  Class 4: Backhand Topspin
  Class 5: Forehand Slice

Discussion and Results
The first models we tried were baseline neural networks, first with a single hidden layer and secondly with four hidden layers. Both were able to achieve high classification accuracy. However, due to constraints in data size, we decided to implement a CNN, as this likely will be better at accounting for variability in data if more is collected. This CNN performs consistently better than the simpler deep networks, with an average classification accuracy of 0.98 on the test set. As we run the model with augmented data of increasing variability (as would be seen in real life), the CNN continues to outperform other models, giving our best result.

Future Modifications and Special Thanks
Physical constraints requiring consistent repetitive motion of tennis strokes, fatigue, and considerable time means that we have a severe lack of data. Future work could involve:
- Collecting more data to further train and refine the model.
- Adding variability with data from more than one professional athlete.
- Expand the number of strokes classifications, as there are more than five strokes which are used in tennis.
We would like to thank TuringSense for their enthusiasm and support for our project. Please visit their website for more information and for updates on their amazing innovations.

References

Contact
Ellen Roper
Stanford University
Email: emper@stanford.edu
Tristan Goswalti
Stanford University
Email: tgoswalti@stanford.edu
TuringSense Website: www.turingsense.com