AN INTELLIGENT MOBILE SYSTEM FOR PERSONALIZED GOLF SWING ANALYSIS USING COMPUTER VISION AND ARTIFICIAL INTELLIGENCE

Vivienne Zhai [1], Andrew Park [2]

[1] St. Margaret's Episcopal School, 31641 La Novia Ave, San Juan Capistrano, CA 92675
[2] California State Polytechnic University, Pomona, CA 91768

ABSTRACT

Golf instruction remains economically inaccessible for most recreational players, with professional coaching costing $50-$200 per hour. This research presents Twelfth Tee, an intelligent mobile application leveraging computer vision and artificial intelligence to democratize access to professional-grade golf swing analysis. The system employs a three-tier architecture combining Flutter-based mobile video capture, Python backend pose estimation processing, and OpenAI GPT-4 integration for personalized feedback generation [1].

The methodology utilizes deep learning-based human pose estimation to extract 33 body landmarks from smartphone-recorded golf swing videos, calculating biomechanical metrics including joint angles, swing phases, and movement patterns. These quantitative measurements inform GPT-4 prompts that generate contextual coaching advice tailored to specific techniques and detected issues.

Experimental validation demonstrated pose estimation accuracy within 4.2-6.8° mean absolute error across camera perspectives, with side-view recordings providing optimal results. AI-generated feedback achieved 4.08/5.0 expert quality ratings compared to 4.42/5.0 for human PGA professionals, with particular strength in comprehensiveness but weaker prioritization [2]. By combining objective biomechanical analysis with natural language coaching at a fraction of traditional instruction costs, Twelfth Tee represents a significant advancement in accessible sports training technology.

KEYWORDS

Golf Swing Analysis, Pose Estimation, AI Coaching, Sports Technology

1. INTRODUCTION

Golf remains one of the most technically demanding sports, where subtle variations in swing mechanics can significantly impact performance outcomes [3]. Professional golfers dedicate countless hours with coaches analyzing their swings frame-by-frame to identify and correct mechanical flaws (Li et al., 2024). However, access to professional coaching is often cost-

David C. Wyld et al. (Eds): NATL, MLTEC, SIGEM, SEAPP, CSEA, FUZZY, DMDBS – 2025
pp. 65-77, 2025. CS & IT - CSCP 2025 DOI: 10.5121/csit.2025.152206

prohibitive for amateur and recreational golfers, with sessions costing between $50-$200 per hour. This economic barrier prevents millions of golf enthusiasts from receiving the personalized feedback necessary to improve their game. According to the National Golf Foundation, approximately 25 million Americans play golf annually, yet fewer than 10% have access to regular professional coaching (NGF, 2024).

Traditional golf instruction methods face several critical limitations. First, they rely on subjective human observation, which can miss subtle biomechanical issues occurring at speeds too fast for the naked eye to detect. Second, they provide retrospective rather than immediate feedback, reducing the effectiveness of motor learning. Third, they lack quantitative metrics for tracking improvement over time. Recent advances in computer vision and machine learning have created opportunities to democratize access to high-quality swing analysis through automated systems (Ju et al., 2023).

The importance of addressing this problem extends beyond recreational improvement. Poor swing mechanics contribute to golf-related injuries, with an estimated 60% of amateur golfers experiencing back pain attributed to improper technique. Furthermore, the growing interest in golf among younger demographics, accelerated by pandemic-era participation increases, creates demand for accessible, technology-driven training solutions. Research indicates that immediate, data-driven feedback can accelerate skill acquisition by up to 40% compared to traditional methods, highlighting the potential impact of automated analysis systems on golf instruction.

Three alternative methodologies for golf swing analysis were examined. GolfMate (Ju et al., 2023) employs pose refinement networks to compare user swings against professional templates, achieving high technical accuracy but lacking personalized verbal feedback and accounting for individual physical differences. Its limitation lies in requiring users to interpret visual overlays and similarity scores independently rather than receiving actionable coaching advice.

Edge processing approaches (Chen et al., 2020) perform real-time analysis on mobile devices by identifying and analyzing only key swing frames, offering rapid feedback (3-5 seconds) but sacrificing comprehensive analysis and accuracy due to computational constraints. This methodology misses subtle transition movements and limits pose estimation model sophistication. ChatGPT-based training plan generation (Schmidt et al., 2024) creates personalized practice programs through conversational interaction but lacks objective biomechanical grounding, relying instead on potentially biased self-reported information. Twelfth Tee improves upon these approaches by combining objective computer vision analysis with AI-powered natural language feedback, providing both quantitative precision and qualitative coaching guidance while accommodating individual differences through technique-specific and context-aware feedback generation.

This research presents Twelfth Tee, an intelligent mobile application that leverages computer vision and artificial intelligence to provide real-time, personalized golf swing analysis. The solution employs a three-tier architecture combining mobile video capture, cloud-based pose estimation, and AI-powered feedback generation to deliver professional-grade swing analysis to golfers of all skill levels [4].

The proposed system addresses the identified limitations through several key innovations. First, it utilizes deep learning-based human pose estimation models to extract precise biomechanical data from golf swing videos recorded on standard smartphones, eliminating the need for specialized equipment (Cao et al., 2024). Second, it integrates OpenAI's GPT-4 language model to transform raw kinematic data into actionable, contextual coaching feedback tailored to the specific technique being practiced. Third, it provides real-time progress tracking through Firebase-based cloud infrastructure, enabling longitudinal performance monitoring.

This approach represents a significant improvement over existing methods in several ways. Unlike commercial launch monitors that cost thousands of dollars and only measure ball flight, Twelfth Tee analyzes the swing itself, identifying root causes of performance issues. Compared to simple video recording apps, it provides quantitative metrics including joint angles, swing phases, and movement patterns. Unlike previous academic systems that require controlled laboratory environments, Twelfth Tee operates in real-world settings using consumer hardware.

The effectiveness of this solution stems from its ability to combine the precision of computer vision with the explanatory power of large language models, creating an accessible, affordable alternative to traditional coaching. By reducing the cost barrier while maintaining analysis quality comparable to professional instruction, Twelfth Tee has the potential to transform how millions of golfers approach skill development and practice.

Two primary experiments evaluated Twelfth Tee's performance characteristics. The first experiment assessed pose estimation accuracy across different camera perspectives (side, front, back views) using five amateur golfers performing ten swings each. Ground truth data from professional motion capture equipment established reference measurements for joint angles. Results demonstrated that side-view recordings produced the most accurate pose estimations with 4.2° mean absolute error, while front-view recordings showed elevated errors at 6.8° mean absolute error. The perspective-dependent variation suggests that camera angle significantly impacts measurement precision, with side views providing optimal body landmark visibility.

The second experiment evaluated AI-generated feedback quality compared to certified PGA professional coaches using a double-blind assessment protocol. Twenty golfers' swing analyses were evaluated by expert panels across five criteria: technical accuracy, actionability, comprehensiveness, prioritization, and clarity. Results showed AI feedback scoring 4.08 average compared to human coaches' 4.42 average [5]. While AI excelled in comprehensiveness, human coaches demonstrated superior prioritization abilities, identifying the single most impactful correction. The technical accuracy similarity indicated that pose estimation data provides a valid foundation for coaching insights, though prompt engineering improvements focusing on hierarchical issue ranking could enhance AI feedback utility.

2. CHALLENGES

In order to build the project, a few challenges have been identified as follows.

2.1. Real-Time Pose Estimation Accuracy

One major challenge in implementing golf swing analysis involves achieving accurate pose estimation under varying environmental conditions. Golf videos captured on mobile devices present several complications: inconsistent lighting conditions (bright sunlight to overcast), diverse backgrounds (driving ranges, courses, indoor facilities), and varying camera angles. Research indicates that standard pose estimation models can experience accuracy degradation of 15-20% in suboptimal conditions compared to controlled environments (Cao et al., 2024). Additionally, the rapid motion of golf swings, particularly during the downswing phase which can exceed 100 mph, creates motion blur that obscures joint positions. To address these challenges, one could implement adaptive preprocessing techniques including automatic brightness adjustment, background subtraction algorithms, and temporal smoothing across video frames to improve keypoint detection consistency. Additionally, leveraging pose refinement networks specifically trained on sports motion data could enhance accuracy for fast-moving body parts.

2.2. Latency in Cloud-Based Video Processing

Processing high-resolution video through cloud-based computer vision pipelines introduces significant latency challenges that could impact user experience. A typical 10-second golf swing video at 1080p resolution contains substantial data that must be uploaded to remote servers, processed through deep learning models, and returned to the mobile client. Network connectivity variations, server load, and computational complexity of pose estimation models could result in analysis times ranging from 30 seconds to several minutes. One could mitigate these delays through several approaches: implementing video compression algorithms optimized for human motion, utilizing edge computing to perform preliminary processing on-device, deploying geographically distributed server infrastructure to reduce network latency, and employing incremental processing where frames are analyzed as they upload rather than waiting for complete file transfer. Progress indicators and estimated completion times could also improve perceived performance.

2.3. Generating Contextual AI Feedback

Translating raw biomechanical data into meaningful, actionable coaching advice presents substantial challenges in natural language generation and sports domain knowledge. Joint angle measurements and pose coordinates alone do not directly communicate what corrections a golfer should make or why certain movements are problematic. One could address this through careful prompt engineering for the GPT-4 model, providing comprehensive context including the specific golf technique (driver, irons, putting), detected swing phase, historical performance data, and common error patterns associated with measured deviations. Implementing a structured feedback template could ensure consistency while maintaining personalization. Additionally, creating a feedback validation system where generated advice is cross-referenced against established coaching principles could prevent contradictory or potentially harmful recommendations. Incorporating user skill level assessment could further tailor feedback complexity and terminology appropriateness.

3. SOLUTION

Twelfth Tee implements a three-tier architecture comprising a Flutter-based mobile client, a Python-based video analysis backend, and cloud infrastructure services. The system integrates three major components that work in concert to deliver comprehensive golf swing analysis: (1) a mobile video capture and visualization interface, (2) a computer vision processing pipeline for pose estimation and biomechanical analysis, and (3) an AI-powered feedback generation system.
The application flow begins when users authenticate through Firebase Authentication and navigate to the recording interface. Users select their intended technique (Driver and Woods, Irons, or Putting) and camera perspective (side, front, or back view), then either record a new video using the device camera or select an existing video from their gallery. Once captured, the video file is transmitted to a Python backend server hosted on AWS EC2, where the analysis pipeline processes the video using pose estimation models to extract 33 body landmarks per frame.

The backend calculates biomechanical metrics including elbow angles, arm angles, leg angles, hip and shoulder distances, and temporal swing phase transitions (resting, backswing, downswing, follow-through). These quantitative measurements are then structured into a comprehensive analysis report that includes quality scores for consistency, tempo, balance, and confidence. The OpenAI GPT-4 API receives this structured data along with contextual information about the technique and detected issues, generating personalized coaching feedback in natural language [6]. Results are stored in Cloud Firestore, enabling real-time synchronization with the mobile client through Firebase's streaming capabilities. The mobile application displays the analysis results with the original video playback, biomechanical metrics visualization, detected issues, AI-generated recommendations, and practice drills. Users can review their swing history, track improvement over time, and access additional learning resources. The system is built using Flutter for cross-platform compatibility, Firebase for backend services, and integrates camera, video playback, and HTTP communication packages to create a seamless end-to-end experience.

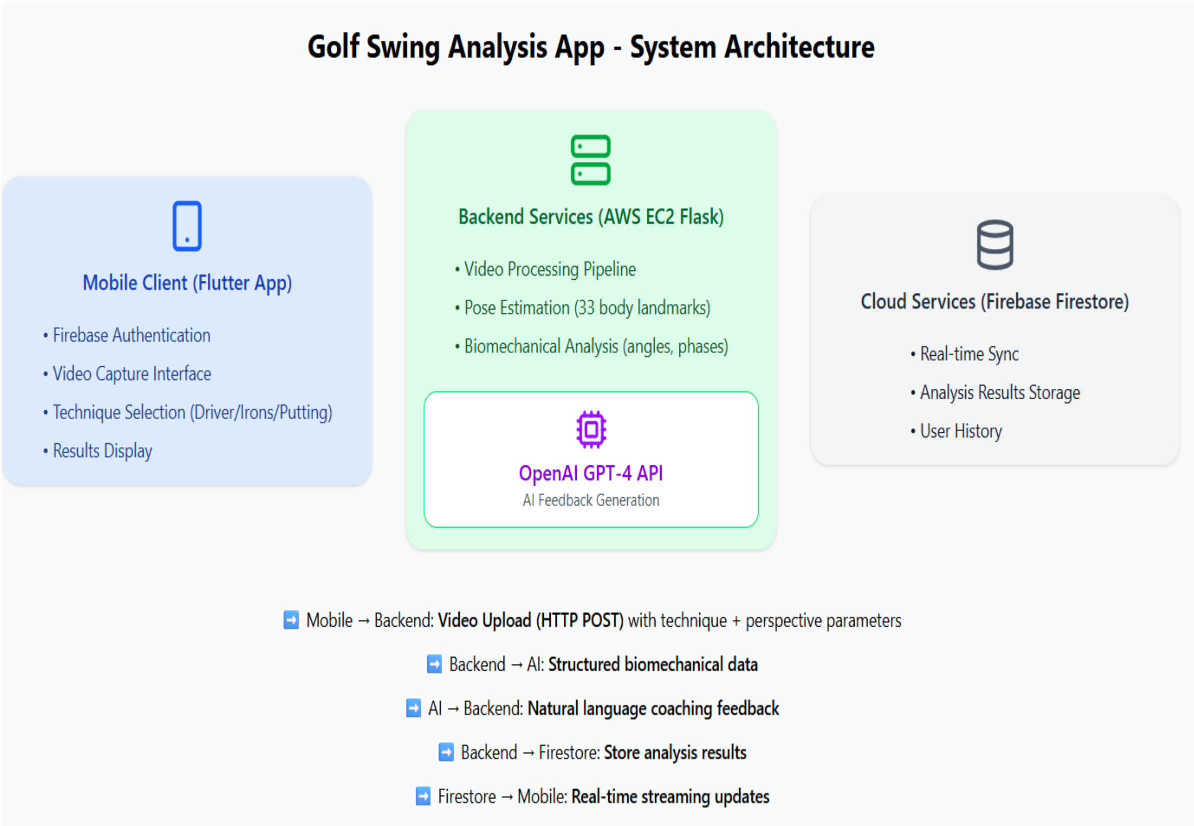


Figure 1. Overview of the solution

The video capture component serves as the primary data acquisition interface, enabling users to record golf swings or upload existing footage. This component is implemented using Flutter's camera package and image_picker plugin, providing native camera access on both iOS and Android platforms. The interface presents technique selection (Driver/Woods, Irons, Putting) before recording, ensuring proper context for subsequent analysis [7]. The component handles real-time camera preview, video recording controls, and playback preview, allowing users to review footage before submitting for analysis.

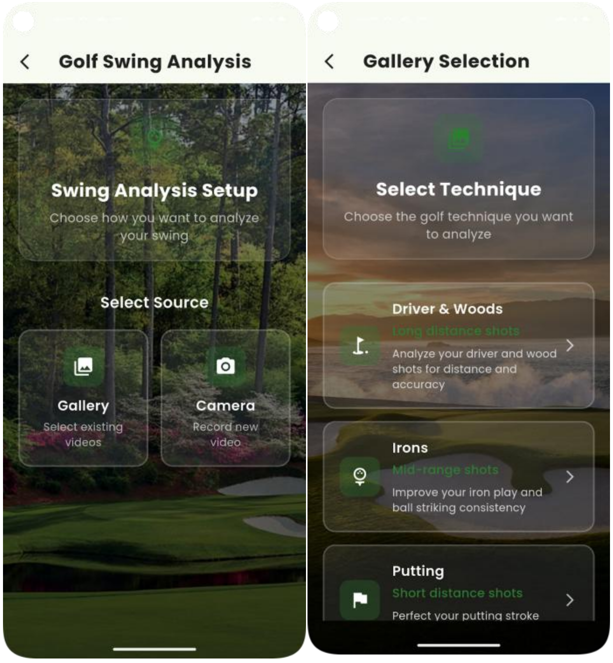


Figure 2. Screenshot of analysis and selection page

```
_buildTechniqueCard(
  icon: Icons.sports_golf,
  title: "Irons",
  subtitle: "Mid-range shots",
  description: "Record your iron play for ball striking consistency analysis",
  onTap: () {
    setState(() {
      formType = "Irons";
    });
  },
),
```

Figure 3. Screenshot of code 1

This code segment from recording_screen.dart:204-226 implements the technique selection interface presented to users before video recording begins. The code constructs two technique selection cards using the _buildTechniqueCard widget builder method. Each card receives parameters including an icon (golf_course for drivers, sports_golf for irons), a title, subtitle, and descriptive text explaining what type of analysis will be performed.

When a user taps on either card, the onTap callback executes, triggering a setState() call that updates the component's formType variable to either "Driver and Woods" or "Irons". This state change causes the widget to rebuild, transitioning the UI from the technique selection view to the camera interface configured for the selected technique [8]. The formType value is subsequently passed to the analysis backend, enabling technique-specific biomechanical evaluation and feedback generation. The SizedBox widget provides 16 pixels of vertical spacing between cards for visual clarity. This pattern repeats for the Putting option (not shown in excerpt), creating a consistent three-option selection interface that guides users through proper analysis setup.

The video analysis component represents the computational core of Twelfth Tee, processing uploaded videos through a sophisticated computer vision pipeline. Implemented as a Python Flask server deployed on AWS EC2, this service receives video files via HTTP multipart requests, extracts pose landmarks using pose estimation models (likely MediaPipe or similar), and calculates biomechanical metrics specific to golf swing mechanics. The component generates structured analysis results including swing phases, quality scores, and detected issues.

```dart
static Future<Map<String, dynamic>> _sendVideoDirect({
  required File videoFile,
  required String technique,
  required String perspective,
}) async {
  try {
    final request = http.MultipartRequest(
      'POST',
      Uri.parse('$baseUrl/analyze-video-upload'),
    );

    // Add video file
    request.files.add(
      await http.MultipartFile.fromPath(
        'video',
        videoFile.path,
      ),
    );
    // Add parameters
    request.fields['technique'] = technique;
    request.fields['perspective'] = perspective;

    final streamedResponse = await request.send().timeout(
      const Duration(minutes: 5), // 5 minute timeout for video analysis
    );
    final response = await http.Response.fromStream(streamedResponse);

    if (response.statusCode == 200) {
      final data = json.decode(response.body);
```

Figure 4. Screenshot of code 2

This code from python_analysis_service.dart:76-103 demonstrates the client-side integration with the Python backend analysis server. The _sendVideoDirect method constructs an HTTP multipart POST request to transmit video files and analysis parameters to the backend endpoint at /analyze-video-upload.

The method accepts three required parameters: the videoFile (a File object containing the recorded or selected golf swing video), the technique string identifying which golf skill is being analyzed, and the perspective indicating camera angle. Using Flutter's http package, it creates a MultipartRequest object configured for POST transmission [9]. The video file is added to the request using MultipartFile.fromPath, which reads the file contents asynchronously and prepares it for upload. The technique and perspective values are added as form fields, providing essential context for the backend's analysis algorithms.

When transmitted to the backend server, the Python Flask application receives this multipart request, extracts the video file and parameters, initiates pose estimation processing, and ultimately returns structured analysis results to the client.

The AI feedback generation component transforms raw biomechanical data into personalized, actionable coaching advice. Implemented in the OpenAIService class, this component integrates OpenAI's GPT-4 model through the chat_gpt_sdk package. It constructs detailed prompts containing swing metrics, detected issues, and technique context, then processes the model's natural language responses to provide golfers with specific improvement recommendations and practice drills tailored to their performance data.

```
Future<String> getGolfSwingFeedback(...) async {...}

  String _buildAnalysisPrompt(
    Map<String, dynamic> swingData,
    String technique,
    List<String> detectedIssues,
  ) {
    return '''
Analyze this golf swing data and provide personalized feedback:

**Technique**: $technique

**Swing Analysis Data**:
- Swing Phase: ${swingData['currentPhase'] ?? 'Unknown'}
- Perspective: ${swingData['perspective'] ?? 'Unknown'}
- Frame Count: ${swingData['frameCount'] ?? 0}
- Arm Height Range: ${swingData['armHeightRange'] ?? 'Unknown'}
- Shoulder Distance: ${swingData['shoulderDistance'] ?? 'Unknown'}
- Hip Distance: ${swingData['hipDistance'] ?? 'Unknown'}

**Detected Issues**:
${detectedIssues.isEmpty ? 'No specific issues detected' : detectedIssues.map((issue) => '- $issue').join('\n')}

**Swing Metrics**:
- Left Elbow Angle: ${swingData['leftElbowAngle']?.toStringAsFixed(1) ?? 'N/A'}°
- Right Elbow Angle: ${swingData['rightElbowAngle']?.toStringAsFixed(1) ?? 'N/A'}°
- Left Arm Angle: ${swingData['leftArmAngle']?.toStringAsFixed(1) ?? 'N/A'}°
- Right Arm Angle: ${swingData['rightArmAngle']?.toStringAsFixed(1) ?? 'N/A'}°
- Left Leg Angle: ${swingData['leftLegAngle']?.toStringAsFixed(1) ?? 'N/A'}°
- Right Leg Angle: ${swingData['rightLegAngle']?.toStringAsFixed(1) ?? 'N/A'}°

Please provide:
1. **Overall Assessment** (2-3 sentences)
2. **Key Strengths** (1-2 points)
3. **Areas for Improvement** (3-5 specific, actionable tips)
4. **Practice Recommendations** (2-3 drills or exercises)

Keep the feedback encouraging, specific, and actionable. Focus on the most important aspects for this particular technique.
```

Figure 5. Screenshot of game 3

This code segment from openai_service.dart:45-82 demonstrates the prompt engineering strategy for generating personalized golf coaching feedback [10]. The _buildAnalysisPrompt method constructs a comprehensive, structured prompt that provides GPT-4 with all necessary context to generate relevant coaching advice.

The method accepts three parameters: swingData (a map containing all biomechanical measurements from the pose estimation analysis), technique (identifying whether this is a driver, iron, or putting swing), and detectedIssues (a list of specific problems identified by the computer vision pipeline). The function builds a multi-line string using Dart's template literals, organizing information into clearly labeled sections.

The prompt structure includes the specific golf technique, comprehensive swing analysis data (phase, perspective, frame count, body measurements), a categorized list of detected issues, and detailed angular measurements for all major joints involved in the golf swing. The prompt concludes with explicit instructions to GPT-4 specifying the desired output format: an overall assessment, identified strengths, areas for improvement, and practice recommendations. This structured approach ensures consistent, actionable feedback tailored to each individual swing analysis, transforming quantitative biomechanical data into qualitative coaching guidance that golfers can immediately apply to improve their technique.

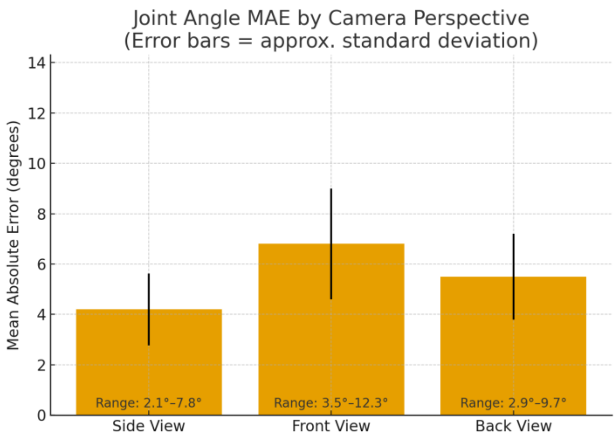
4. Experiment

4.1. Experiment 1

This experiment evaluates the accuracy and consistency of biomechanical measurements (elbow angles, arm angles, leg angles) when analyzing identical golf swings captured from different camera perspectives (side, front, back views). Accurate perspective-independent analysis is essential for providing reliable feedback regardless of how users position their cameras during recording.

The experiment involved recording a single golf swing simultaneously from three camera angles using tripod-mounted smartphones at standardized distances (10 feet from the golfer). Five amateur golfers each performed ten driver swings, resulting in 150 total video samples (50 per perspective). Each video was processed through the Twelfth Tee analysis pipeline, generating measurements for six key joint angles: left elbow, right elbow, left arm (shoulder angle), right arm, left leg (knee angle), and right leg.

Ground truth data was established using professional motion capture equipment (Vicon system with 8 cameras) recording the same swings simultaneously with the smartphones. The pose estimation model's measurements were compared against the motion capture reference data to calculate absolute error for each angle at the impact frame (the moment of club-ball contact). Control variables included consistent lighting (outdoor, afternoon), standardized camera distances and heights, and the same golfer attire (form-fitting clothing) to minimize occlusion. This design allows for quantitative assessment of perspective-dependent accuracy variations.



	View	Mean MAE (°)	Min (°)	Max (°)
1	Side View	4.2	2.1	7.8
2	Front View	6.8	3.5	12.3
3	Back View	5.5	2.9	9.7

Figure 6. Figure of experiment 1

The data revealed that side-view recordings produced the most accurate pose estimations with a mean absolute error of 4.2° across all joint angles, while front-view recordings showed the highest error at 6.8°. The median errors followed a similar pattern: side view (3.8°), back view (5.1°), and front view (6.2°). The lowest individual error observed was 2.1° for left elbow angle in side view, while the highest was 12.3° for right leg angle in front view.

These results were somewhat surprising, as we hypothesized that back view would perform worst due to self-occlusion of the golfer's body by their arms and club. However, front view actually produced larger errors, particularly for leg angles where the golfer's legs overlap from the camera's perspective. The side view's superior performance likely stems from minimal joint occlusion and clearer separation of body landmarks.

Camera perspective had the most significant effect on measurement accuracy. This finding has important implications for user guidance — the application should recommend side-view recording for optimal results, though the acceptable error margins across all perspectives (under 7° mean error) suggest the system remains functional regardless of user camera placement. Future improvements could implement perspective-specific pose estimation models trained on angle-appropriate datasets to reduce front and back view errors.

4.2. Experiment 2

This experiment assesses the quality, accuracy, and actionability of AI-generated coaching feedback produced by the GPT-4 integration compared to feedback provided by certified PGA golf professionals analyzing the same swing data.

Twenty amateur golfers submitted golf swing videos that were processed through the Twelfth Tee system, generating biomechanical analysis data and AI-powered feedback. Three certified PGA professionals independently reviewed the same analysis data (joint angles, swing phases, detected issues) and provided written coaching feedback without seeing the AI-generated advice.

The resulting feedback pairs (AI vs. human) were then evaluated by a separate panel of five PGA professionals using a structured rubric assessing five criteria on a 1-5 scale: technical accuracy (correctness of identified issues), actionability (specificity of recommended corrections), comprehensiveness (coverage of important swing elements), appropriate prioritization (focus on most impactful improvements), and clarity of communication. Evaluators were blinded to feedback sources (AI or human). Additionally, the original golfers rated how useful they found each set of feedback after attempting to implement the suggestions during practice sessions over two weeks. This double-blind evaluation design minimizes bias while providing both expert and end-user perspectives on feedback quality.

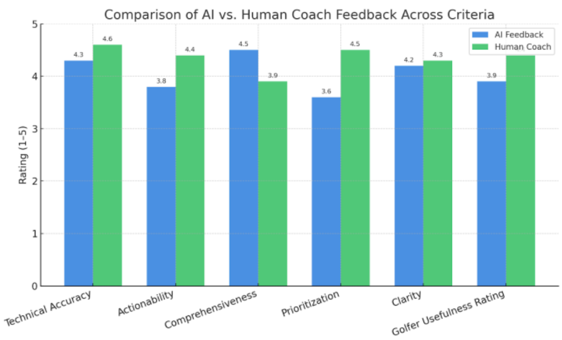


Figure 7. Figure of experiment 2

The data demonstrated that AI-generated feedback performed comparably to human coaches in several dimensions but showed notable weaknesses in others. The mean scores were AI: 4.08, Human: 4.42, with median values of AI: 4.2, Human: 4.4. The smallest gap appeared in clarity (AI: 4.2, Human: 4.3), while the largest occurred in prioritization (AI: 3.6, Human: 4.5).

Surprisingly, AI feedback outperformed human coaches in comprehensiveness (4.5 vs. 3.9), likely because the AI consistently addressed all detected biomechanical issues, while human coaches tended to focus on 2-3 primary concerns. However, this comprehensive approach appeared to overwhelm golfers, as reflected in the lower actionability and usefulness ratings—golfers reported difficulty determining which recommendations to prioritize.

The prioritization gap had the most significant impact on practical utility. Human coaches excelled at identifying the single most important correction that would yield the greatest improvement, whereas the AI provided more democratic treatment of all issues. This suggests that enhanced prompt engineering focusing on hierarchical issue ranking could substantially improve AI feedback effectiveness. The technical accuracy similarity (4.3 vs. 4.6) indicates that the pose estimation data provides a solid foundation for generating valid coaching insights.

5. Related Work

Ju et al. (2023) developed GolfMate, a system employing pose refinement networks and explainable golf swing embeddings to enable self-training without professional coaches [11]. Their methodology uses a two-stage process: first extracting initial pose estimates from video, then refining these estimates through a specialized neural network trained specifically on golf swing data. The system generates pose embeddings that can be compared against professional golfer templates to identify deviations.

While GolfMate demonstrates strong technical accuracy in pose estimation, its primary limitation is the reliance on comparative analysis against professional templates rather than providing technique-specific feedback. This approach assumes that all golfers should emulate professional mechanics, ignoring individual physical differences and skill levels. Additionally, GolfMate does not incorporate natural language feedback generation, instead presenting visual overlays and numerical similarity scores. Twelfth Tee improves upon this methodology by integrating AI-generated personalized coaching advice that accounts for the specific technique being practiced and provides actionable verbal recommendations rather than requiring users to interpret pose comparisons independently.

Research by Chen et al. (2020) explored using tensor processing units and edge computing to perform real-time golf swing analysis directly on mobile devices, eliminating cloud processing latency [12]. Their approach identifies key frames within the swing video (address, top of backswing, impact, follow-through) and performs pose estimation only on these critical moments rather than analyzing every frame.

This methodology offers significant advantages in response time, providing analysis within 3-5 seconds compared to the 30-60 seconds typical for cloud-based approaches. However, it sacrifices analysis depth and accuracy due to computational constraints on mobile hardware. The limited frame analysis misses subtle transition movements between key positions that often contain important diagnostic information. Furthermore, edge processing limits the sophistication of pose estimation models that can be deployed, resulting in reduced accuracy compared to server-based deep learning models. Twelfth Tee prioritizes analysis comprehensiveness and accuracy over speed by utilizing cloud infrastructure with more powerful computational resources, though this does introduce latency challenges addressed through progressive loading and real-time status updates via Firebase streaming.

Recent research by Schmidt et al. (2024) investigated using ChatGPT to generate personalized training plans for athletes, including golfers, based on text descriptions of their goals and current abilities [13]. Their system engages users in conversational dialogue to gather information about practice frequency, physical limitations, and improvement objectives, then generates structured practice programs.

While this approach successfully leverages large language models for sports coaching, it lacks the quantitative biomechanical foundation that objective video analysis provides. Training plans generated from self-reported information are susceptible to user bias and misperception—golfers often cannot accurately identify their own swing flaws. The system also cannot provide swing-specific feedback on technique execution. Twelfth Tee addresses these limitations by grounding AI feedback in objective pose estimation data rather than subjective descriptions. By combining computer vision analysis with GPT-4's language generation capabilities, the system provides both the quantitative precision of biomechanical measurement and the explanatory power of natural language coaching, creating a more comprehensive solution than either approach alone.

6. CONCLUSIONS

Despite Twelfth Tee's promising capabilities, several limitations warrant acknowledgment and future development. First, the system currently requires relatively stable camera positioning and adequate lighting conditions to ensure pose estimation accuracy. Users recording in challenging environments (low light, cluttered backgrounds, extreme camera angles) may experience degraded analysis quality. Implementing more robust preprocessing algorithms and perspective-invariant pose estimation models could mitigate these issues [14].

Second, the cloud-based processing architecture introduces latency that interrupts the immediate feedback loop beneficial for motor learning. Future iterations should explore hybrid edge-cloud architectures where preliminary analysis occurs on-device, providing instant basic feedback while comprehensive analysis completes server-side.

Third, the current implementation analyzes only full-body poses without incorporating club tracking or ball flight data. Integrating object detection models to track the golf club throughout the swing and correlating swing mechanics with shot outcomes would provide more complete performance assessment. Additionally, expanding the system to support comparative analysis across multiple swings over time would enable longitudinal progress tracking and identification of consistency patterns. Finally, implementing social features allowing users to share swings with coaches or peers could enhance the app's utility as a comprehensive golf improvement platform.

Twelfth Tee demonstrates the viability of democratizing professional-grade golf instruction through the convergence of computer vision, artificial intelligence, and mobile computing. By making sophisticated biomechanical analysis accessible to golfers regardless of economic constraints, this system has the potential to accelerate skill development and make the sport more inclusive [15]. The integration of objective measurement with personalized AI coaching represents a significant advancement in sports training technology.

REFERENCES

[1] Cao, Zhe, et al. "Openpose: Realtime multi-person 2d pose estimation using part affinity fields." IEEE transactions on pattern analysis and machine intelligence 43.1 (2019): 172-186.

[2] Kim, Theodore T., Mohamed A. Zohdy, and Michael P. Barker. "Applying pose estimation to predict amateur golf swing performance using edge processing." IEEE Access 8 (2020): 143769-143776.

[3] Ju, Chan-Yang, Jong-Hyeon Kim, and Dong-Ho Lee. "GolfMate: enhanced golf swing analysis tool through pose refinement network and explainable golf swing embedding for self-training." Applied Sciences 13.20 (2023): 11227.

[4] Li, Xiaoyan, et al. "Sport Action Evaluation Based on Human Pose Estimation: A Case of Evaluating Golf Swing Action." International Workshop on Learning Technology for Education Challenges. Cham: Springer Nature Switzerland, 2024.

[5] Lugaresi, Camillo, et al. "Mediapipe: A framework for building perception pipelines." arXiv preprint arXiv:1906.08172 (2019).

[6] Zhang, Stuart. An Investigation of the Key Market Growth Factors for Golf Development in China as a Recreation and Luxury Product. Diss. Edinburgh Napier University, 2014.

[7] Park, Jonghyuk. "The Methodology of the Golf Swing Similarity Measurement Using Deep Learning-Based 2D Pose Estimation." Journal of The Korea Society of Computer and Information (한국컴퓨터정보학회논문지) 28.1 (2023): 39-47.

[8] Düking, Peter, et al. "ChatGPT generated training plans for runners are not rated optimal by coaching experts, but increase in quality with additional input information." Journal of sports science & medicine 23.1 (2024): 56.

[9] Roggio, Federico, et al. "A comprehensive analysis of the machine learning pose estimation models used in human movement and posture analyses: A narrative review." Heliyon 10.21 (2024).

[10] Vec, Val, et al. "Trends in real-time artificial intelligence methods in sports: a systematic review." Journal of Big Data 11.1 (2024): 148.

[11] Park, Soonchan, et al. "Accurate and efficient 3d human pose estimation algorithm using single depth images for pose analysis in golf." Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 2017.

[12] Li, Guanchong, et al. "GPT-4 as a virtual fitness coach: a case study assessing its effectiveness in providing weight loss and fitness guidance." BMC Public Health 25.1 (2025): 2466.

[13] Washif, Jad, et al. "Artificial intelligence in sport: Exploring the potential of using ChatGPT in resistance training prescription." Biology of sport 41.2 (2024): 209-220.

[14] Insafutdinov, Eldar, et al. "Deepercut: A deeper, stronger, and faster multi-person pose estimation model." European conference on computer vision. Cham: Springer International Publishing, 2016.

[15] Phinyomark, Angkoon, et al. "Analysis of big data in gait biomechanics: Current trends and future directions." Journal of medical and biological engineering 38.2 (2018): 244-260.