CaddieSet: A Golf Swing Dataset with Human Joint Features and Ball Information

Supplementary Material

A. Extracting Joint Coordinates: Model Setup

In this study, we used the SwingNet architecture, which is a combination of Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) with a lightweight deep neural network architecture. The goal of this architecture is to detect and predict human postures during golf swing events. To detect eight key golf swing events, we employed a sequence mapping model fine-tuned by GolfDB and CaddieSet, which achieved an accuracy of 78.0%. This represents an improvement of approximately 2% over the baseline SwingNet model. Notably, accurately labeling the Address and Finish events was particularly challenging, as previously noted in [16]. This challenge arises due to the subjective nature of labeling and the inherent difficulty in precisely localizing these events temporally. After excluding the Address and Finish events, the accuracy for detecting the remaining six events increased to 94.1%, compared to 91.8% with the vanilla SwingNet model. Additionally, when testing on the MS COCO val2017 dataset [14], the Faster R-CNN detector achieved a human average precision (AP) of 56.4. However, when paired with HRNet for pose estimation, the AP increased significantly to 74.9, which demonstrates the robustness of the models used for joint coordinate extraction, ensuring that the extracted joint data is reliable for further analysis.

B. Detailed Feature Analysis for Swing-Related Metrics

Based on domain knowledge, we utilized 15 swing-related metrics to extract features for each of the eight golf swing events. This allowed us to generate a total of 40 swing-related features. These features capture the joint angles and movements of various body parts during the swing, with each feature assigned to a specific phase of the swing. For example, the STANCE-RATIO is measured during the Address phase to capture the golfer's initial stance, while HIP-ROTATION is measured up to the Impact phase to track pelvic rotation. These 40 features were analyzed in relation to BallSpeed and other key metrics, and the influence of each feature on the swing performance was evaluated.

B.1. Key Swing Metrics and Their Role

The 15 key metrics are crucial for understanding the dynamics of a golf swing. For example, metrics such as SHOULDER-ANGLE, HIP-ROTATION, and WEIGHT-SHIFT are vital for generating power and controlling the swing's direction. Meanwhile, HEAD-LOC, SHOULDER-LOC, and STANCE-RATIO ensure that the swing remains consistent and accurate, facilitating better ball striking.

B.2. Visualization of Features in the Swing Process

We present a detailed analysis of how these features interact and affect ball trajectory. The heatmap in Figure 5 demonstrates the correlation between swing-related features and BallSpeed, showing that features such as STANCE-RATIO and SHOULDER-ANGLE are strongly correlated with ball speed and distance.

C. Benchmark Evaluation: Model Setup

For benchmarking, we employed various vision-based models including ResNet18, MobileNet_V3, and ViT-B/16. These models were chosen for their widespread use in computer vision tasks and their expected strong performance when adapted to our swing videos. Specifically, we concatenated image patches from eight swing sequences into a single input, forming a combined 320×160 input image, which was processed by the models for prediction. Each image represents one stage of the swing, capturing the progression over time. The benchmark models were evaluated using various metrics such as Accuracy (Acc), Area Under the ROC Curve (AUC), and Mean Squared Error (MSE), as shown in the table S.1.

Table S.1. Hyperparameter settings for the machine learning methods used in the experiments.

Methods	Hyperparameters
LR	penalty='l2', C=1.0, solver='lbfgs'
SVM	C=1.0, kernel='rbf', probability=True
RF	n_estimators=100, max_features='sqrt'
XGBoost	use_label_encoder=False, eval_metric='logloss'

D. Down-the-Line (DTL) View Analysis

D.1. Data Collection from the DTL View

For a more comprehensive analysis, we also collected data from the Down-the-Line (DTL) view. This view allows us to capture key joint movements from a different perspective, providing additional insights into the swing mechanics. To extract swing-related features from the DTL view, we utilized nine metrics, which are table S.2. The swing-related features extracted from the DTL view also demonstrate a close relationship with ball information, as illustrated in Figure S.1

D.2. DTL View Experimental Results

CaddieSet for the DTL view comprises 833 samples, with 666 samples used for training and 167 for testing. Benchmarking these features showed significant correlations with ball speed, as detailed in the table S.3.

Table S.2. Description of the 9 metrics for DTL view: Each indicator measures specific joint information in the golf swing. These metrics help evaluate and improve the efficiency and effectiveness of the swing. They are based on insights from domain experts.

Feature	**Description**	**Measurement**
SPINE-ANGLE	Spine angle relative to horizontal	degree
LOWER-ANGLE	Angle formed by right pelvis, knee, and ankle	degree
SHOULDER-ANGLE	Shoulder angle relative to horizontal	degree
LEFT-ARM-ANGLE	Angle formed by left shoulder, elbow, and wrist	degree
RIGHT-ARM-ANGLE	Angle formed by right shoulder, elbow, and wrist	degree
HIP-LINE	Movement of hip relative to Address	ratio
HIP-ANGLE	Rotation degree of pelvis relative to Address	degree
RIGHT-DISTANCE	Gap between right elbow and the torso	ratio
LEFT-LEG-ANGLE	Angle formed by left pelvis, knee, and ankle	degree

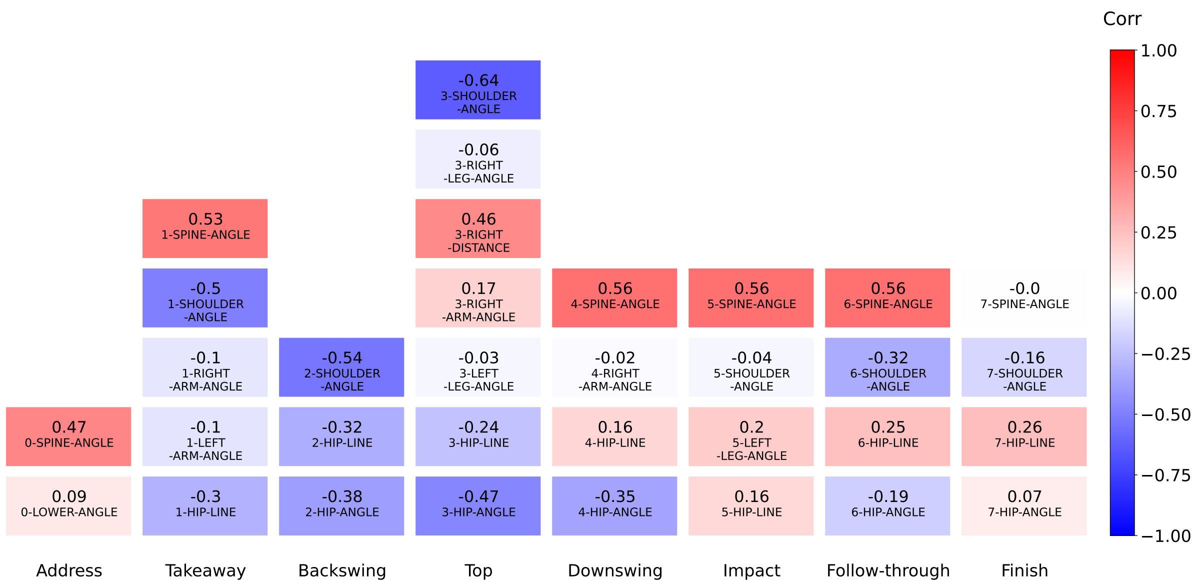

Figure S.1. Correlation heatmap between swing-related features from DTL view and BallSpeed. For example, 1-SPINE-ANGLE shows a significant positive correlation of 0.53. This angle, measured during the *Takeaway*, is crucial for maintaining a stable swing plane and generating torque, contributing to increased ball speed. Additionally, hip-related features such as 3-HIP-ANGLE and 4-HIP-LINE show substantial correlations with ball speed, highlighting their importance in ensuring effective hip rotation and position during the swing.

Table S.3. Benchmark comparison of model performance on the target variables DirectionAngle, SpinAxis, and BallSpeed on DTL view. The best experimental results are in bold.

Method	DirectionAngle		SpinAxis		BallSpeed
	Acc	**AUC**	**Acc**	**AUC**	**MSE**
ResNet18	**0.9162**	0.8190	0.7186	0.7200	74.74
MobileNet_V3	0.8922	0.7962	**0.7246**	0.7267	37.61
ViT-B/16	0.8802	0.7648	0.6527	0.5712	48.27
LR	0.9102	0.8483	0.7126	0.7376	10.27
SVM	0.8743	0.8033	0.6527	0.7214	44.12
RF	0.9042	0.8366	0.7186	0.7176	**7.35**
XGBoost	0.9042	0.8418	0.6946	0.7248	8.25
NAM	0.8623	**0.9025**	0.6048	**0.7434**	9.12