An Integrated Framework for Running Posture Analysis: Fusing UAV Vision with Deep Learning Models

Zuyuan Wang

School of Cyberspace Security (Xingu Industrial College), Chengdu University of Information Technology,
Chengdu 610225, China

Abstract: ***With the acceleration of modern life and rising health awareness, running has become one of the most popular forms of exercise due to its simplicity and lack of venue restrictions. However, incorrect running form can easily lead to sports injuries such as knee joint damage and plantar fasciitis. Traditional motion-analysis systems, while accurate, are expensive and require a laboratory environment, making them neither economical nor convenient for ordinary runners. To address these issues, this paper innovatively combines drones with deep learning to achieve running-form correction. A drone equipped with a vision sensor captures real-time video of the runner; the YOLOv8 algorithm performs human-pose recognition, and the MediaPipe Pose model reconstructs 3D poses, enabling low-cost, portable motion analysis outdoors. The system automatically detects abnormal running form and provides real-time, visual correction suggestions via a mobile app, helping runners adjust their posture promptly and prevent injuries. Compared with traditional solutions, the proposed system offers flexible deployment, ease of use, and low cost, providing the general public with professional, scientific running guidance.***

Keywords: Drone; Pose recognition; Running-form analysis; Motion-form correction.

1. INTRODUCTION

In recent years, drone technology has matured and been applied in many fields, bringing new possibilities to motion analysis. For capturing human movement, traditional cameras are limited in height and mobility when covering sports with large ranges and high speeds, whereas drone aerial photography offers unrestricted altitude and flexible speed and direction changes. Most current drone-based motion-analysis systems remain at the stage of simple video recording, lacking an integrated solution that includes real-time tracking, intelligent analysis, and instant feedback. In the niche area of running-form correction, no systematic technical solution has yet been established.

To fill this gap, this paper proposes a complete drone- and deep-learning-based running-form correction system. It innovatively integrates autonomous drone tracking, edge computing, and cloud-based intelligent analysis. First, a visual-tracking algorithm with motion compensation is designed to mitigate drone-shake during filming. Second, a gait-analysis model fusing spatiotemporal features is built to enable dynamic pose assessment in unstructured environments. Finally, a lightweight client–cloud collaborative architecture is developed. Compared with existing solutions, the proposed system offers significant advantages in mobility, real-time performance, and analytical dimensions. Economic and supply chain research is advanced by Tang, Yu, and Liu (2025) through their investigation of supply chain coordination with dynamic pricing advertising and consumer welfare impacts [1], while manufacturing automation progresses through Xie and Chen's (2025) Maestro system for multi-agent task recognition and optimization [2]. Digital advertising technologies show substantial innovation with Zhu's (2025) RAID system for reliability automation in large-scale ad platforms [3], Zhang's (2025) CrossPlatformStack enabling high-availability deployment across meta services [4], Hu's (2025) GenPlayAds for procedural playable 3D ad creation [5], and Li, Lin, and Zhang's (2025) privacy-preserving framework incorporating federated learning and differential privacy [6]. Recommendation systems further evolve through Li, Wang, and Lin's (2025) graph neural network enhanced sequential recommendation method for cross-platform ad campaigns [7]. Urban computing and public infrastructure benefit from Xu's (2025) CivicMorph for generative public space modeling [8], while communication systems advance through Tu's (2025) AutoNetTest for intelligent 5G network automation [9]. Data analytics is enhanced by Xie and Liu's (2025) DataFuse for multimodal interview analytics [10], and platform stability is strengthened through Zhu's (2025) ReliBridge as a scalable LLM-based backbone for small businesses [11]. Content creation is revolutionized by Hu's (2025) few-shot neural editors for 3D animation [12], while computer vision foundations are strengthened by Chen et al.'s (2022) one-stage object referring with gaze estimation [13]. Energy systems optimization features Gao et al.'s probabilistic planning research (2018, 2019) for minigrid balancing with renewables and storage [14-15], and materials science characterization is advanced by Zhang and Needleman's (2021) research on power-law creep parameter identification [16]. Recruitment technology evolves with Li et al.'s (2025) integration of GPT and hierarchical graph neural networks for resume-job matching [17], and time-series analysis progresses through Su et al.'s (2025) WaveLST-Trans model for financial anomaly detection [18] and Zhang et al.'s (2025) MamNet for network traffic forecasting [19]. Autonomous driving technology is significantly advanced by Peng et al.'s (2025) NavigScene framework for beyond-visual-range navigation [20], while motion recognition progresses through Guo's (2025) IMU-based data completion with LSTM [21]. Software architecture innovations include Zhou's (2025) performance monitoring in microservices architecture [22], data security through Zhang's (2025) blockchain-based medical data sharing [23], analytical methodologies through Yu's (2025) Python applications in market analysis [24], and marketing optimization through Liu's (2025) empirical analysis based on 4P theory [25].

2. SYSTEM DESIGN

2.1 System Framework

This system adopts a three-tier modular architecture consisting of a tracking drone, a client-side app, and a data-analysis server, forming a closed “capture–analyze–feedback” loop for athletic posture. 1) Data-capture tier: a vision-sensor-equipped drone performs dynamic tracking via a multi-stage PID (Proportional-Integral-Derivative) control algorithm and YOLOv8 pose recognition; 2) Interaction tier: a mobile app provides remote control, parameter tuning, and data visualization; 3) Analysis tier: the server-side MediaPipe Pose model reconstructs 3D poses, then, together with a gait knowledge base and Dynamic Time Warping (DTW), generates corrective recommendations.

2.2 Drone-Side Design: Follow-Shot Collection of Running Gait Data

2.2.1 UAV-End Hardware Design

1) Core flight components: lightweight carbon-fiber frame, low-noise three-blade propellers, 2000KV brushless motors, powered by 6s model-aircraft batteries, electronic speed controllers rated for continuous current above 20A, STM32F4-based flight controller; overall design weight ≈1.5 kg, single continuous flight time ≥15 min.

2) Perception and Navigation Module: The JY901B attitude sensor is selected; this sensor employs a magnetic field sensor combined with high-dynamic Kalman filtering to provide long-term stable heading-angle data. A laser TOF (Time of Flight) ranging sensor mounted beneath the UAV enables altitude-hold flight. For vision, a high-definition camera is carried.

3) Raspberry Pi: an edge-computing device that provides network services and the computing power required for target tracking.

2.2.2 UAV Control Algorithms

The low-level attitude control employs a cascaded PID structure: the outer attitude loop outputs the desired angular rate to the inner rate loop, achieving closed-loop regulation through real-time feedback from attitude sensors. The mid-level position control is also based on a cascaded PID design; the outer position loop receives the desired position from the upper-level visual-tracking command, combines it with monocular range feedback to generate a desired velocity command for the inner velocity loop, which then realizes closed-loop velocity control by integrating the triaxial acceleration measured by the attitude sensors.

The upper-level human-following controller performs object detection and pose estimation using a YOLOv8 model deployed on an edge-computing device. During tracking, the drone’s motion can be divided into the active motion required for target tracking and the passive motion caused by environmental disturbances; the superposition of the two introduces inter-frame motion distortion. To address this, the system first compensates for camera motion in the raw video stream to eliminate background distortion, then extracts keypoint information of the target via the pose-estimation model. Based on the target’s 2D coordinates and size in the image plane, the relative distance between the drone and the target and the deviation of the target’s center are computed; this deviation is fed into the mid-level position controller to align the target center with the camera’s optical axis. Meanwhile, by analyzing the spatial azimuth of the line connecting the target’s shoulders, the yaw angle is calculated and a heading-control command is generated so that the drone can continuously adjust its orientation to maintain persistent tracking of the moving target.

2.2.3 Network Communication

The drone's communication system is built on a Raspberry Pi platform and establishes physical- and data-link-layer connections with a smartphone via Wi-Fi Direct. At the application layer, flight-control commands and status data are encapsulated using MAVLink (Micro Air Vehicle Link), while low-latency video streaming is implemented with RTP (Real-time Transport Protocol) to satisfy the dual requirements of real-time remote control and video transmission.

2.3 Client Design

The client is implemented as a mobile app whose core is a clean and intuitive interaction design that delivers an excellent user experience. The app comprises four main functional modules: drone remote control, parameter tuning, motion-pose data visualization and analysis, and user-center management.

2.3.1 Drone Remote-Control Module

This mobile control module enables short-range remote control of the drone and real-time video monitoring from the phone. The human-machine interface adopts a layered design guided by the principles of intuitiveness, safety, and functional priority. The main screen centers on the live video feed, overlaid with flight-parameter information. A top status bar displays critical flight data in real time, while dual virtual joysticks are configured on both sides: the left stick adjusts throttle and yaw, and the right stick controls pitch and roll. The bottom function area integrates one-touch buttons for return-to-home, hover, and image capture, ensuring that emergency functions can be triggered instantly.

2.3.2 Drone Parameter-Tuning Module

The parameter-configuration interface provides key parameter tuning for the flight-control system, including PID controller gains, follow distance, altitude, and angle settings. The interface combines interactive sliders with functional buttons, and each parameter group is accompanied by a detailed description label. To improve tuning efficiency, the system supports an extended view that offers parameter explanations, tuning-experience guidance, and historical parameter records.

2.3.3 Visualized Motion-Pose Data Analysis Module

The running-form analysis visualization interface adopts a structured design, presenting processed motion-analysis results through multimodal data. The interface uses a vertically scrollable layout and includes the following functional modules: a fixed top navigation bar, a dynamic skeletal animation rendering area, a core motion-parameter card, a joint-angle time-series curve and posture-comparison heat-map analysis area, and a fixed bottom module that displays motion-improvement suggestions. This design realizes full-process visualization from real-time motion reproduction to quantitative metric analysis.

2.3.4 User Center Module

This module provides user-account management functions, including registration/login authentication, personal-information maintenance, historical-data queries, and system settings. First-time use requires login verification; otherwise, the user is forcibly redirected to the authentication flow. Personal information supports editing of fields such as avatar, nickname, gender, and body data. Historical data is presented as a list showing analysis time and a summary of running-form issues; clicking an entry jumps to the visualization-analysis interface to view details. System settings include account management and notification-permission configuration.

2.4 Server-Side Design: Data Analysis

2.4.1 Human Pose Skeleton Recognition and 3D Reconstruction

The server side uses the MediaPipe Pose human-skeleton pose model to detect 33 anatomical keypoints, outputting the 2D coordinates (x, y) and relative depth z of each point in the image. For 3D reconstruction, a predefined human-proportional model is first employed; through monocular depth estimation and bone-length constraints, the 2D pixel coordinates are back-projected into 3D space, with the hip center set as the coordinate-system origin and the Z axis pointing toward the camera. Finally, the 3D point cloud and skeletal connections in the world coordinate system are rendered as 3D graphics.

2.4.2 Gait Analysis and Key Parameter Extraction

First, time-series alignment and Kalman filtering smooth and denoise the raw 3D joint coordinates, eliminating motion-capture jitter. Next, extreme points in the vertical displacement of the ankle joint are used to detect gait key events, dividing the motion into stance, propulsion, and swing phases; continuous motion is segmented into individual gait cycles and normalized in time and space. During parameter extraction, key metrics are calculated across three dimensions: spatiotemporal parameters, joint-dynamics parameters, and symmetry indices. Spatiotemporal parameters include cadence, stride length, and ground-contact time; joint-dynamics parameters include knee-flexion angle, pelvic-tilt angle, and ankle-flexion angle.

2.4.3 Gait Feature Knowledge Base

Build a knowledge base grounded in gait-phase time-slice analysis; use the DTW algorithm to align gait cycles across different users, and, guided by key biomechanical events, divide each cycle into three primary phases—stance, propulsion, and swing—then further subdivide them into 15–20 time slices. Within each slice, extract three categories of features: spatiotemporal gait parameters, joint kinematics, and symmetry indices. Compute similarity between the streaming data and the multi-dimensional feature templates of the corresponding slice in the knowledge base, ultimately matching abnormal running patterns. Feed the analysis results, spatiotemporal gait parameters, and joint motion parameters into a large language model to generate personalized corrective exercise recommendations.

2.5 Database Design

The system is built on a MySQL relational database whose architecture comprises four functional modules: 1) User Management handles account authentication and permission control; 2) Drone Management stores device parameters and flight logs; 3) Data Acquisition records video metadata and time-series motion-pose data; 4) Result Analysis manages gait-feature metrics and abnormal-pattern data.

2.5.1 User Management Module

The User Management module’s database design includes a user-info table and a user-profile table, adopting a “basic info + extended info” split to keep core account data concise while flexibly storing personalized user information.

2.5.2 Drone Management Module

The Drone Management module’s database design contains a drone-device table and a drone-config table, managing basic drone information and personalized configuration parameters respectively, enabling full-lifecycle device management and flexible control.

2.5.3 Data Acquisition Module

The Data Acquisition module’s database design focuses on the various data generated during UAV recording, creating a recording-session table, a video-record table, and a human-pose table to achieve structured storage of collected data. The recording-session table records macro-level information for a complete UAV recording task; the video-record table stores metadata for video files generated within a single session; the human-pose table stores joint coordinates output by the pose-detection model and preprocessing pipeline.

2.5.4 Result Analysis Module

The Result Analysis module’s database design centers on gait-analysis tasks, metric data, and abnormal findings, using three core tables to structure the pose-analysis results. The gait-analysis-task table records meta-information for a single gait-analysis task, serving as the primary container for results; the gait-metric table stores the various quantitative metrics computed during gait analysis; the gait-issue table logs abnormalities or recommendations identified during analysis, providing users with correction guidance.

2.6 Network Communication Design

The app and server adopt a layered communication architecture: basic control commands and user authentication are handled via HTTPS; large files such as videos use chunked upload with resume capability and background task management to ensure stable transmission over mobile networks; real-time data is pushed through persistent WebSocket connections for state synchronization.

3. SUMMARY

This design proposes a technical solution for a motion-pose correction system based on UAVs and deep learning. Through a three-tier architecture of autonomous UAV tracking and filming, real-time reception via a mobile client, and intelligent analysis on the server side, it realizes a closed loop of outdoor motion-pose acquisition, analysis, and feedback. The system innovatively combines a multi-stage PID control algorithm with visual motion compensation to ensure filming stability, employs the MediaPipe Pose model for 3D pose reconstruction, and designs a layered communication protocol to guarantee data-transmission efficiency, providing a feasible solution to the mobility, real-time, and applicability limitations of traditional motion-analysis systems.

FUNDING:

This work is supported by the Chengdu University of Information Technology Undergraduate Education Teaching Research and Reform Project & Undergraduate Teaching Engineering Project No. JYJG2022135, No. JYJG2024169, and the Undergraduate Innovation Training Program “UAV and Deep-Learning-Based Motion-Pose Correction System (No. X202510621266).”

REFERENCES

[1] Bingrong Yu. Design and Implementation of an Android Fitness Posture-Correction APP Based on Pose

[2] Tang, H., Yu, Z., & Liu, H. (2025). Supply Chain Coordination with Dynamic Pricing Advertising and Consumer Welfare An Economic Application. Journal of Industrial Engineering and Applied Science, 3(5), 1–6.

[3] Xie, Minhui, and Shujian Chen. "Maestro: Multi-Agent Enhanced System for Task Recognition and Optimization in Manufacturing Lines." Authorea Preprints (2025).

[4] Zhu, Bingxin. "RAID: Reliability Automation through Intelligent Detection in Large-Scale Ad Systems." (2025).

[5] Zhang, Yuhan. "CrossPlatformStack: Enabling High Availability and Safe Deployment for Products Across Meta Services." (2025).

[6] Hu, Xiao. "GenPlayAds: Procedural Playable 3D Ad Creation via Generative Model." (2025).

[7] Li, X., Lin, Y., & Zhang, Y. (2025). A Privacy-Preserving Framework for Advertising Personalization Incorporating Federated Learning and Differential Privacy. arXiv preprint arXiv:2507.12098.

[8] Li, X., Wang, X., & Lin, Y. (2025). Graph Neural Network Enhanced Sequential Recommendation Method for Cross-Platform Ad Campaign. arXiv preprint arXiv:2507.08959.

[9] Xu, Haoran. "CivicMorph: Generative Modeling for Public Space Form Development." (2025).

[10] Tu, Tongwei. "AutoNetTest: A Platform-Aware Framework for Intelligent 5G Network Test Automation and Issue Diagnosis." (2025).

[11] Xie, Minhui, and Boyan Liu. "DataFuse: Optimizing Interview Analytics Through Multimodal Data Integration and Real-Time Visualization." (2025).

[12] Zhu, Bingxin. "ReliBridge: Scalable LLM-Based Backbone for Small Business Platform Stability." (2025).

[13] Hu, Xiao. "Learning to Animate: Few-Shot Neural Editors for 3D SMEs." (2025).

[14] Chen, J., Zhang, X., Wu, Y., Ghosh, S., Natarajan, P., Chang, S. F., & Allebach, J. (2022). One-stage object referring with gaze estimation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 5021-5030).

[15] Gao, W.; Tayal, D., Gorinevsky, D.: Probabilistic planning of minigrid with renewables and storage in Western Australia. In: 2019 IEEE Power & Energy Society General Meeting (PESGM) (2019). https://doi.org/10.1109/pesgm40551.2019.8973483

[16] W. Gao and D. Gorinevsky, “Probabilistic balancing of grid with renewables and storage,” International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), 2018.

[17] Zhang, Yupeng, and Alan Needleman. "On the identification of power-law creep parameters from conical indentation." Proceedings of the Royal Society A 477.2252 (2021): 20210233.

[18] Li, Huaxu, et al. "Enhancing Intelligent Recruitment With Generative Pretrained Transformer and Hierarchical Graph Neural Networks: Optimizing Resume-Job Matching With Deep Learning and Graph-Based Modeling." Journal of Organizational and End User Computing (JOEUC) 37.1 (2025): 1-24.

[19] Su, Tian, et al. "Anomaly Detection and Risk Early Warning System for Financial Time Series Based on the WaveLST-Trans Model." (2025).

[20] Zhang, Yujun, et al. "MamNet: A Novel Hybrid Model for Time-Series Forecasting and Frequency Pattern Analysis in Network Traffic." arXiv preprint arXiv:2507.00304 (2025).

[21] Peng, Qucheng, Chen Bai, Guoxiang Zhang, Bo Xu, Xiaotong Liu, Xiaoyin Zheng, Chen Chen, and Cheng Lu. "NavigScene: Bridging Local Perception and Global Navigation for Beyond-Visual-Range Autonomous Driving." arXiv preprint arXiv:2507.05227 (2025).

[22] Guo, Y. (2025, May). IMUs Based Real-Time Data Completion for Motion Recognition With LSTM. In Forum on Research and Innovation Management (Vol. 3, No. 6).

[23] Zhou, Z. (2025). Research on Software Performance Monitoring and Optimization Strategies in Microservices Architecture. Artificial Intelligence Technology Research, 2(9).

[24] Zhang, T. (2025). Research and Application of Blockchain-Based Medical Data Security Sharing Technology. Artificial Intelligence Technology Research, 2(9).

[25] Yu, Z. (2025). Advanced Applications of Python in Market Trend Analysis Research. MODERN ECONOMICS, 6(1), 115.

[26] Liu, Huanyu. "Research on Digital Marketing Strategy Optimization Based on 4P Theory and Its Empirical Analysis."