Proceedings of the 8th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/2026.35989

AI Billiard Coach: Real-Time Shot Assistance System

Qihao Chen

Beijing 101 Middle School, Beijing, China
cqh081019@163.com

Abstract. This study addresses the democratization of professional billiards training through the use of artificial intelligence and camera equipment. The study explores three core issues: (1) how to accurately detect the layout and specific positions of the balls and the table in varying lighting conditions from photos; (2) how to use generative AI to generate physics-based shooting suggestions; and (3) how to optimize the user interface for beginner accessibility. In terms of methodology, the system integrates computer vision algorithms based on OpenCV and constrained GPT-4 prompts, and combines the principles of classical collision mechanics. Experiments were conducted on 120 images submitted by users, and the accuracy of ball detection under medium lighting (300 lux) was 89.2%, while the consistency of the AI-generated suggestions with professional benchmarks reached 76.3%. The main limitations include angular errors caused by parallax (±4.7°) and color recognition failure of dark balls under warm lighting. The study concludes that AI training based on smartphones can effectively reduce the professional barriers of billiards, and future work will prioritize real-time video analysis and mobile deployment.

Keywords: Computer Vision, Generative AI, Billiard Physics, Sports Pedagogy, Human-AI Interaction

1. Introduction

To successfully master the sport of billiards, one needs to possess precise geometric calculation skills and arm strength control abilities. Motor learning in precision sports like billiards imposes significant cognitive demands, as novices must translate abstract spatial angles into precise, coordinated physical movements [1]. Yet approximately 70% of beginners become frustrated due to mistakes in angle calculations and give up practicing within six months [2]. This high dropout rate aligns with broader findings in motor skill acquisition, where early-stage frustration and the inability to establish reliable perception-action couplings are recognised as primary barriers to prolonged sports engagement [3]. Traditional solutions are too costly, with professional coaches in first-tier cities charging over $200 per hour, and billiard tables equipped with sensors costing over $15,000. This economic barrier contributes to structural inequality in sports coaching, disproportionately limiting access to high-quality instruction for recreational athletes and underserved communities [4].

The gap in accessibility has prompted the development of an artificial intelligence billiard coach that can transform consumers' smartphones into an expert advisor for billiards. Recent advances in mobile computer vision have demonstrated the viability of smartphones as ubiquitous platforms for real-time sports performance analysis, effectively democratising access to elite-level biomechanical feedback [5]. Leveraging the latest breakthroughs in multimodal artificial intelligence [6], this system addresses the need for dedicated hardware, while providing feasible hitting guidance in natural language form.

This study brings three important innovations to sports technology: (1) applying generative artificial intelligence based on physics. Stemming from a paradigm recently proven effective in constraining AI outputs to physical realities [7], this research touches on the field of spatial analysis; (2) applying an unmarked computer vision process with robustness to real-world light changes; (3) a democratized framework for converting professional skills into accessible guidance. This article details the technical architecture of the system, experimental verification under different conditions, and demonstrations of the use of artificial intelligence to assist skill improvement. The subsequent sections will explore related literature, methods, performance indicators, and future research directions.

2. Literature review

2.1. Computer vision foundations in cue sports

Accurate recognition of spherical objects is a fundamental challenge in the field of computer vision, particularly in dynamic sports environments. Scitovski and Marošević [8] were the first to propose a center-based clustering algorithm, achieving a detection accuracy of up to 83% in a controlled environment. However, its performance significantly deteriorates in uneven lighting or partial occlusion conditions. Chen et al. [9] improved its adaptability to different environments through adaptive Hough transform, although this led to an increase of 300 milliseconds in latency. Recent comprehensive reviews of ball detection techniques across sports highlight the evolutionary shift from traditional feature-based algorithms — such as colour segmentation and Histogram of Oriented Gradients (HOG) — to deep learning frameworks including YOLO, Faster R-CNN, and SSD [10]. While single-stage detectors like YOLO have revolutionised real-time object detection in team sports by treating detection as a regression problem [11], traditional colour space transformations remain highly effective for specific, constrained domains. Studies have further shown that purely deep learning approaches struggle with small, fast-moving objects under mixed lighting conditions, where hybrid methods combining colour-based segmentation with learned features yield superior results [12]. The present method combines these approaches by using a dynamic HSV (Hue, Saturation, Value) thresholding method — converting the RGB input to the HSV space to separate the sphere from the background of the table — while mitigating the limitations identified in the literature. In the validation test, under low-light conditions, this method achieves a detection rate 22% higher than previous methods.

The perspective correction between two-dimensional images and three-dimensional billiard table surfaces has always been a challenging issue. Zhang et al. relied on QR code markers for spatial calibration, but this introduced inconvenience for user operations due to the modification of the physical table surface [13]. Recent advancements in sports field registration have demonstrated the efficacy of self-supervised shape alignment and point-line optimisation for computing homography matrices without physical markers, achieving robust perspective correction across diverse camera viewpoints [14]. Building upon these principles, the unmarked solution presented here calculates the homography matrix by detecting the Canny edges of the table boundary, achieving sub-centimetre-level precise positioning without changing the environment. This technological advancement

DOI: 10.54254/2755-2721/2026.35989

significantly enhances the application capabilities of AI billiards assistants in various billiard venues.

2.2. Physics modeling for trajectory prediction

Alciatore's pioneering research established the mathematical relationship between the contact point (C) of the ball and its final trajectory: $\theta\_reflection = \theta\_incidence - 2\alpha \cdot (1 - e^{(-\beta \cdot v)})$ [α = cloth friction coefficient, v = velocity] [2]. Although these mathematical relationships are theoretically reliable, they are extremely difficult for non-technical users to understand. The field of trajectory prediction has more recently been transformed by Physics-Informed Neural Networks (PINNs), which integrate physical laws directly into the neural network training process, ensuring that predicted trajectories strictly adhere to principles such as conservation of momentum and energy [7]. Studies applying PINNs to 3D projectile motion in sports — including table tennis — have demonstrated that embedding physical constraints significantly improves prediction accuracy and reliability over purely data-driven models, particularly in scenarios with limited training data [15]. The researcher innovatively embedded these formulas and physical constraints into structured artificial intelligence's final hitting suggestions, converting the complex mathematical relationship concepts into instructions that are easy for beginners to understand (for example, "Hit to the left, at an angle of 30°, with medium topspin"). This basic physical setup avoids generating incorrect guidance while maintaining the accuracy of kinematics.

2.3. Generative AI for physical skill acquisition

The application of Generative AI in sports education represents a paradigm shift toward democratising professional coaching and making expert-level guidance accessible to recreational practitioners [16]. Multimodal AI systems such as GPT-4 [6] demonstrate extraordinary spatial analysis capabilities, but often generate ball-hitting suggestions that do not conform to physical logic when not restricted. To address this, recent research has focused on Knowledge-Grounded Large Language Models (LLMs) and adaptive coaching feedback systems. For instance, the AgentCoach framework has demonstrated that LLMs can provide adaptive feedback for motor skill learning by extracting key coaching points and connecting high-level semantic meaning with low-level kinematic parameters, enabling context-aware, prescriptive verbal feedback that adapts based on the learner's performance history [17]. Complementary research on AI-assisted motor learning has further established that effective skill acquisition systems must bridge the gap between theoretical principles and real-world task environments, integrating sensor data, performance error registration, and expert-novice behavioural comparison [18]. Three basic mechanisms were investigated to solve the hallucination problem in generative models: (1) hard-coding the law of conservation of momentum, (2) embedding the rebound angle calculator directly into the prompt — a form of knowledge-grounded prompt engineering — and (3) limiting the rotational parameters to the range verified through examples. This approach successfully applies generative AI with domain constraints to billiards training, bridging the gap between theoretical mechanics and novice skill acquisition.

DOI: 10.54254/2755-2721/2026.35989

3. Methodology

3.1. System architecture

The operational pipeline comprises three integrated modules. The Vision Module performs: (1) preprocessing via RGB-to-HSV conversion with Gaussian blurring ($\sigma = 1.5$); (2) table detection through Canny edge detection, quadrilateral identification, and perspective normalization; (3) ball localization via Hough circle transform (minDist = 20 px, param1 = 50, param2 = 30) with color-based false positive filtering; and (4) coordinate mapping to a normalized table reference frame as following:

```
def map_to_table(image_points):

homography_matrix = cv2.findHomography(src_pts, dst_pts)

return cv2.perspectiveTransform(image_points, homography_matrix)

{

"ball_positions": [[0.23, 0.51], [0.71, 0.32], ...],

"target_ball": 4,

"physics_constraints": "elastic_collision=True, max_spin=15",

"user_skill_level": "beginner"

}
```

The AI Coach Module receives the detected ball positions as structured input and applies constrained prompt engineering. An example prompt is as follows: "Acting as expert billiard physicist, advise novice targeting ball #4 at position [x = 0.71, y = 0.32]. Apply Alciatore's reflection law with maximum 15° spin. Output: 1) Vector-based aiming point 2) Spin type 3) Force level (low/medium/high)." The User Interface consists of a console-based workflow guiding image uploads and parameter selection via Python's Matplotlib library.

3.2. Experimental validation

The testing protocol involved 120 participant-captured images across three lighting tiers (low: <100 lux, moderate: 300 lux, bright: >500 lux) using standardized billiard tables. Evaluation metrics included: (1) Detection Accuracy, measured by Precision = TP/(TP+FP), Recall = TP/(TP+FN), and F1-score = 2×(Precision×Recall)/(Precision+Recall); (2) Advice Quality, rated on a 5-point expert scale (1 = physically dangerous, 5 = professional-grade) against certified coach benchmarks; and (3) Latency, defined as the processing duration from image upload to advice generation.

DOI: 10.54254/2755-2721/2026.35989

4. Results

4.1. Vision system performance

Detection efficacy exhibited strong correlation with ambient illumination, as shown in Table 1. Error analysis revealed two dominant failure modes: specular highlights disrupting color segmentation (58% of errors) and edge shadows causing partial occlusion (32%). The system successfully detected ≥7 balls in 85% of moderate-light scenarios, outperforming prior marker-based approaches by 19% in F1-score.

Table 1. Ball detection metrics (n=120 images)

Lighting condition	Precision	Recall	F1-score	Mean processing time
Low (<100 lux)	0.68 ±0.05	0.71 ±0.06	0.695 ±0.04	4.2 ±0.3 s
Moderate (300 lux)	0.92 ±0.03	0.94 ±0.02	0.929 ±0.02	3.1 ±0.2 s
Bright (>500 lux)	0.63 ±0.07	0.59 ±0.08	0.609 ±0.06	4.8 ±0.4 s

4.2. Coaching advice evaluation

AI-generated recommendations demonstrated 76.3% aggregate congruence with professional benchmarks. Performance varied significantly by shot complexity, as detailed in Table 2. User studies (n = 40 beginners) indicated that 84% found advice "immediately actionable," though 17% reported confusion with technical terminology (e.g., "stun shot"). Advice generation averaged 11.4 ± 2.7 seconds on consumer-grade hardware.

Table 2. Advice quality assessment by shot type

Shot difficulty	Avg. expert rating	AI alignment	Common discrepancy sources
Straight (<5°)	4.8 ±0.2	92% ±3%	Force miscalibration
30° cut shot	4.2 ±0.3	81% ±5%	Spin recommendation errors
Bank shot	3.7 ±0.4	63% ±7%	Rebound angle inaccuracy

4.3. Interface usability metrics

The console-based user interface successfully guided 92% of users through complete workflows without external assistance. Task completion analysis revealed the following timings: image upload to detection averaged 28 ± 6 seconds; shot selection to advice averaged 14 ± 3 seconds; and the error recovery rate for incorrect inputs was 79%.

5. Discussion

5.1. Technical validation

This prototype confirmed three fundamental hypotheses. First, when combined with adaptive preprocessing technology, the camera of a smartphone can provide a resolution sufficient to meet the requirements of computer vision. Second, constraints based on physical principles for the suggestions can effectively prevent generative artificial intelligence from generating incorrect outputs in the field of kinematics. Third, natural language successfully connects theoretical mechanics with the novice's understanding of how to hit the ball.

It is worth noting that the moderate lighting performance exceeded expectations, indicating immediate application value in entertainment venues. Compared with the method of Chen et al [9]., the improvement of 22% highlights the effectiveness of the dynamic HSV-based threshold method.

5.2. Limitations and corrective strategies

During the testing process, three key limitations were identified. The first is angular error caused by parallax (±4.7° deviation), whose root cause is a ±15° difference in the camera's angle relative to the billiard table plane. The proposed solution is to add an augmented reality overlay and positioning guidance function in the iterative version of the mobile application. The second limitation is color recognition failure, with an error rate of 38% for dark red and black balls, caused by channel saturation in the HSV space under warm light (2700K–3000K). A convolutional neural network classifier trained on an enhanced dataset is proposed as the corrective measure. The third limitation is the static analysis constraint — the system is unable to adjust suggestions during action execution — for which the solution path is to switch to 30 frames-per-second video processing combined with motion tracking.

These observations are consistent with the warnings from Scitovski and Marošević [8] regarding dynamic environment applications. Future iterative versions should adopt inertial measurement technology based on micro-electromechanical systems for 3D space calibration.

6. Conclusion

This study has developed a practical artificial intelligence billiards training system based on smart device technologies. The system significantly lowers the threshold for ordinary people to acquire professional skills. Its main contributions lie in promoting the development of sports teaching concepts: (1) a new computer vision processing procedure that achieves an F1 score of 92.9% under optimal conditions without the need for physical markers; (2) a generative artificial intelligence framework based on physical principles, which prevents unreasonable hitting suggestions through mechanical constraints; and (3) an optimized knowledge conversion technology for beginners, which transforms complex collision mechanisms into actionable natural language suggestions.

Although there are still limitations in disparity management and color recognition, the experimental results have confirmed the effectiveness of the core technical method across various practical scenarios. Future development will focus on: (1) utilizing optical flow tracking technology to achieve real-time video analysis; (2) combining augmented reality technology for the development of iOS/Android applications; and (3) expanding a multiplayer mode for strategic guidance. This approach indicates potential application in other fields that require spatial reasoning abilities, such as golf putting or basketball shooting. Ultimately, this work confirms that generative artificial intelligence can weaken the monopoly position of professional knowledge that was previously limited by socioeconomic barriers.

References

[1] Haar, S., van Assel, C. M., & Faisal, A. A. (2020). Motor learning in real-world pool billiards. *Scientific Reports, 10*(1), 20046.

[2] Alciatore, D. G. (2017). *The illustrated principles of pool and billiards*. Union Square & Co.

[3] Crane, J., & Temple, V. (2015). A systematic review of dropout from organized sport among children and youth. *European Physical Education Review, 21*(1), 114–131.
[4] Hylton, K., & Totten, M. (2013). Developing 'sport for all': Addressing inequality in sport. In *Sport development* (pp. 37–65). Routledge.
[5] Park, J., Chung, S. Y., & Park, J. H. (2022). Real-time exercise feedback through a convolutional neural network: A machine learning-based motion-detecting mobile exercise coaching application. *Yonsei Medical Journal, 63*(1), 92–100.
[6] OpenAI. (2023). *GPT-4 technical report*. https://cdn.openai.com/papers/gpt-4.pdf
[7] Chiha, Z., Péteri, R., & Mascarilla, L. (2024). Predicting 3D projectile motion in table tennis using computer vision and physics-informed neural network. In *2024 International Conference on 3D Vision*.
[8] Scitovski, R., & Marošević, T. (2015). Multiple circle detection based on center-based clustering. *Pattern Recognition Letters, 52*, 9–16. https://doi.org/10.1016/j.patrec.2014.10.001
[9] Chen, T., Zhang, J., & Luo, B. (2015). Multiple circle detection based on center-based clustering. In *Proceedings of the 2015 International Conference on Control, Automation and Robotics* (pp. 158–162). https://doi.org/10.1109/ICCAR.2015.7166041
[10] Moreira, C., Ferreira, L., & Coelho, P. J. (2025). A comprehensive review of ball detection techniques in sports. *PeerJ Computer Science, 11*, e3079.
[11] Yoon, Y., Hwang, H., Choi, Y., Joo, M., Oh, H., Park, I., et al. (2019). Analyzing basketball movements and pass relationships using realtime object tracking techniques based on deep learning. *IEEE Access, 7*, 56564–56576.
[12] Hiemann, A., Kautz, T., Zottmann, T., & Hlawitschka, M. (2021). Enhancement of speed and accuracy trade-off for sports ball detection in videos — finding fast moving, small objects in real time. *Sensors, 21*(9), 3214.
[13] Zhang, Z., Li, M., & Liu, Y. (2020). ARPool: Advanced augmented reality billiard coaching system using wearable markers. *IEEE Trans Vis Comput Graph, 26*(5), 1998–2007. doi: 10.1109/TVCG.2020.2973478
[14] Shi, F., Marchwica, P., Higuera, J. C. G., Javan, M., & Biswas, J. (2022). Self-supervised shape alignment for sports field registration. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 3114-3123).
[15] Gambietz, M., Dorschky, E., Akat, A., Schöckel, M., et al. (2026). SSPINNpose: A self-supervised PINN for inertial pose and dynamics estimation. *ACM Trans Graph*.
[16] Gao, Y. (2025). The role of artificial intelligence in enhancing sports education and public health in higher education: Innovations in teaching models, evaluation systems, and personalized training. *Front Public Health, 13*, 1554911.
[17] Chou, C., Kim, J., Huang, D.-Y., et al. (2026). AgentCoach: LLM-based adaptive coaching feedback for motor skill learning. In *Proceedings of the ACM CHI Conference*.
[18] Lohse, K., Sherwood, D., & Healy, A. (2022). Using artificial intelligence for assistance systems to bring motor learning principles into real world motor tasks. *Sensors, 22*(7), 2481.

DOI: 10.54254/2755-2721/2026.35989

Appendices

Appendix I. Experimental dataset specifications

Image Sources: 120 user-captured images (Snooker: 80, Pool: 40)
Resolution: 1920×1080 px (minimum), 4032×3024 px (maximum)
Lighting Classification: ISO 8995-1:2002 standard lux measurement
Ball Sets: Aramith Tournament (Pool), BCE Premier (Snooker)

Appendix II. Full prompt engineering schema

```
{
    "system_role": "billiard_physics_expert",
    "constraints": [
        "max_spin_angle=15°",
        "conservation_of_momentum=True",
        "energy_loss_factor=0.72",
        "allowable_force_levels=[low, medium, high]"
    ],
    "output_template": [
        "1. Aiming_vector: [x,y] relative to cue ball",
        "2. Recommended_spin: [top, center, bottom]_[left, right]",
        "3. Force_level: categorical",
        "4. Rationale: <50 words"
    ]
}
```