Computer Vision: Golf Pro

Parker Sullins



Abstract

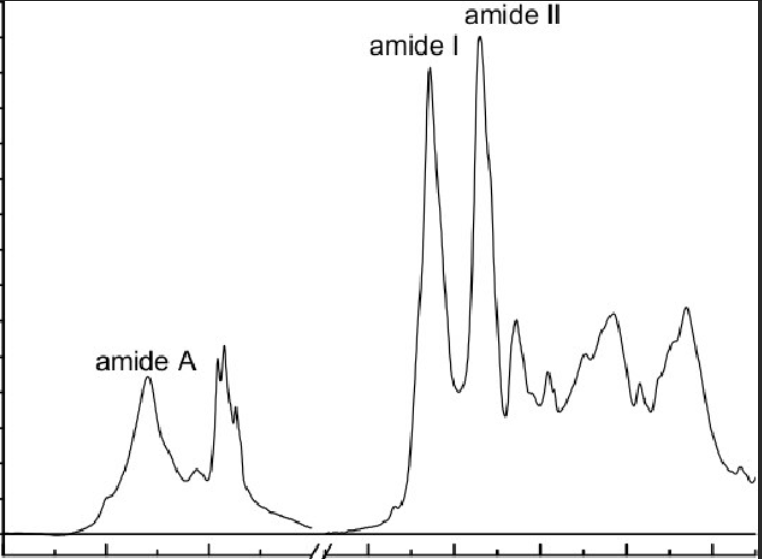
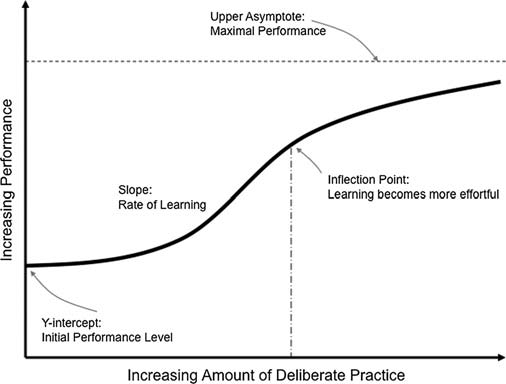
in this paper, we attempt to rectify many of the pitfalls that come with learning golf. Using computer vision and machine learning, we aimed to produce full 3D tracking of key skeletal points with the intent of providing insights on how a beginner golf compares to a pro on key metrics. For this study, we focused solely on the driver, the longest club in a player’s arsenal, and the most used club. Using a pre-trained machine learning, we tracked 33 key points, each generating four features, an X, Y, Z coordinate tuple and a visibility metric. This generated 133 features to extract per a frame. We also attempted to normalize the total frames of a swing so that each swing took approximately three seconds and was composed of 180 frames. Using the 133 key point features and the current frame feature, we were able to accurately predict whether a golf was an amateur or a pro with 98% accuracy on testing data. The models we tested were Linear Regression, Gradient Boosting Classifier, and Random Forest. Random Forest performed the best and was selected as the preferred model for classifying golfers as either Pro or Amateur. The goal of this research was to explore whether a machine learning model could successfully identify a highly skilled players swing versus a low skill players swing, and if so, understand why.

1 Introduction

Currently, golf is a highly prohibitive sport to newcomers. Beginners must shell out hundreds of dollars for clubs and equipment, and that is all before having ever swung a golf club. As most beginners soon find out, golf is not only expensive, but requires a large amount of time and effort to obtain reasonable competency. This issue is only worsened by the nature of golf as an athletic venture. Often the only feedback a player receives is watching the flight of the ball after impact. It often takes weeks for a beginner to even begin striking a ball consistently, let alone possessing any accuracy. Below is a typical learning curve one could expect when attempting to gain skill in an area, and to the right is the typical learning curve one might experience when learning golf.

Golf Learning Curve

Typical Learning Curve

­

Time

Perceived Skill



Figure 1: Typical learning curve for skill acquisition

Golf is an atypical sport, and the learning curve reflects this. The learning curve of golf is steep, unpredictable, and highly discouraging. The most common method to reduce early friction when learning golf is to higher an instructor. Cost of an instructor typically range from $75 to $150, and players can only expect results after numerous lessons. We hope to change the seemingly insurmountable challenges of learning golf for the layman. We hope to be able to provide personal insights into mistakes golfers are making so they may correct their swing without the assistance of a trainer. Ideally, players would be able to record their swing from their smartphones, upload them through a mobile app, and receive feedback on their swing after the video is processed on a cloud server. This would greatly reduce the cost of learning golf, while providing exceptional feedback that could only otherwise be accessible by paying for a coach.

2 Team Organization

Team organization consisted of solo research performed by Parker Sullins. Additionally, several amateur golfers volunteered their time, and without them this would not have been possible. Also, special thanks to Taylor Made for providing excellent footage of professional golfers.

3 Related Work

Currently single camera swing analysis is in the research phase, and no current products can provide motion tracking feedback on a player’s body orientation relative to the golf ball. However, sensor based three-dimensional tracking systems are available. Currently sensor-based tracking consists of attaching sensor to locations throughout the body and the golf club. Using multiple high-speed cameras, they can track player movement exceptionally well using motion tracking technology developed for film making. The draw backs of this system are that a user must attach numerous tracking nodes which can be cumbersome when performing athletic movements. Additionally, they require several high-speed specialty cameras, and require a studio that must be professionally set up to make the environment acceptable for the systems software to run correctly. These systems range from tens of thousands of dollars on the low side, to upward of one hundred thousand dollars. This means such systems are practically impossible to obtain by the average golfer.



Figure 2: The OPTITRACK'S GEARS GOLF SYSTEM

Notice the multiple highspeed cameras in red and motion tracking body suit, hat, and shoes in green.

There also exist flight tracking systems that have been on the market for years and perform competently when simulating the flight path and distance of a golf ball after impact. These systems range from hundreds to thousands of dollars and are more obtainable to the average golfer. However, these systems fail to address the underlying issue that beginner golfers face. These systems do not provide feedback on swing form or tell a user how to correct their swing.

Academic Research in this field is ongoing, and several studies exist which try to solve a similar problem using machine learning and computer vision. These papers use computer vision to classify the golf swing into various states, and to identify which state a golfer is in throughout a swing. The goal of this is to see if machine learning can correctly identify the key sequences in a golf swing. This research is not geared at helping golfers and is only concerned with motion tracking of the golf sequence, and still makes use of a motion capture suit. Additionally, several papers attempt to use inertial sensors to measure physical characteristics of a golfer’s swing, such as angular velocity and gyroscope alignment of a golfer body throughout the swing. These papers were successful at motion tracking and digital signal processing of the golf swing, but they were unsuccessful at understanding what makes a good versus bad golf swing, and do not provide feedback to a player. Currently, this is the only research into golf motion tracking using only computer vison, without use of a motion capture suit.

4 Data Acquisition and Processing

4.1 Data Collection

In this paper, golf swing data was collected from various volunteer, amateur golfers who have some experience with golf, but have never has formal instruction. Data collection from the amateur golfers consisted of a single iPhone 10X camera, utilizing 60 fps in MOV format. This format is commonly used in smartphones as it allows for better compression and video quality when compared to MP4. Videos were taking approximately 17 feet behind the golfers and slightly off center to the right. This was done to best match the professional footage of pro golfers produced by Taylor Made’s golfing YouTube channel. Twelve amateur swings were recorded from three different golfers. Additionally, twelve professional swings were downloaded as MP4’s and used as the data for the professional swing. All videos were converted to MP4’s for compatibility with the OpenCV library. Videos were slowed or sped up so that all swings took approximately three seconds to complete. The videos were then exported at 60 frames per second so that each video was segmented into roughly 180 frames each. All videos were then fit to a 9:16 ration, and golfers were placed squarely in the middle, with similar distance from the golfer to the edges for all videos. Data was collected on each key point generated by Google’s MediaPipe, a machine learning library for live and streaming media and mapped to the appropriate frame. Data generated by MediaPipe’s model was then fit with a supervised Random Forest Classifier. The classifier optimized a solution for a binary dependent variable, Professional or Amateur golfer, and reported a probability for that video being of a pro or amateur.

4.2 Data Description

Data was collected from thirty-three key point body segments and mapped to the frame in which they occurred. Each point was composed of a four-part tuple, an X, Y, and Z geometric coordinate and visibility estimation. X and Y coordinates are generated relative to the image dimensions and the Z coordinate is estimated by the model. Visibility represents the confidence in which the model feels it can accurately pinpoint where a certain segment on the body is at each frame.

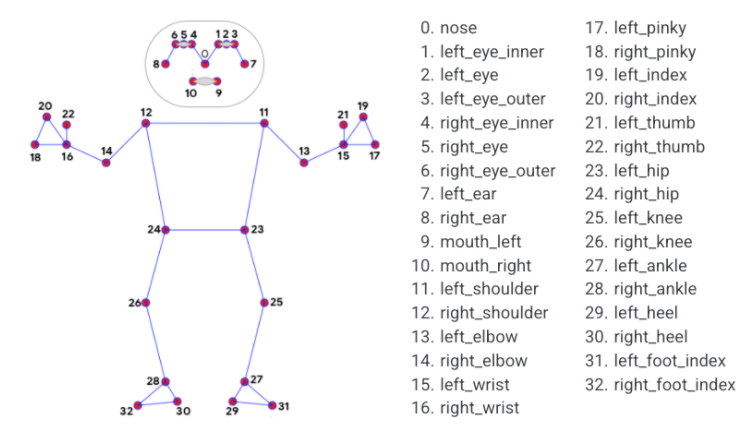
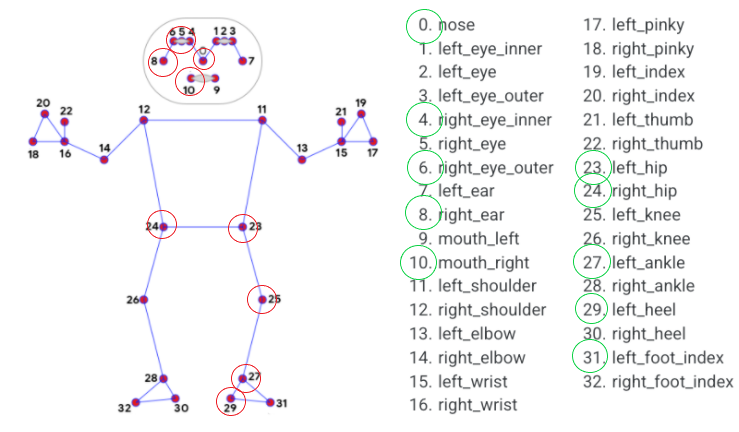


Figure 3: A diagram showing where each feature maps to in the MediaPipe Pose Model

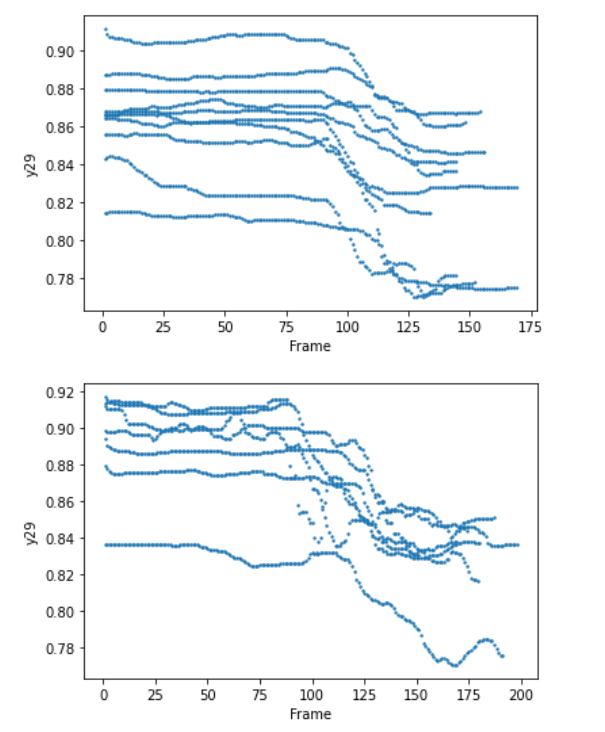
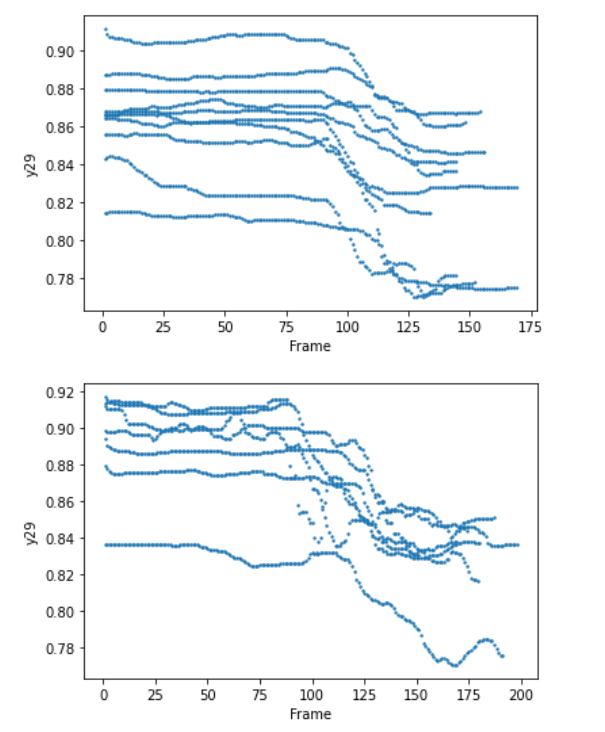
We collected over 3,000 rows of data for each of the 134 features, generating over 1,000,000 data points.

4.3 Visualization

Data visualization was used to explore relationships between various features and how they were classified by the Random Forest model. Feature Importance was calculated and used to determine which features to explore based on their relative weight in the model. Below is a illustration of the 10 key points with the highest contribution to the classification of a swing.



The results of the feature importance analysis are somewhat unexpected. For example, no feature related to the upper body is recorded until the 40th most important feature, and no upper body feature contributed more to the model than its relative size in the feature set, i.e. 1:134. The most important feature was the left heel, which is unexpected as the left heel is typically stationary throughout the golf swing with little to no movement.



Amateur Golfer

Pro Golfer

Keep in mind that movement is calculated relative to the joints position in the 1020px by 720px frame, and thus the physcial coordinate is not as important as the pattern generated by the movement. Here we have the left heel’s change in the y-axis through out the swing. Intrestingly, I do not see anything that seems to be a signifigant difference between the pro and amatuer, in fact, the amatuers seem to have smoother movements throughout the swing. Yet, this feature contributes over 6% to the models decision making.

5 Software and Developing Tools

5.1 Software

5.2 Computer Setup

5.3 Sensor Hardware

6 Milestones

7 Results and Discussion

8 Conclusion

9 Future Work

10 References