International Journal of Advanced AI Applications

Volume2.Issue4.Apr.2026

Online ISSN 3104-9338

Print ISSN 3104-932X

Hong Kong Dawn Clarity Press Limited

http://www.dawnclarity.press/index.php/ijaaa

TABLE OF CONTENTS

Online ISSN: 3104-9338
Print ISSN: 3104-932X

A Multi-Strategy Method for Spot Center Localization and Error Measurement in Engineering Applications

Xiaofeng Peng*, Min Liao

Geely University of China, China

Received: February 9, 2026

Revised: February 25, 2026

Accepted: February 26, 2026

Published online: March 18, 2026

To appear in: *International Journal of Advanced AI Applications,* Vol. 2, No. 4 (April 2026)

* Corresponding Author: Xiaofeng Peng (pengxiaofengp452@gmail.com)

Abstract. To address the challenges of existing automatic spot-center positioning methods in engineering applications, namely the difficulty in simultaneously achieving real-time performance, positioning accuracy, and detection stability, this study proposes a multi-strategy approach for spot-center positioning and error measurement. This method integrates the center of mass approach, elliptical fitting, and two-dimensional Gaussian fitting. It selects strategies based on real-time and accuracy constraints for different detection scenarios, while introducing a unified error measurement model for the quantitative evaluation of positioning results. By establishing a pixel-to-physical coordinate mapping, a systematic evaluation of spot center localization and aiming error was achieved using metrics such as the root mean square error. Experiments demonstrated that under both static and dynamic conditions, this method exhibits excellent robustness and stability, meeting the comprehensive requirements for accuracy and real-time performance in engineering applications.

Keywords: *Multi-strategy Fusion; Spot Center Positioning; Error Measurement; Performance Trade-off; Deviation Calculation*

1. Introduction

Laser aiming technology is widely applied in industrial measurements and equipment calibrations, where its precision and stability directly impact system reliability and safety [1-2]. While existing methods have extensively studied theoretical accuracy, traditional single-positioning approaches struggle to balance precision, efficiency, and stability under complex lighting conditions, background interference, and variable spot shapes, which limits their engineering applications. Therefore, constructing a multi-strategy fusion positioning

framework for complex engineering environments—enabling the complementary advantages of different positioning methods, dynamically adjusting weights to balance accuracy and computational efficiency, and integrating positioning results with target surface calibration for quantitative error measurement [3] has become a critical challenge for enhancing the practicality and robustness of visual inspection systems.

Existing automated spot center localization methods: Scholars worldwide have proposed various computer vision-based detection methods. Current laser spot center localization methods primarily fall into three categories: first, methods based on gray-level centroid calculation, which offer high computational efficiency but weak interference resistance [4]; second, methods based on contour geometric modeling, which exhibit good stability but rely on edge extraction quality [5]; and third, methods based on two-dimensional Gaussian distribution modeling, which offer high accuracy but involve complex computations [6-7]. Existing research primarily focuses on optimizing individual methods, lacking a systematic analysis of the trade-offs between accuracy, efficiency, and stability across different approaches, thereby limiting their application in complex engineering environments.

To address these issues, this study proposes a multi-strategy approach for engineering-oriented spot center localization and error measurement. The key contributions of this study are as:

(1) Establishing a multi-strategy localization framework that allows flexible selection based on engineering requirements, achieving an effective trade-off between detection accuracy and computational efficiency.

(2) Integrating a target surface calibration model to establish a unified laser aiming error measurement method, enabling the quantitative calculation of deviation distance and direction [8-9].

(3) Validating the robustness and stability of the proposed method under complex illumination conditions through comparative experiments [10].

2. Multi-Strategy Spot Center Localization and Error Measurement Method

2.1. System Design

To address the diverse requirements for detection accuracy and real-time performance in engineering applications, this study proposes a multi-strategy framework for spot-center positioning and error measurement. The overall system architecture is illustrated in Figure 1.

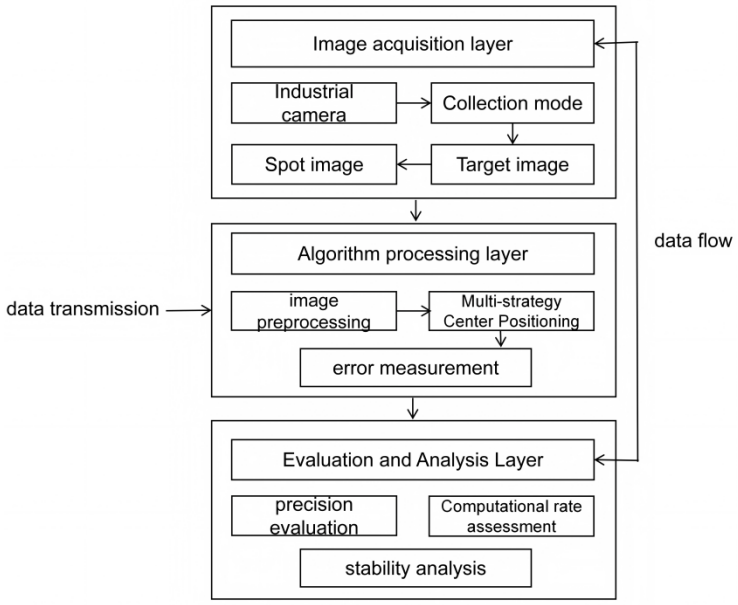

Figure 1. Framework of the multi-strategy spot center localization and error measurement system

The algorithm processing layer of this framework undergoes a focused design comprising three key steps:

(1) Spot image preprocessing: Denoising, enhancement, and segmentation are performed on the raw target surface image to obtain a stable spot region [11-12].

① Noise reduction: Employ a 3×3 median filter to eliminate isolated noise and salt-and-pepper noise while preserving edge structures; in dynamic scenes, overlay a 5×5 Gaussian filter ($\sigma = 1.0 - 1.5$) to suppress random noise, ensuring continuous grey-level distribution and facilitating subsequent Gaussian fitting.

② Image enhancement: Performed grey-level normalisation, mapping pixels to [0,255]; employed Contrast-Enhancing Localised Adaptive Histogram Equalisation (CLAHE) to boost local contrast, followed by γ correction ($\gamma = 0.8 - 1.2$)to accentuate the difference between the light spot core and background, thereby improving segmentation stability.

③ Threshold Segmentation: Under static, uniform illumination conditions, employ Otsu's global threshold segmentation. For dynamic or fluctuating illumination scenarios, utilise local mean adaptive threshold segmentation ($\text{window size } 21 \times 21, \mathrm{C} = 5 - 10$) . Following binarisation, perform 3×3 structural element opening operations and connected component analysis, retaining the largest connected region as the valid light spot area.

(2) Multi-strategy spot center localization: Select an appropriate center localization strategy based on engineering application requirements to extract the spot center [13-14]. Upon obtaining a stable spot region, the system enters the multi-strategy centre localisation phase.

Unlike traditional methods employing a single fixed algorithm, this approach supports three centre localisation strategies: grey-level centroid method, elliptical least-squares fitting, and two-dimensional Gaussian fitting.

(3) Error visual measurement: The target surface calibration model is combined to convert pixel coordinates to physical coordinates to calculate aiming error offsets [15].

Through this framework, different center positioning strategies can be flexibly switched under a unified data flow and error model, thereby adapting to the comprehensive demands of accuracy, efficiency, and stability in various engineering inspection scenarios.

2.2. Related Methods

2.2.1. Multi-Strategy Spot Center Localization

To address the diverse spot center localization demands across inspection scenarios, this study proposes a multi-strategy approach. This method integrates three common localization algorithms—centroid-based, elliptical fitting, and two-dimensional (2D) Gaussian fitting—within a unified framework. By selecting strategies based on real-time and accuracy constraints for different scenarios, optimal localization precision and robustness are ensured while maintaining real-time performance.

(1) Gray-Scale Centroid Method

The gray-level centroid method treats the laser spot as a two-dimensional distribution. The "center of mass" of the spot is calculated by performing a weighted average of the gray-level values of pixels within the spot region, using this as the spot center position.

Since the spot brightness follows a Gaussian-like distribution, its center can be obtained through brightness-weighted summation: Let the spot region be Ω, the pixel coordinates in the image be (x, y), and the corresponding grayscale value be $I(x, y)$. The spot center coordinates can be expressed as

$$x_c = \frac{\sum_{(x,y)\in\Omega} x \cdot I(x,y)}{\sum_{(x,y)\in\Omega} I\,(x,y)}, \quad y_c = \frac{\sum_{(x,y)\in\Omega} y \cdot I(x,y)}{\sum_{(x,y)\in\Omega} I\,(x,y)}$$

In the image coordinate system, (x, y) represents the horizontal and vertical coordinates of a pixel. The value $I(x, y)$ indicates the gray level of the corresponding pixel, showcasing its light intensity. Additionally, Ω defines the spot region obtained through threshold segmentation or preprocessing.

This method features a simple algorithmic structure and high computational efficiency,

enabling rapid spot-center localization with a low computational overhead. It is suitable for engineering scenarios that demand high real-time performance, where the spot shape is regular and background interference is minimal. However, it is sensitive to grayscale distribution and may exhibit positioning shifts under conditions of noise interference, spot asymmetry, or local saturation, which limits its robustness and accuracy.

(2) Elliptical Least Squares Fitting

The elliptical least-squares fitting method first extracts the edges of the spot, approximates the spot contour as an elliptical model, fits the elliptical parameters using the least-squares method, and uses the center of the ellipse as the spot center.

Fitting a quadratic curve:

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$$

Solve for parameters A, B, C, D, E, F via least squares. Under elliptical constraints, the center coordinates are given by

$$x_c = \frac{2CD - BE}{B^2 - 4AC}, \quad y_c = \frac{2AE - BD}{B^2 - 4AC}$$

In this context, x and y denote the pixel coordinates of the spot's edge points. The parameters $A \sim F$ represent the ellipse model parameters derived from fitting the set of edge points.

This method determines the center position by geometrically modeling the spot edges, thereby reducing the dependence on a uniform light intensity distribution. It maintains good positioning stability even under spot deformation or center grayscale saturation, making it suitable for detection scenarios with slightly irregular spot contours or elliptical deformations. However, it relies heavily on edge extraction quality and has high computational complexity, potentially leading to reduced positioning accuracy in complex backgrounds or under blurred edge conditions.

(3) Two-Dimensional Gaussian Fitting

The 2D Gaussian fitting method assumes that the intensity distribution of the laser spot approximates a 2D Gaussian function. This distribution was fitted using nonlinear least squares, with the center of the fitted function representing the center of the spot.

The 2D Gaussian model can be expressed as

$$I(x, y) = I_0 \exp\left(-\frac{(x - x_0)^2}{2\sigma_x^2} - \frac{(y - y_0)^2}{2\sigma_y^2}\right) + I_b$$

In this context, $I(x, y)$ represents the gray value at pixel coordinates (x, y). The variable I_0 denotes the peak intensity of the spot, while (x_t, y_t) indicates the center of the Gaussian distribution, marking the coordinates of the spot's center. The parameters σ_x and σ_y define the expansion scales of the spot in the x and y directions, respectively. Lastly, I_b refers to the background gray value bias term.

This method, based on a two-dimensional Gaussian distribution model approximation of the spot, exhibits strong noise resistance and high positioning accuracy. It enables sub-pixel-level center extraction and is suitable for high-precision detection and calibration scenarios. However, its computational complexity and sensitivity to initial parameters limit its engineering applicability when the spot deviates from the Gaussian assumption or when a high real-time performance is required.

2.2.2. Error Measurement

To enable the comparability and quantitative analysis of laser aiming errors across different spot-center positioning strategies, this study establishes a unified error measurement method based on multi-strategy positioning results. This method uses target calibration to establish a mapping relationship between the pixel and physical coordinates, providing a unified description of the deviation of the spot center relative to the target center.

The method first obtains the pixel-to-physical coordinate scaling factor k through calibration. It then converts the pixel coordinates of the spot center (x_c, y_c) to a physical coordinate system with the target center as the origin (X, Y). The conversion relationship is as follows:

$$X = k(x_c - x_t), \quad Y = k(y_c - y_t)$$

where (x_t, y_t) denotes the target center pixel coordinate system.

Based on this, the aiming errors are uniformly described in terms of their magnitude and direction as follows:

① Deviation Distance: The magnitude of the offset between the spot center and the target center is introduced as the deviation distance d, defined as the "Euclidean distance" between the spot center and the target center in the physical coordinate system:

$$d = \sqrt{(X)^2 + (Y)^2}$$

where: X, Y are the coordinate components of the spot center in the physical coordinate system.

② Deviation Angle: The directional characteristic of the spot center's offset. The deviation

angle θ is defined as the azimuth angle of the spot center relative to the target center in the physical coordinate system:

$$\theta = \arctan\left(\frac{Y}{X}\right)$$

where: θ denotes the angle between the offset direction of the spot center and the X − axis; X and Y represent the coordinate components of the spot center in the physical coordinate system.

This measurement is independent of specific localization strategies and applies to different localization results, such as the gray-level centroid, elliptical fitting, and two-dimensional Gaussian fitting methods, providing a unified evaluation basis for comparing the performance of multi-strategy approaches.

2.3. Adaptive Spot Localization Strategy Selection

Strategy selection is critical for multi-strategy localization methods. Based on the distinct characteristics of the application scenarios, the most suitable strategy selection mechanism must be identified. The system automatically presets the selection prior to operation, according to the engineering application requirements. The specific selection principles were as follows:

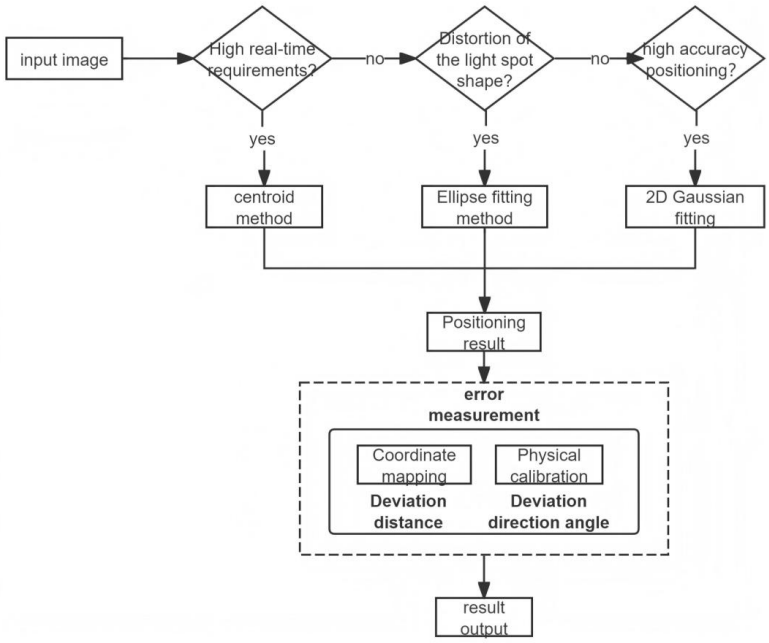

Figure 2. Multi-strategy spot center localization selection architecture

① Scenarios with high real-time requirements: For dynamic monitoring (e.g., moving laser targeting), the centroid method is recommended. This approach calculates the center of gravity of the spot using grayscale weighting, enabling millisecond-level positioning to meet real-time

detection demands.

② Scenarios with spot shape distortion: When the spot becomes distorted owing to the imaging angles or optical system effects, the elliptical fitting method is recommended. This technique enhances the positioning accuracy by fitting the spot contour with an ellipse, thereby improving the system robustness under suboptimal conditions.

③ High-precision static scenarios: For static detection or calibration requiring extreme accuracy, the 2D Gaussian fitting method is recommended. This technique establishes a 2D Gaussian model based on the brightness distribution of the light spot, achieving sub-pixel positioning accuracy suitable for static high-precision calibration. The specific architecture is illustrated in Figure 2.

3. Experimental Design and Validation

3.1. Experimental Scenario Design

Existing research predominantly focuses on single-spot center localization algorithms or compares multiple algorithms in parallel to analyze performance differences. However, studies lacking a unified system framework for achieving algorithmic cooperative scheduling and adaptive selection remain scarce [16-17]. Consequently, although such methods are effective under ideal experimental conditions, they struggle to simultaneously meet real-time, accuracy, and stability requirements in complex engineering scenarios.

To address these limitations, this study introduces the concept of "multi-strategy unified validation" during the experimental design phase. This approach systematically evaluates the performance of different positioning algorithms under static and dynamic conditions, providing an experimental foundation for the subsequent development of automated multi-strategy selection mechanisms. In static experiments, the laser maintained a fixed orientation while continuously capturing multiple frames to analyze spot extraction stability and the repeatability of different center-based positioning algorithms, thereby verifying their positioning accuracy under ideal conditions. In the dynamic experiments, spot offsets of varying magnitudes and directions were generated by adjusting the laser orientation. Corresponding deviation distances and direction angles are calculated to evaluate the system's adaptability and detection stability under multiple aiming states. The experimental system comprises a laser emission module, target surface apparatus, and image processing module. It employs a violet laser and a violet photosensitive target surface. A mark was placed at the target center, with concentric circles of radii 2/4/6/8/10 cm drawn for pixel-to-physical coordinate mapping and deviation calculation

[18].

3.2. Evaluation Metrics

This study constructs an experimental evaluation metric system based on three dimensions: detection accuracy, computational efficiency, and result stability. This metric system aims to comprehensively reflect the performance differences and engineering adaptability of different positioning strategies in practical-detection scenarios.

(1) Positioning Accuracy Metric

Positioning accuracy measures the deviation between the center of the laser spot and its true position, serving as a core metric for evaluating the laser aiming detection performance. This study uses the calibrated physical coordinates of the target surface as a reference and adopts the deviation distance of the laser spot center as the primary accuracy evaluation indicator.

Under static experimental conditions, multiple consecutive frames were captured to compute the deviation distance of the spot center relative to the target center. The root mean square error (RMSE) serves as a quantitative evaluation metric for positioning accuracy and is calculated as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} d_i^2}$$

Among these,d_i represents the deviation distance of the spot center in the i frame image, where Ndenotes the number of sampled frames. RMSE comprehensively reflects the overall level of positioning error and is suitable for precision comparison analysis between different positioning strategies.

(2) Computational Efficiency Metric

Computational efficiency measures the real-time processing capability of different spot-center localization strategies in engineering applications. This study uses the average processing time per frame as the efficiency evaluation metric, encompassing two primary computational processes: image preprocessing and spot center localization.

Under identical hardware and software conditions, the average processing time required by each positioning strategy for a single frame was statistically recorded to reflect the differences in computational complexity. This metric provides an intuitive indication of the suitability of each method for real-time requirements, serving as a crucial basis for flexible multi-strategy

selection.

(3) Stability and Repeatability Metrics

Stability evaluates the fluctuation of the spot center positioning results under repeated measurements, serving as a key indicator for assessing reliability in engineering applications. This study analyzes the statistical characteristics of the spot center deviation distance using multi-frame continuous acquisition results from static experiments.

Specifically, the standard deviation (SD) of the deviation distance was adopted as the stability evaluation metric, calculated as follows:

$$SD = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(d_i - \bar{d})^2}$$

where $\bar{d}$ is the mean deviation distance; d_i represents the deviation distance of the spot center in the i frame image; and N denotes the number of sampled frames. A smaller standard deviation indicates less fluctuation in positioning results and better system stability.

In the dynamic experiments, the consistency of the deviation calculations under varying offset magnitudes and directions was analyzed to further validate the adaptability and robustness of the proposed method in complex detection scenarios.

3.3. Experimental Results and Analysis

Table 1. Performance Comparison of Different Spot Center Localization Algorithms (Static)

Algorithm Type	Mean Square Error/ mm	Maximum Deviation/ mm	RMSE /mm	Single Frame Processing Time / ms	Standard Deviation SD /mm
Gray-scale Centroid Method	0.082	0.137	0.094	3.1	0.031
Elliptical approximation method	0.041	0.089	0.052	8.6	0.018
Two-dimensional Gaussian fit	0.028	0.061	0.036	21.3	0.021
Gray-scale Centroid Method	0.082	0.137	0.094	3.1	0.031

Under identical experimental conditions, the detection performance of the system in static scenarios was analyzed comparatively. As shown in Table 1., the three spot-center localization

algorithms exhibit distinct strengths. The gray-level centroid method achieves the shortest single-frame computation time, making it suitable for high-realtime scenarios. However, its mean error and RMSE are relatively large, and its positioning accuracy is significantly affected by spot morphology and noise. The elliptical fitting method outperforms the gray-level centroid method across all error metrics, demonstrating stability against spot deformation with moderate computation time. The 2D Gaussian fitting method achieves the highest positioning accuracy and best error control; however, its computation is time-consuming, making it unsuitable for direct application in high-real-time scenarios.

Table 2. Performance Comparison of Spot Center Localization Algorithms (Dynamic)

Algorithm Type	**Mean Absolute Distance Error/mm**	**Azimuth Angle Error/°**	**Single Frame Time Fluctuation/ms**	**Maximum Transient Deviation / mm**	**Algorithm Type**
Gray-scale Centroid Method	0.035	2.7	±0.4	0.152	Gray-scale Centroid Method
Ellipse Fitting Method	0.022	1.8	±0.7	0.098	Ellipse Fitting Method
Two-dimensional Gaussian fitting	0.019	1.5	±1.2	0.074	Two-dimensional Gaussian fitting
Gray-scale Centroid Method	0.035	2.7	±0.4	0.152	Gray-scale Centroid Method

Table 3. Relationship Between Application Scenario Requirements and Automatic Spot Positioning Strategy Selection

Application Scenario Type	**Primary Requirements**	**System-Automatically Selected Positioning Strategy**
Dynamic Real-Time Monitoring	High real-time performance, rapid response	Gray-scale Centroid Method
Spot morphology distortion	Robustness and Stability	Ellipse Fitting Method
Static high-precision detection	High Accuracy, Low Fluctuation	2D Gaussian fitting method
Dynamic Real-Time Monitoring	High real-time performance, rapid response	Gray-scale Centroid Method

Under identical experimental conditions, the detection performance of the system in dynamic scenarios was compared and analyzed. As shown in Table 2., all three spot-center localization algorithms demonstrated excellent adaptability and robustness in dynamic-targeting scenarios. The gray-level centroid method offers significant advantages in terms of computational speed and real-time performance, making it suitable for high-real-time dynamic tracking applications.

The elliptical fitting method demonstrated good robustness against spot deformation and exhibited stable direction and deviation control. The two-dimensional Gaussian fitting method maintains optimal accuracy and is suitable for high-precision slow-change processes.

The combined results of the static and dynamic experiments revealed that the performance advantages of different spot-center localization algorithms under varying operating conditions exhibited significant scenario dependency. Therefore, within a unified detection framework, the system dynamically selects localization strategies based on real-time requirements, spot characteristics, and accuracy demands, as listed in Table 3. The system automatically selects the corresponding localization strategy prior to the operation according to the predefined application scenario requirements and maintains consistency throughout the experiment.

4. Conclusions

4.1. Research Summary

The multi-strategy spot center localization and error measurement method proposed in this study for engineering applications integrates geometric center methods, centroid methods, and Gaussian distribution modeling. By dynamically adjusting the strategy weights, it overcomes the limitations of single strategies under complex conditions, such as varying illumination and background interference. In addition, combined with the target surface calibration model, the proposed error measurement method enables the quantitative calculation of spot localization errors, thereby enhancing the overall accuracy and robustness of the system.

4.2. Limitations of the Current Approach

Although the multi-strategy spot center localization method proposed in this study demonstrated good performance in experiments, it still has some shortcomings and limitations, mainly reflected in the following aspects:

(1) Insufficient Intelligent Strategy Selection: Current strategy selection and weight adjustment rely on preset rules and simple error models, lacking intelligent optimization in dynamic or unknown environments.

(2) Adaptive Error Compensation: Error compensation based on target surface calibration models may be affected by target surface variations and equipment wear in practical applications, necessitating solutions for dynamic corrections and calibration updates.

(3) Robustness in Complex Noise Environments: While proposed method performs well in conventional scenarios, its robustness in extreme noise conditions requires improvement. Enhancing the noise suppression capabilities is a key focus of future research.

4.3. Future Outlook

Future research will focus on exploring intelligent strategy selection mechanisms based on machine learning or adaptive control to further enhance the adaptability and accuracy of the system. Additionally, the introduction of dynamic calibration and adaptive error compensation mechanisms significantly improved long-term stability and precision. Moving forward, we will expand the application scope of this method, exploring its potential in fields such as Unmanned Aerial Vehicles (UAVs), autonomous driving, and robotic vision navigation. We will also strengthen its robustness in extremely noisy environments to accommodate more complex engineering scenarios.

References

[1] Zheng Hong, & Shi Qingsong. (2002). 激光瞄准大轴半径测量方法的研究与应用 [Research and Application of Laser-Aimed Large Shaft Radius Measurement Method]. *Optoelectronic engineering*, *29*(4), 28-31. [in Chinese]

[2] Wu Yue, & Wang Xinlun. (2014). 一种激光瞄准具瞄准偏差的判读系统 [A laser sight aiming deviation interpretation system]. *Mechanical Research and Applications*, *27*(4), 128-130.[in Chinese]

[3] Guan, Y. G., Li, Q., He, J. Z., Gao, X. Y., Zhou, W. C., & Wei, J. F. (2015). Effects of laser facula detector array parameters on measurement accuracy of far-field beam quality. *Optik*, *126*(23), 3906-3911. [in Chinese]

[4] Wang Bing, Zhi Qinchuan, Zhang Zhongxuan, Geng Guohua, & Zhou Mingquan. (2004). 灰度图像质心快速算法 [Fast algorithm for the centroid of grayscale images]. *Journal of Computer-Aided Design and Graphics*, *16*(10), 1360-1365. [in Chinese]

[5] Pan Deng, Li Yanli, Gao Dong, & Zheng Jianhua. (2020). 基于椭圆拟合的多光斑/重叠光斑中心提取方法 [Method for Extracting Centres of Multiple/Overlapping Spots Based on Elliptical Fitting]. *Acta Optica Sinica*, *40*(14), 1410001. [in Chinese]

[6] Shi Meihong, Mao Jianghui, Liang Ying, & Long Shizhong. (2007). 一种强高斯噪声的图像滤波方法 [An image filtering method for strong Gaussian noise]. *Computer Applications*, *27*(7), 1637-1640. [in Chinese]

[7] Yang, Y., Zhou, Y., & Huang, H. (2025, October). Introducing Unbiased Depth into 2D Gaussian Splatting for High-accuracy Surface Reconstruction. In *Computer Graphics Forum* (Vol. 44, No. 7, p. e70252). [in Chinese]

[8] Zhao Dongfeng, Dai Yaping, Yin Xianhua, Shao Ping, Huaneng, & Li Haitao. (2004). 高功率激光装置靶场光学系统的误差分析 [Error analysis of the optical system of a high-power laser device target field]. *China Laser*, *31*(12), 1425-1428. [in Chinese]

[9] Gao, R., Zhao, M., Liu, H., Qi, K., Wang, S., Li, P., ... & Luo, Z. (2024). Analysis of laser spot centroiding errors for laser acquisition system in gravitational wave detection missions. *Optics Express*, *32*(27), 48556-48570. [in Chinese]

[10] SHEN, L., ZHANG, W., LONG, J., GUO, K., LI, X., LIU, Y., ... & DAI, H. (2024). Design of teaching experiments based on high-precision laser spot centroid positioning technology. *Experimental Technology and Management*, *41*(8), 188-194. [in Chinese]

[11] Liu Danping, & Hu Yu. (2005). 提高光斑图像质心精度的去噪方法 [Denoising method to improve the centroid accuracy of spot images]. *Optoelectronic Engineering*, *32*(8), 56-

58. [in Chinese]
[12] Xu Yaming, Shu Jinfang, & An Dongdong. (2014). 自适应阈值激光光斑中心定位方法研究 [Research on the adaptive threshold laser spot centre positioning method]. *Urban survey*, (4), 5-7. [in Chinese]
[13] Jiang Jiawen, Kang Jiehu, & Wu Bin. (2021). 激光光斑中心高精度定位补偿算法研究 [Research on High-Precision Positioning Compensation Algorithm for Laser Spot Centre]. *Laser & Optoelectronics Progress*, *58*(14), 1412002. [in Chinese]
[14] Shen, L., Guo, K., Zhou, Q., Li, X., Sun, H., Deng, Z., ... & Long, J. (2026). A novel laser spot centroid positioning method based on variable-coefficient error compensation. *Optics & Laser Technology*, *194*, 114468. [in Chinese]
[15] Zhao Xuefeng, Zhong Xiaomin, & Lan Yihua. (2011). 标定靶面平行成像平面时 Tsai 算法的改进 [Improvement of Tsai algorithm when calibrating the target surface parallel to the imaging plane]. *Computer Engineering and Design*, *32*(3), 1019-1022. [in Chinese]
[16] Tang Guanqun. (2009). 几种激光光斑中心定位算法的比较 [.Comparison of several laser spot centre positioning algorithms]. *Journal of Beijing Institute of Mechanical Engineering*, *24*(1), 61-64. [in Chinese]
[17] Zhu, J., Xu, Z., Fu, D., & Hu, C. (2019). Laser spot center detection and comparison test. *Photonic Sensors*, *9*(1), 49-52. [in Chinese]
[18] Mulleria Babu, S., & SBND collaboration. (2025). The UV laser calibration system for measuring the electric field in the SBND. *Journal of Instrumentation*, *20*(06), C06055. [in Chinese]

Online ISSN: 3104-9338
Print ISSN: 3104-932X

Design of Intelligent Disinfection Doormat System Based on STM32

Biyuan Tang, Xu Wang*

Chengdu College of University of Electronic Science and Technology of China, China

Received: March 10, 2026

Revised: March 15, 2026

Accepted: March 17, 2026

Published online: March 24, 2026

To appear in: *International Journal of Advanced AI Applications,* Vol. 2, No. 4 (April 2026)

* Corresponding Author: Xu Wang (34106831@qq.com)

Abstract. At present, disinfection at the entrances and exits of public places relies heavily on manual operation. To improve the disinfection efficiency and intelligent level of areas with high-frequency personnel flow, this study adopts a modular design method and a centralized control strategy of the main control chip. Taking the STM32 chip as the system's core control unit, it integrates a stress sensing module, an alcohol atomization disinfection module, an OLED status display module, a Bluetooth communication module and a buzzer prompt module. The system structure is optimized via modular division and function integration, with reasonable control logic and execution process designed to achieve coordinated and stable operation of each module. Finally, a physical prototype is fabricated and performance tests are conducted. Test results demonstrate that the system offers fast response, stable operation and high disinfection efficiency, enabling automatic induction disinfection. It is suitable for high-traffic public places and exhibits excellent practicality and promotional value.

Keywords: *STM32F103 Microcontroller; Intelligent Disinfection Doormat; Alcohol Atomization; Stress Sensing; Performance Test*

1. Introduction

With the development of the times, the public's health awareness has been significantly enhanced, and the prevention and control of health and safety are urgent. Selecting a safe and efficient disinfection method for public place entrances has become an urgent practical engineering challenge.

Considering the requirements for intelligence, safety and convenience in disinfection, many researchers have proposed the design concept of intelligent disinfection doormats in recent years. As early as 2021, a research team has proposed a design idea that integrates gravity sensing technology with ultraviolet disinfection technology, achieving ideal disinfection effect [1]. In 2022, relevant researchers took the STM32 chip as the core control unit, constructed a basic framework of

intelligent disinfection doormat including disinfection execution module and human-computer interaction module, providing important technical support for the engineering implementation of the system [2]. In 2023, relevant scholars optimized the design of intelligent alcohol atomization disinfection equipment based on the TRIZ inventive principles, effectively addressing the safety shortcomings of traditional disinfection methods and further enhancing the operational reliability of intelligent disinfection devices [3]. From the above research, it can be seen that the design method of intelligent disinfection doormat has an important impact on the overall performance, structural rationality and development cycle of the product.

Therefore, this study adopts the design idea combining function and structural form, optimizes the matching relationship between system functions and hardware structure, and while shortening the design cycle, enables the system to meet the actual needs of efficient, safe and automatic disinfection in public places.

2. Overall Design of the Intelligent Doormat Control System

The smart doormat disinfection system is primarily designed for the entrance areas of public places with high foot traffic and density, such as hotels, shopping malls, and hospitals. These areas serve as mandatory passageways for personnel entry and exit, and their floor hygiene directly impacts the risk of pathogens spreading into the indoor environment, making them a critical point for infection prevention and control. To address this issue, this system aims to develop an intelligent doormat with automated operation capabilities and efficient sterilization, to minimize the possibility of pathogens carried on soles entering the indoor space, thereby strengthening the hygienic protection of public areas.

Table 1. Correspondence between system core functions and hardware modules.

Core Functional Module	Automatic Sensing	Disinfection Processing	Human-Computer Interaction	Wireless Communication	System Power Supply
Hardware Implementation Unit	Stress Sensing Module	Alcohol Disinfection Module	Display Module	Bluetooth Module	Power Module

To meet practical application requirements, the product is endowed with the following key functions: First, the system needs a highly sensitive step-sensing mechanism, capable of detecting personnel entry and exit in real-time and autonomously initiating the disinfection process, achieving unattended fully automated operation. Second, it must integrate an immediate and harmless disinfection mechanism to ensure effective sterilization of the soles. Third, it should be equipped with an intuitive human-computer interaction interface, displaying the working status in real-time and providing clear operational feedback for users and administrators. Additionally, the system

needs to support short-range wireless communication functionality, facilitating remote status monitoring and parameter setting by staff via mobile terminals or a management backend. Finally, the system should have a continuous and stable power supply solution to meet the demands of long-term, uninterrupted operation.

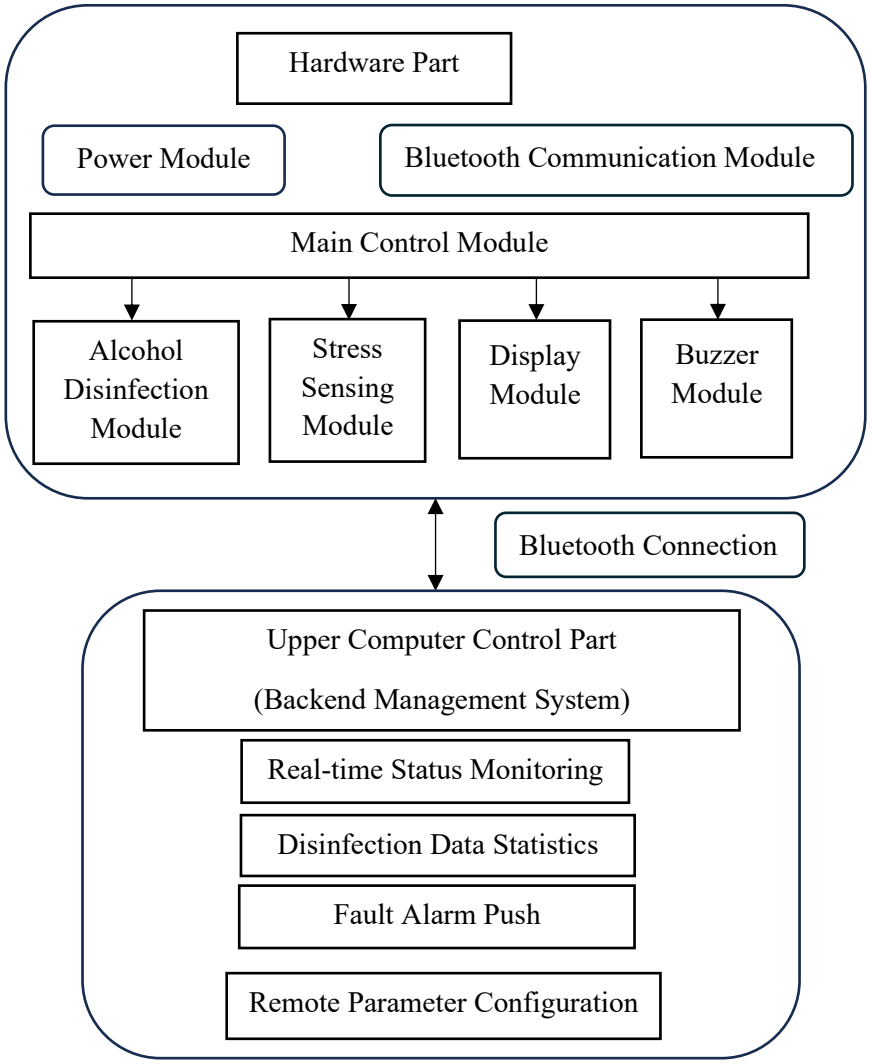

Figure 1. Overall layout of the system.

At the hardware implementation level, this system adopts a modular design concept, establishing a hardware architecture centered around an STM32 main controller. Each core function is realized by a dedicated hardware module, with the specific correspondences shown in Table 1. Specifically, the automatic sensing function is accomplished by the stress sensing module, the disinfection processing function is achieved by the alcohol disinfection module, human-computer interaction is conducted via the display module, the wireless communication function is supported by the Bluetooth module, and stable power supply for the entire system is ensured by the power module.

Through the collaborative work of each module and the organic integration via the system bus, a fully functional and efficient intelligent disinfection doormat system is collectively formed.

3. Hardware Design of the System

The hardware design of the smart disinfection doormat system is based on the overall system architecture shown in Figure 1. A comprehensive discussion is carried out closely around core functional units such as the STM32 main control chip, the alcohol atomization disinfection module, and the stress sensing module. The work sequentially covers the selection of key components for each module, performance matching, and the schematic design of the hardware circuits, while also considering interface compatibility between modules and circuit stability. This forms a complete and implementable hardware solution, laying a solid and reliable hardware foundation for the subsequent software programming development, program debugging, and the fabrication, assembly, and testing of the physical prototype.

3.1. Main Control Module Design

As the core unit of the entire system, the main control module must possess strong processing capability, rich peripheral interfaces, and excellent real-time response performance to realize the coordinated scheduling and stable operation of multiple modules such as stress sensing, disinfection control, status display, and wireless communication. Based on a comprehensive consideration of the system's processing performance requirements, development cost control, and development ecosystem support, combined with the relevant research results in References [4-5], this design selects the STM32F103C8T6 microcontroller as the main control chip, and its minimum system is shown in Figure 2.

This type of microcontroller is based on the high-performance ARM Cortex-M3 core with a main frequency of up to 72 MHz Compared with traditional 8-bit or 16-bit single-chip microcomputers (such as the 51-series single-chip microcomputers, which usually have a main frequency of 12 MHz), its data processing speed is significantly improved. It can efficiently complete real-time multi-task scheduling, ensuring the system's rapid collection and response to sensor signals, as well as the coordinated and smooth operation of each functional module.

Rich peripheral resources are the core basis for the selection of the main control chip: first, the chip integrates multiple advanced timers, which can generate high-precision PWM signals to accurately control the atomization volume of the alcohol atomization module and ensure the stability of disinfection effect; second, it has built-in various standard communication interfaces such as I2C, SPI, and USART, which can realize seamless connection with OLED display screen, Bluetooth communication module, and digital sensor, simplifying the complexity of hardware connection;

third, the chip has a built-in ADC conversion unit, providing convenient and reliable hardware support for the collection of analog signals output by the stress sensor.

Table 2. STM32F103C8T6 Main Parameters.

Kernel	Flash Capacity	CCM	Main Frequency	Peripheral Resources	Kernel
Cortex-M3	64 KB	20 KB	72 MHz	PWM,SPI,I2C,USART,etc.	Cortex-M3

In addition, the STM32F103C8T6 is equipped with 64KB Flash storage space and 20KB SRAM running memory, and its storage capacity is sufficient to carry the complex program logic and data cache requirements of the system. This type of chip has an extremely high popularity in the market. It not only has the advantages of low cost and stable supply, but also has a mature STM32Cube development ecosystem and abundant technical data and development cases, which can effectively reduce development difficulty and significantly shorten the product development cycle.

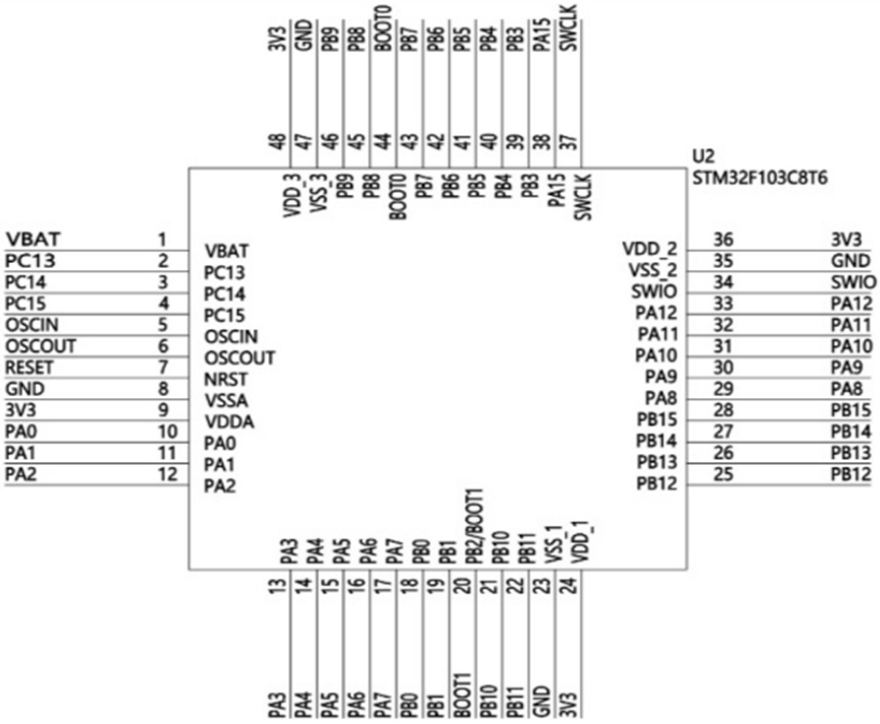

Figure 2. STM32F103C8T6 main controller.

In summary, the STM32F103C8T6 microcontroller achieves the optimal balance between performance, function, and cost, can fully meet the main control requirements of the intelligent disinfection doormat system, and is the optimal selection for this design.

3.2. Alcohol Disinfection Module Design

This module utilizes ultrasonic atomization technology. Compared to traditional liquid spraying disinfection methods, it offers the advantages of uniform atomization, no residue, and high disinfection efficiency, effectively preventing issues such as slippery floors and disinfectant waste.

.

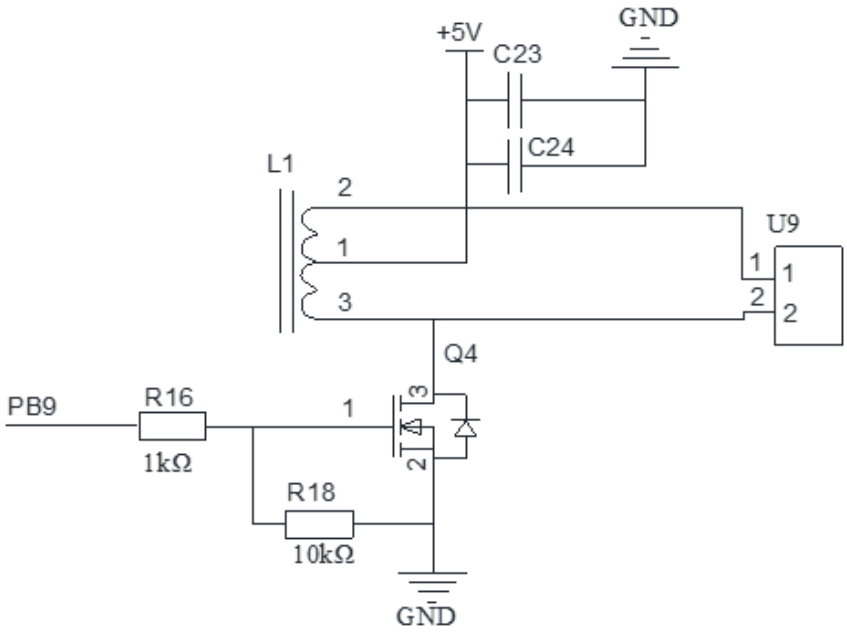

Figure 3. Alcohol disinfection module.

To achieve efficient atomization and ensure that disinfectant particles uniformly cover the disinfection area, this design employs a high-performance piezoelectric ceramic transducer as the core atomization component. The mechanical vibration frequency generated during its operation ranges from 1.7 to 2.4 MHz Within this frequency range, a 75% concentration alcohol disinfectant can be efficiently broken down into fine particles of 1–5 μm [6]. Droplets of this particle size ensure the disinfection effect while evaporating quickly to avoid residue. The system's main controller precisely adjusts the spray volume by regulating the PWM signal. The specific circuit design of the alcohol disinfection module is shown in Figure 3.

3.3. Stress Sensing Module Design

The stress sensing module serves as the signal acquisition core of this intelligent disinfection doormat system. It is primarily responsible for real-time detection of pressure signals generated when a person steps on the doormat. Once an effective gravity trigger signal is sensed, it immediately sends a trigger command to the main controller to initiate the subsequent disinfection action. The detection performance of this module directly determines the timeliness and reliability of the system's response. Therefore, during the design process, key considerations such as detection sensitivity, linearity, and anti-interference capability must be prioritized to ensure accurate identification of valid stepping signals, thereby avoiding false triggers or missed detections.

This system employs the IMS thin-film resistive sensor as the core stress detection component. This device offers excellent linear characteristics and detection accuracy, along with a wide sensing range. It is well-suited for various stress detection scenarios encountered in the practical use of the

doormat, meeting the system's requirements for precise acquisition of stepping signals [7]. The specific circuit design of the stress sensing module is shown in Figure 4.

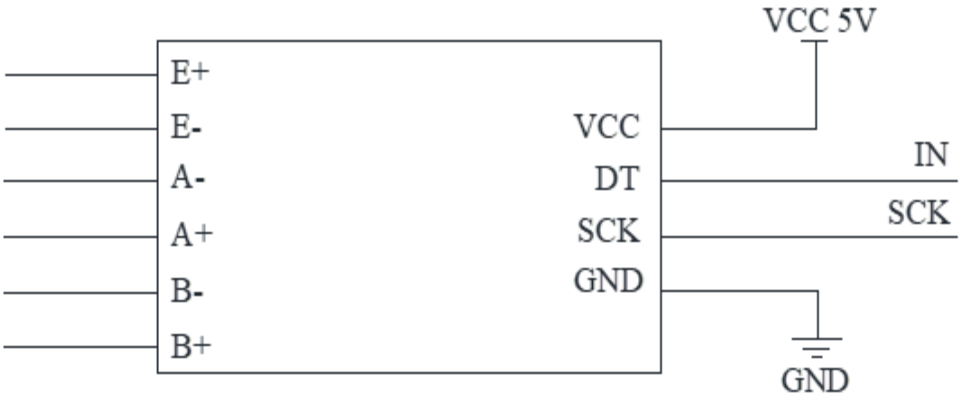

Figure 4. Stress sensing module.

3.4. Bluetooth Module Design

This wireless communication module is mainly used to realize short-range data interaction between the system and mobile devices, facilitating on-site operation and status checking of the doormat by staff. From the perspective of actual application scenario characteristics, intelligent disinfection doormats are mostly deployed at indoor entrance areas of public places such as hotels, shopping malls and hospitals. Such spaces are relatively enclosed, and the typical operation radius of staff is usually no more than 10 meters. Therefore, the system has no stringent requirements for wireless communication indicators such as transmission distance and wall-penetrating capability, but instead prioritizes communication stability, low power consumption, and hardware compatibility with the STM32F103C8T6 main control chip. Based on this application requirement, this study finally selected Bluetooth communication technology as the wireless interaction solution, and precisely adopted the E104-BT5032A Bluetooth module [8]; this module adopts the BLE5.0 low-power Bluetooth protocol, which can effectively reduce the overall power consumption of the system while ensuring the reliability of data transmission, and is highly compatible with the low-power operating modes of the STM32F103C8T6.

In terms of hardware adaptation, the E104-BT5032A Bluetooth module achieves seamless connection with the STM32F103C8T6 main control chip through the USART serial communication interface — the STM32F103C8T6 has 3 built-in USART peripheral interfaces, and USART2 is selected as the communication port in this design. This port supports full-duplex asynchronous communication and configurable hardware flow control, fully meeting the data receiving and transmitting requirements of the Bluetooth module. The effective communication distance of this module can reach 50-70 meters in an open environment, and even in actual indoor application scenarios with multiple wall obstructions, the effective communication distance can still cover the

operating range around the disinfection doormat, which is sufficient to meet the needs of staff to complete parameter setting, status query and other operations through mobile devices. The core technical parameters of this module are shown in Table 3. Its supply voltage of 3.6V is compatible with the output of the 3.3V/5V power module of the STM32F103C8T6 system, eliminating the need for additional level conversion circuits and simplifying the complexity of hardware integration.

Table 3. Basic parameters of E104-BT5032A.

Maximum Transmit Power	Bluetooth Protocol	Effective Distance	Receiving Sensitivity	Supply Voltage	Maximum Transmit Power
3.8dBm	BLE5.0	50-70m	-96dBm	3.6V	3.8dBm

3.5. Display Module Design

To provide real-time feedback on the operational status of the disinfection doormat—such as disinfection start/stop status, disinfection duration, and equipment operation status—the system is configured with a display module as the human-computer interaction interface. Based on the system's human-computer interaction requirements and installation space constraints, this study selects a 0.96-inch OLED display with a resolution of 128×64, paired with an SSD1306 driver chip, as the core hardware of the display module.

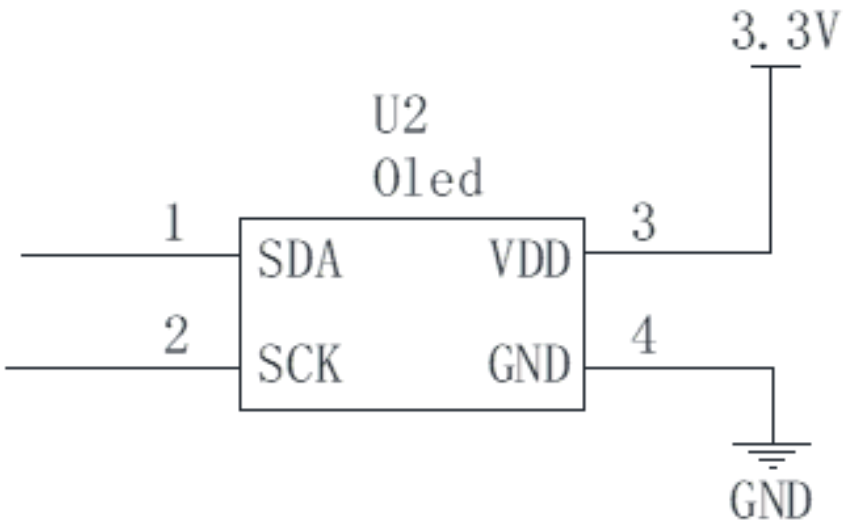

Figure 5. Display module.

The display module is directly driven by the STM32F103C8T6 main control chip through the I2C serial communication interface (one of the built-in peripheral interfaces of the STM32F103C8T6). This chip supports standard I2C communication protocols, enabling stable and low-power data transmission between the main controller and the SSD1306 driver chip, which perfectly matches the low-power characteristics of the OLED display. The display module features self-luminous properties, eliminating the need for an additional backlight unit, and can clearly present information even in the absence of backlighting. It also offers fast response times and low power consumption,

enabling timely and accurate feedback on various operational statuses of the disinfection doormat [9].

The specific circuit design of the display module is shown in Figure 5.

3.6. Buzzer Module Design

The buzzer module, as a crucial status feedback unit of the smart disinfection doormat system, is primarily used to provide clear auditory cues at the start or end of disinfection, or when system malfunctions occur. It assists personnel and entrants in quickly identifying the system's operational status, thereby further enhancing the human-computer interaction functionality.

Table 4 . Main parameters of the buzzer module.

Drive Mode	Operating Frequency	Operating Temperature Range	Operating Voltage	Drive Mode	Operating Frequency
High/Low Level Triggering	2.3KHz	-20℃~+70℃	4~8V	High/Low Level Triggering	2.3KHz

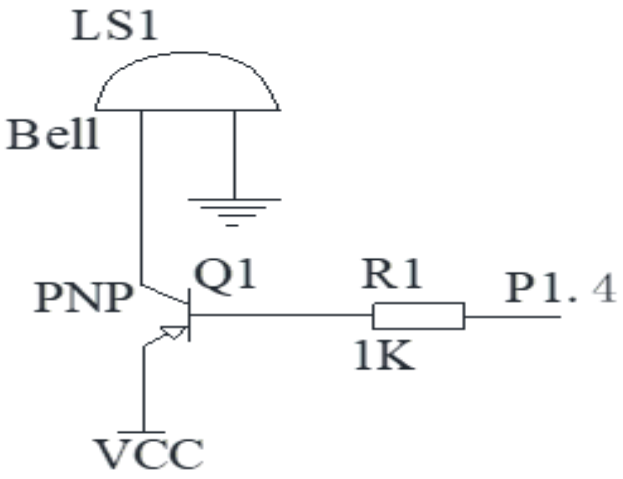

Figure 6. Buzzer module.

This design employs an active buzzer [10], which integrates an internal oscillation circuit, eliminating the need for an additional oscillation driving circuit. Its simple structure helps reduce the number of peripheral components, simplifies the hardware circuit design, and lowers the difficulty of system development and debugging. The driving circuit of the buzzer utilizes a 9012 transistor as the switching element. The main controller outputs control signals through an I/O port to manage the buzzer's operation, with a driving current of up to 30 mA, ensuring stable performance. The buzzer emits a fixed frequency of 2.3 kHz, producing a clear and distinguishable sound. With a sound pressure level, no less than 85 dB measured at 1 meter directly in front, its acoustic characteristics ensure that the alert tones are effectively recognizable against background noise in

typical environments. The main parameters of the buzzer module are listed in Table 4, and its circuit design is illustrated in Figure 6.

4. Software Design of the System

4.1. Workflow of the System

After the system is powered on, the main controller STM32F103C8T6 first performs system initialization, sequentially configuring the I/O ports, interrupt parameters, timers, and operating modes of each peripheral module to ensure that all modules can respond correctly to the commands of the main controller.

After the initialization is complete, the system enters the main loop mode, continuously monitoring signals from the force sensing module and waiting for a valid trigger. When a person steps onto the disinfection doormat, the sensor within the force sensing module detects the pressure change in real time, and the collected signal is conditioned and transmitted to the main controller. The main controller performs software filtering on the received force value to eliminate interference and then compares it with a preset threshold. If the detected force value exceeds the preset threshold, it is determined as a valid step trigger, and the disinfection process is immediately initiated.

This process specifically includes:

(1) System Initialization: After the system is powered on, the main controller STM32F103C8T6 first performs system initialization, sequentially configuring the I/O ports, interrupt parameters, timers, and operating modes of each peripheral module to ensure that all modules can respond correctly to the commands of the main controller.

(2) Standby Monitoring: Once the initialization is complete, the system enters the main loop mode, continuously monitoring signals from the force sensing module and waiting for a valid trigger.

(3) Signal Detection and Judgment: When a person steps onto the mat, the sensor within the force sensing module detects the pressure change in real time. The collected signal is conditioned and transmitted to the main controller. The main controller performs software filtering on the received force value to eliminate interference and then compares it with a preset threshold. If the detected force value exceeds the preset threshold, it is determined as a valid step trigger, and the disinfection process is immediately initiated.

(4) Disinfection Process Execution: The disinfection process specifically includes: the main controller activates the alcohol atomization module by outputting precise PWM signals to achieve spray disinfection; controls the display module to clearly show "Disinfecting" on the OLED screen; drives the buzzer to emit an audible alert, informing personnel that disinfection is in progress; and simultaneously starts a timer to precisely control the disinfection duration according to preset

parameters.

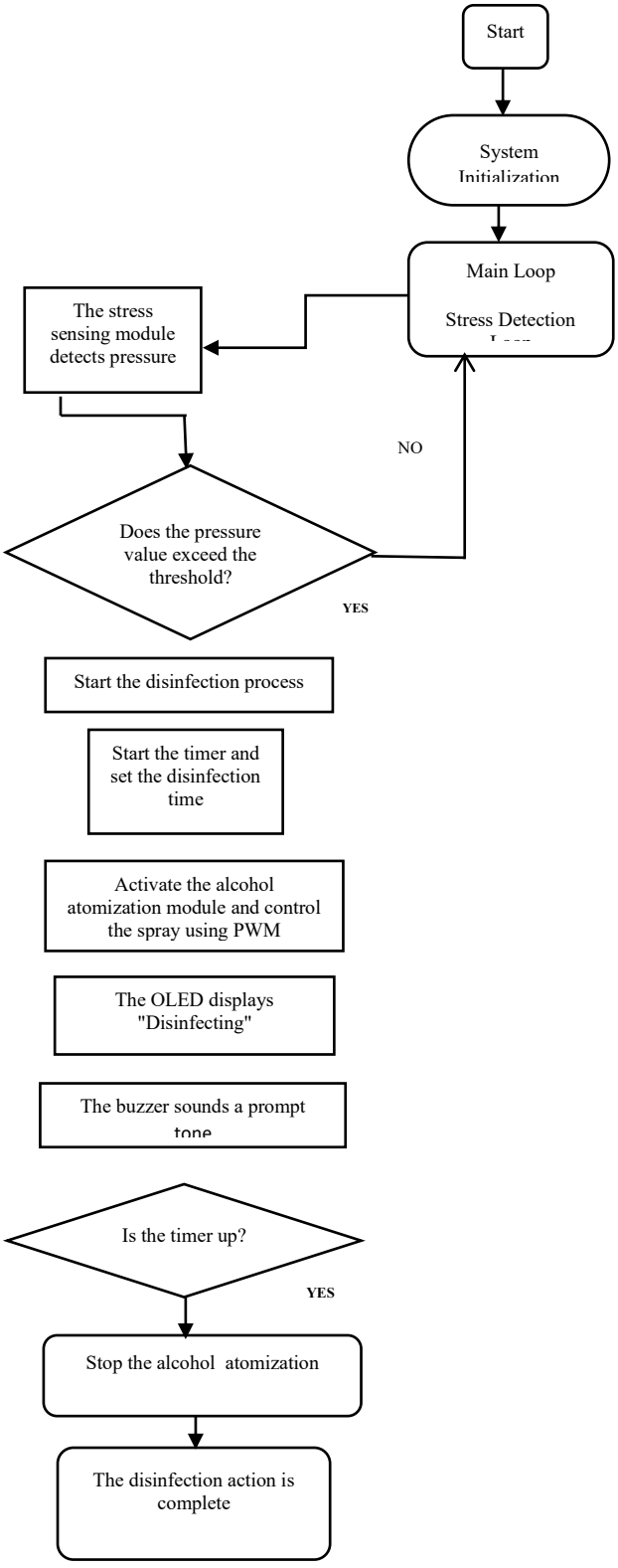

Figure 7. System workflow.

(5) Disinfection Completion and Reset: Once the disinfection duration reaches the preset value, the process automatically ends: the buzzer emits another alert to indicate completion; the OLED display updates to show "Disinfection Complete"; the system then returns to the main loop, continuing to monitor signals from the force sensing module for the next valid trigger.

(6) Remote Control Monitoring: Throughout the system's operation, the Bluetooth module remains in a stable listening state, capable of receiving remote control commands from mobile devices in real time. Personnel can use mobile devices to manually start disinfection, adjust parameters, and perform other operations, enhancing the convenience of system control.

The specific software workflow of the system is shown in Figure 7.

4.2. Upper computer Design

To further enhance the convenience and intelligence of system management, and to meet the requirements for centralized management of the disinfection doormat in the background, this system is equipped with a companion upper computer management platform. The upper computer is primarily designed to facilitate background monitoring of the disinfection doormat's operation by management personnel, enabling real-time insight into the core status of the doormat (such as whether the alcohol level is sufficient, daily disinfection user count, equipment operation faults, etc.) and its operational dynamics, thereby achieving refined management of the equipment.

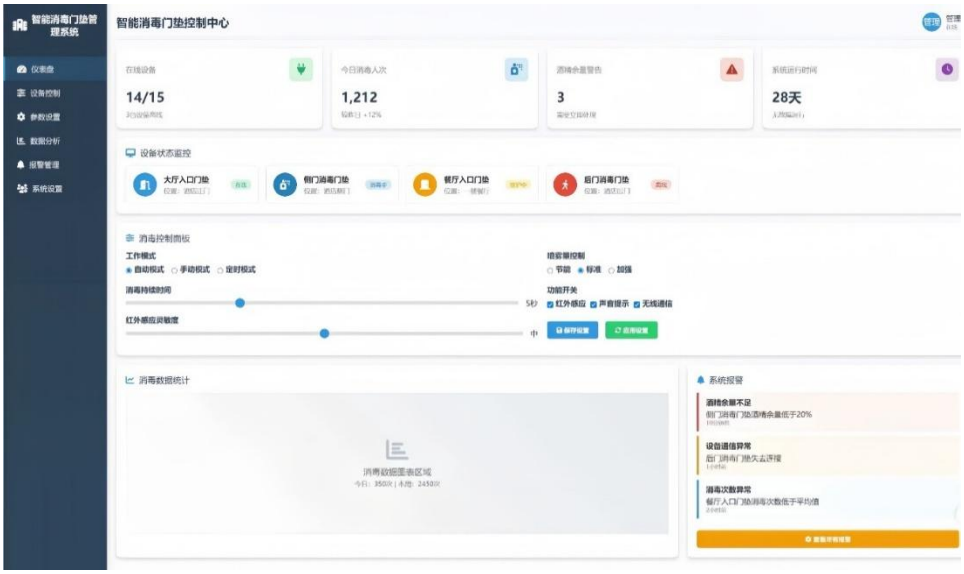

Figure 8. Disinfection doormat management system.

The interface of the disinfection doormat management system us shown in Figure 8. (Chinese Interface)

Based on actual management requirements, the upper computer design is mainly divided into six core modules, each with clear functional positioning and collaborative operation. These modules include:1. Dashboard Module, 2. Device Control Module,3. Parameter Setting Module, 4. Data Analysis Module, 5. Alarm Management Module, 6. System Setting Module. Each module performs its specific functions, collectively enabling comprehensive background control of the disinfection doormat system. [11-12]

5. System Performance Testing

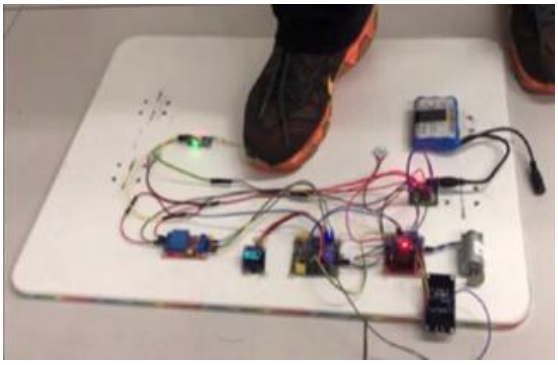

Figure 9. Test process.

After successfully completing the intricate hardware assembly and comprehensive software debugging of the innovative smart disinfection doormat, an initial prototype model was developed to verify the rationality and overall feasibility of the system structure design. This prototype aimed to assess the reliability and smooth collaboration among each individual module, as well as the effectiveness of the core functions integrated within the complete system. Following this important development phase, systematic performance testing was conducted on the prototype to ensure its stable operational capabilities. The primary tests focused on four key performance indicators: response time, operational stability, sensor measurement accuracy, and disinfection efficiency, all of which are essential for evaluating the practical performance and usability of the doormat in various real-world applications. This thorough evaluation process is crucial to ensuring that the doormat meets the necessary standards for modern disinfection solutions.

Table 5. Performance Test Results provides a comprehensive quantitative verification and evaluation of the core performance of the smart disinfection doormat system designed in this study. The table systematically presents the test methods, test results, and preset standards for four key performance indicators: response time, operational stability, sensor measurement accuracy, and disinfection efficiency, clearly reflecting the compliance status of each indicator. From the verification results presented in the table, all test indicators strictly meet the preset design requirements, with no deviation from the preset standards. This fully demonstrates that the system meets design expectations in terms of real-time response, long-term operational reliability, signal detection accuracy, and disinfection effectiveness, making it suitable for practical application scenarios such as hotels and other public places. In summary, the test results in Table 5 effectively validate the scientific nature, rationality, and engineering feasibility of the smart disinfection doormat system designed in this study, providing support for the practical application and promotion of the system, as well as offering insights for the optimization design of similar smart disinfection devices.

Table 5. Performance test results.

Test Indicators	Test Method	Test Results	Preset Standards
Response Time	Simulate normal stepping, record the total time from signal detection to deactivation initiation and display update, repeat the test 30 times, and calculate the average.	The response time (0.32 s on average, maximum 0.5 s) is defined as the total duration from pressure signal exceeding the preset threshold to the initiation of alcohol atomization and OLED status display update. (including the 0.08 s sliding mean software filtering time for stress sensor signals)	Response Time≤1s
Alcohol Consumption	Test the alcohol consumption under continuous effective trigger (one trigger per 5 minutes), record the consumption for 1 hour and calculate the hourly average.	Average consumption: 8.5mL/h, no alcohol leakage	≤10mL/h (economical and practical)
Sensor Measurement Accuracy	Use loads of 40kg, 60kg, and 80kg to simulate stepping, repeating the test 10 times for each weight.	Measurement error≤±2%.	Measurement error≤±3%, capable of distinguishing between valid stepping and false triggers.
Disinfection Efficiency	Use 75% alcohol to disinfect samples contaminated with E. coli for 5 seconds, and test the disinfection rate.	Disinfection rate≥99.2%.	Disinfection rate ≥ 99%, complies with relevant standards for disinfection in public places.
False Trigger Rate	Set 4 interference scenarios: light stepping (≤20kg), pet stepping (5-15kg), rolling luggage (10-30kg), environmental vibration; repeat 100 times for each scenario, count the false trigger times.	False trigger rate =0.02%,no missed detection of valid triggers (≥40kg)	≤0.5%
Power Supply Durability	Test with DC 5V/2A power supply (mains power) and 3.7V 10000mAh lithium battery (backup power); record continuous operation time under battery power.	Battery power: continuous operation for 36h (standby) / 12h (continuous trigger); no power supply interruption	Standby ≥24h, continuous trigger ≥ 10h
Response Time	Simulate normal stepping, record the total time from signal detection to deactivation initiation and display update, repeat the test 30 times, and calculate the average.	The response time (0.32 s on average, maximum 0.5 s) is defined as the total duration from pressure signal exceeding the preset threshold to the initiation of alcohol atomization and OLED status display update. (including the 0.08 s sliding mean software filtering time for stress sensor signals)	Response Time≤1s

6. Conclusion

During the design process of the intelligent disinfection doormat control system based on STM32, guided by the actual application scenario requirements of public places such as hotels and the preset core functions of the system, the responsibilities and positioning of main functional modules including alcohol atomization disinfection, stress sensing, OLED display prompt, Bluetooth communication and buzzer status prompt were clarified, and the overall system architecture layout and the logical connection design between various modules were completed. On this basis, the hardware circuit design, component selection and debugging work with STM32F103C8T6 microcontroller as the core were focused on, and the compilation, debugging and optimization of supporting control programs were completed synchronously, thus constructing a complete and implementable intelligent disinfection doormat control system. The performance test results show that all functions of the system operate stably, trigger accurately, respond quickly and have good disinfection effect, basically achieving the expected design goals and being fully capable of meeting the daily disinfection application needs of public places.

In future research, I will further optimize the functional and structural design of the doormat, improve the detailed issues such as false trigger protection and energy consumption control, continuously enhance the comprehensive performance, practicality and economy of the product, and promote the extensive popularization and application of the intelligent disinfection doormat in more public places such as hotels, shopping malls and hospitals.

References

[1] Zhang, S. Q., & Wang, X. (2021). 基于专家评价法的智能消毒门垫设计 [Design of Intelligent Disinfection Door Mat Based on Expert Evaluation Method]. *Electromechanical Technology*, 28–30. [in Chinese]

[2] Pan, J. X. (2022). 基于 STM32 的智能消毒门垫控制系统设计 [Design of an Intelligent Disinfection Door Mat Control System Based on STM32]. *Electronics making,* 30(23), 32–34, 79. [in Chinese]

[3] Guo, Y. Y., & Wang, X. (2023). 基于 TRIZ 发明原理的智能消毒门垫设计 [Intelligent Disinfection Door Mat Design Based on TRIZ Inventive Principles]. *Electromechanical Technology,* (2), 20–22. [in Chinese]

[4] Zhang, P. (2020). 工业产品形态设计(第 3 版) [Industrial Product Form Design (3rd Edition)]. *Beijing Institute of Technology Press.* [in Chinese]

[5] Zhu, R. L., Xiao, Y. L., Chen, R. X., & Liu, Y. (2023). 基于 STM32F407 的 IBP 仿真系统设计与实现 [Design and Implementation of IBP Simulation System Based on STM32F407]. *Communication World,* 32(6), 175–177. [in Chinese]

[6] Zhang, H., Pan, Y. C., Zhang, Z. F., & Li, T. (2024). 基于 STM32 单片机的智能加湿器设计 [Design of an Intelligent Humidifier Based on STM32 Microcontroller]. *Technological innovation,* (19), 103–106. [in Chinese]

[7] Zhang, Y. J., & Wang, X. (2022). 传感器基础与应用 [Fundamentals and Applications of Sensors]. *Sichuan Science and Technology Press.* [in Chinese]

[8] Sun, L. M., & Mou, K. C. (2025). 基于蓝牙传输的音叉物位开关的设计 Design of a Tuning Fork Level Switch Based on Bluetooth Transmission]. *Automation and Instrumentation,* 40(5), 105–107. [in Chinese]

[9] Ma, H. W. (2024). 基于 FPGA 的 OLED 显示屏控制电路设计 [Design of FPGA-based OLED Display Control Circuit]. *Beijing Jiaotong University*. [in Chinese]

[10] Qu, M. (2025). 基于 PLC 的变频蜂鸣器音频调速自动控制系统 [PLC-based Variable Frequency Buzzer Audio Speed Automatic Control System]. *Automation and Instrumentation,* 40(1), 210–213. [in Chinese]

[11] Zhang, H., Tang, J., Tang, C. W., & Yang, J. (2025). 基于 STM32 的高性能温度控制器算法研究与实现 [Research and Implementation of High-Performance Temperature Controller Algorithm Based on STM32]. *Automation Technology and Applications,* 44(4), 2–5. [in Chinese]

[12] Deng, H. (2022). 虚拟仿真平台在嵌入式系统课程中的应用 [Application of Virtual Simulation Platform in Embedded Systems Courses]. *Computer Programming Skills and Maintenance,* (10), 145–147. [in Chinese]

Online ISSN: 3104-9338
Print ISSN: 3104-932X

Clinical Study of Battlefield Acupuncture for Acute Pain and Edema Management in Distal Radius Fractures within an Emergency Setting

Feng Li, Yufang Deng*

Qianxinan Prefecture, Xingyi, 562400, China

Received: March 20, 2026

Revised: March 21, 2026

Accepted: March 22, 2026

Published online: March 26, 2026

To appear in: *International Journal of Advanced AI Applications,* Vol. 2, No. 4 (April 2026)

* Corresponding Author: Yufang Deng (78805751@qq.com)

Abstract. Distal radius fractures (DRF) are highly prevalent traumatic injuries often complicated by severe acute pain and prolonged edema, which hinder early functional recovery and increase healthcare burdens. This prospective, randomized controlled trial evaluated the efficacy of Battlefield Acupuncture (BFA) as an adjunct therapy for 64 patients with DRF during the acute edema phase. Participants were randomly assigned to either a BFA group (n=32), receiving auricular press needles at five specific points combined with standard manual reduction and splinting, or a Control group (n=32) receiving standard care alone. Results demonstrated that the BFA group achieved a superior analgesic effect, with Visual Analogue Scale (VAS) scores decreasing by over 50% within 72 hours post-intervention. Additionally, the BFA group showed significantly faster edema resolution, with the swelling index decreasing by 35.0 ± 4.2% within 72 hours ($P=0.003$), accompanied by a marked increase in serum β-endorphin levels (+42.3 ± 5.8 pg/mL, $P<0.01$) and reduced Substance P. No severe adverse events were observed, and the BFA group reported higher patient satisfaction and optimized medical resource utilization. In conclusion, BFA is a safe, effective, and economically feasible intervention that enhances acute symptom control and the early recovery trajectory of patients with distal radius fractures.

Keywords: *Battlefield Acupuncture; Distal Radius Fracture; Acute Pain Management; Edema Resolution*

1. Introduction

Distal Radius Fractures (DRF) are among the most common traumatic injuries treated in

emergency departments, accounting for approximately one-sixth of all fracture cases. Distal Radius Fractures (DRF) linked to osteoporosis have become a significant public health issue due to their increasing prevalence [3]. Currently, clinical management of DRF primarily relies on manual reduction combined with plaster or splint immobilization, or open reduction and internal fixation. However, these conventional approaches often face challenges in the acute phase, including poorly controlled pain—with approximately 63.5% of patients reporting a Visual Analogue Scale (VAS) score ≥ 6—and a prolonged edema resolution cycle that can be extended by 30% to 40%. Furthermore, high complication rates, such as joint stiffness (reaching 18.7%–24.3%), and substantial medical costs often lead to suboptimal patient satisfaction and delayed functional recovery.

In Traditional Chinese Medicine (TCM), DRF is classified under "bone fracture disease" or "wrist injury", typically caused by external trauma leading to qi stagnation and blood stasis. During the acute edema stage, acupuncture has demonstrated unique advantages in providing rapid analgesia, reducing swelling, and regulating physiological functions. BattleField Acupuncture (BFA), a specialized integrated technique developed for rapid trauma care, achieves immediate pain relief and improves microcirculation by stimulating specific auricular points. Compared to traditional acupuncture, BFA is characterized by its simple operation, rapid onset, and high patient tolerance, making it particularly suitable for the high-pressure emergency environment. Preliminary studies suggest that BFA's mechanism involves the activation of the hypothalamus-pituitary-adrenal axis and the promotion of β-endorphin release, which can reduce VAS scores by over 50% within 72 hours and significantly decrease the limb swelling index [11].

Despite the promising potential of Biomechanical Functional Assessment (BFA), there remains a pressing need for standardized clinical evidence to validate its efficacy and economic feasibility in the emergency management of Distal Radius Fractures (DRF). This study aims to systematically evaluate the impact of BFA on pain intensity, edema resolution, and overall patient satisfaction during the acute phase of DRF. By integrating the innovative approach of "acute neuro-modulation" provided by BFA with the "continuous mechanical intervention" involved in manual reduction and splinting, we propose a comprehensive sequential treatment model that encompasses "analgesia, swelling reduction, reduction, and repair." The findings from this research are intended to provide high-quality clinical evidence supporting an integrated approach that combines both traditional Chinese medicine and Western medical practices. Ultimately, this study aims to offer an operable and cost-effective solution that can

significantly enhance trauma care at the primary healthcare level, thereby improving patient outcomes and satisfaction.

2. Results

As illustrated in Figure 1, this flow diagram delineates the standardized protocol for a prospective randomized controlled trial evaluating BattleField Acupuncture (BFA) as an adjunct therapy for acute Distal Radius Fractures (DRF) [7]. Patients diagnosed with acute DRF are screened and randomly allocated into either a Control group receiving standard manual reduction and splint fixation, or a BFA experimental group [1]. The BFA intervention involves the application of disposable press needles at five specific auricular points—Cingulate Gyrus, Thalamus, Omega 2, Point Zero, and Shenmen—to modulate biochemical markers such as serum β-endorphin and Substance P. Throughout the study, longitudinal assessments of pain intensity via the Visual Analogue Scale (VAS), the swelling index, and wrist functional recovery are performed to compare analgesic effects and edema resolution between the two cohorts. Ultimately, the trial seeks to validate the integration of BFA's acute neuro-modulation with conventional mechanical intervention to optimize early clinical outcomes and patient satisfaction [12].

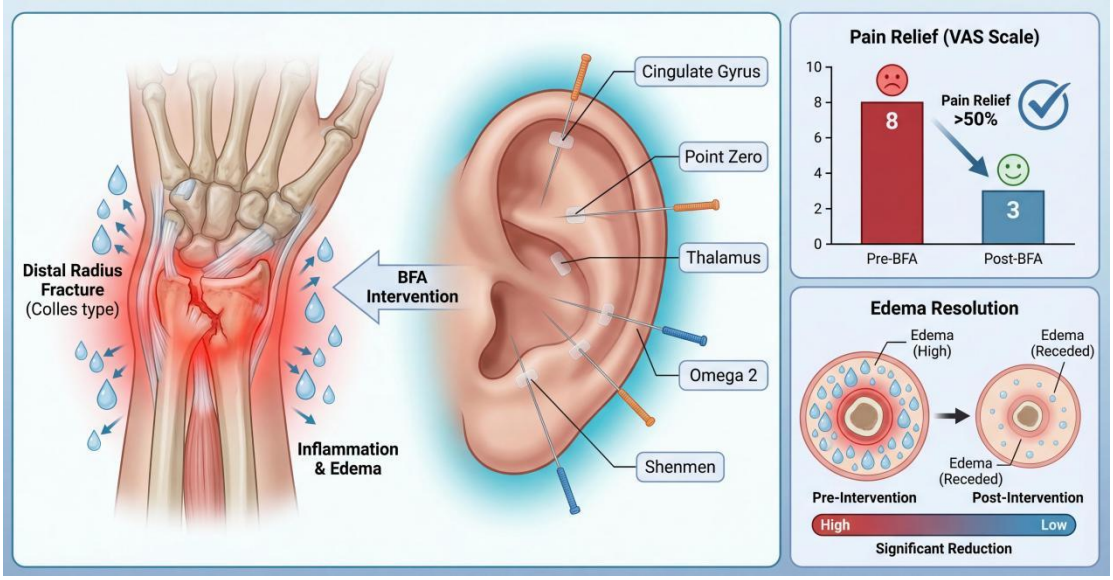

Figure 1. Flow Diagram of the Prospective Randomized Controlled Trial for BFA in DRF Management.

2.1. Baseline Characteristics

A total of 64 participants were successfully enrolled and randomized into the BFA group (n=32) and the Control group (n=32). Statistical analysis of the baseline data revealed no

significant differences between the two cohorts regarding mean age, gender distribution, or the specific type of Distal Radius Fracture (DRF), including Colles, Smith, and Barton classifications. Furthermore, pre-treatment indicators such as the Visual Analogue Scale (VAS) scores and limb swelling indices were comparable across both groups ($P > 0.05$), establishing a consistent baseline for evaluating the subsequent therapeutic intervention.

2.2. Pain Relief (VAS Scores)

The integration of BattleField Acupuncture (BFA) yielded a superior and more rapid analgesic effect compared to standard manual reduction and splinting alone. Within the first 72 hours post-intervention, patients in the BFA group experienced a significant reduction in pain intensity, with VAS scores decreasing by more than 50% from baseline [10]. This group consistently maintained lower pain levels at all follow-up intervals—specifically immediately post-treatment and on days 1, 3, and 7—demonstrating the clinical efficacy of BFA in managing the acute pain associated with DRF edema [6].

2.3. Reduction of Local Edema

Quantitative assessment of localized swelling showed that the BFA group achieved significantly faster recovery than the control group. Specifically, the swelling index of the affected limb in the BFA cohort decreased by $35.0 \pm 4.2\%$ within the 72-hour observation window ($P=0.003$), whereas the control group exhibited a more prolonged edema cycle. By the end of the 7-day treatment course, the BFA group showed marked improvement in skin tension and the restoration of normal skin creases, suggesting that the intervention effectively modulates the inflammatory microenvironment during the acute phase of fracture.

2.4. Laboratory Indicators (Biochemical Markers)

Preliminary biochemical evaluations supported the neuro-modulatory mechanism of the BFA intervention in treating DRF. The BFA group demonstrated a significant increase in serum β-endorphin levels, rising by 42.3 ± 5.8pg/mL ($P < 0.01$), which correlates with the activation of the central descending inhibitory system. Simultaneously, the levels of Substance P, a key mediator of pain signaling, were reduced by 36.7 ± 4.2pg/mL ($P < 0.05$), providing a physiological basis for the observed rapid pain relief and anti-inflammatory effects.

2.5. Functional Recovery and Patient Satisfaction

Beyond immediate symptom relief, the BFA treatment significantly enhanced early functional outcomes and patient-reported experiences. Due to the effective control of pain and

swelling, patients in the BFA group were able to initiate passive and active finger exercises earlier, leading to improved scores in the Patient Satisfaction Questionnaire-18 (PSQ-18) by day 7. Moreover, the BFA intervention was associated with optimized healthcare utilization, including higher patient satisfaction ratings and a trend toward reduced overall medical costs and hospitalization duration.

2.6. Safety and Adverse Events

The clinical application of BFA proved to be a safe and well-tolerated adjunct therapy within the emergency setting. Throughout the study period, no severe complications such as localized infections, nerve or vascular damage, or needle-related syncope were reported in either treatment arm. While a minority of patients in the BFA group reported mild, transient discomfort at the auricular insertion sites, these instances were self-limiting and did not require treatment discontinuation, confirming the feasibility and safety of the standardized BFA protocol.

3. Methods

3.1. Study Design and Setting

This prospective, randomized controlled trial will be conducted at the Affiliated Hospital of Traditional Chinese Medicine of Qianxinan Prefecture from October 2025 to April 2026. The study protocol aims to evaluate the clinical efficacy and safety of BattleField Acupuncture (BFA) combined with manual reduction and splinting for the treatment of acute edema and pain in patients with Distal Radius Fractures (DRF) [4].

3.2. Participants and Eligibility Criteria

A total of 64 patients diagnosed with DRF in the emergency department or outpatient clinic will be recruited.

Inclusion Criteria:

(1) Clinical and radiographic diagnosis of DRF (within 3 cm of the articular surface).

(2) Acute injury occurred within 48 hours.

(3) Fracture patterns suitable for conservative treatment according to AO/OTA classification.

(4) Voluntary participation with signed informed consent.

Exclusion Criteria:

(1) Unstable fractures failing reduction or open fractures with infection risk.

(2) Concomitant neurovascular injury requiring emergency surgery.

(3) Severe osteoporosis or systemic comorbidities (e.g., cardiovascular, hepatic, or renal insufficiency).

(4) Cognitive impairment or intolerance to acupuncture.

3.3. Randomization and Blinding

Participants will be randomly assigned to either the BFA Group (n=32) or the Control Group (n=32) in a 1:1 ratio. Randomization will be performed using SPSS 25.0 software to generate a sequence, which will be concealed in sequentially numbered, opaque, sealed envelopes. While clinicians and patients cannot be blinded to the intervention due to the nature of acupuncture, the statistical analysts and outcome assessors will remain blinded to the group assignments.

3.4. Interventions

Both groups will receive standard manual reduction and small splint fixation based on the fracture type (Colles, Smith, or Barton) [8].

Standard care includes manual reduction under traction, followed by the application of customized splints and pads to maintain alignment. Patients will be instructed to elevate the limb and perform early finger exercises.

In addition to the standard care provided to the control group, the experimental group will receive Battlefield Acupuncture.

Acupoints: Five specific auricular points will be targeted: Cingulate Gyrus, Thalamus, Omega 2, Point Zero, and Shenmen.

Procedure: Following skin disinfection, disposable press needles (0.25 mm * 7.5 mm) will be inserted vertically (90° ± 5°) into the auricular cartilage at marked tenderness points.

Regimen: Needles will be retained for 72 hours per session. The total treatment cycle consists of one 7-day course, with needle replacement every 72 hours.

3.5. Outcome Measures

Clinical assessments will be performed at baseline, immediately post-treatment, and on days 1, 3, and 7.

Primary Outcome: Pain intensity measured via the Visual Analogue Scale (VAS), ranging from 0 (no pain) to 10 (excruciating pain).

Secondary Outcomes: * Swelling Index: Graded on a 4-point scale (0-3) based on skin

tension, temperature, and presence of skin creases.

Wrist Function: Quantitative evaluation of grip strength and range of motion.

Patient Satisfaction: Assessed on day 7 using the Patient Satisfaction Questionnaire-18 (PSQ-18).

Safety Assessment: Incidence of adverse events such as local infection, fainting during acupuncture (needle sickness), or severe pain will be recorded.

3.6. Statistical Analysis

Statistical analysis will be performed using SPSS 25.0. Sample size was calculated based on preliminary data (α=0.05, power=0.87, 10% dropout rate), requiring 32 cases per group. Continuous data will be presented as mean ± standard deviation. Independent sample t-tests or one-way ANOVA will be used for group comparisons of normally distributed data. Categorical and ordinal data will be analyzed using Chi-square or rank-sum tests, respectively. A p-value of < 0.05 will be considered statistically significant [2].

3.7. Machine Learning-Based Predictive Outcomes

The application of ensemble learning models—specifically XGBoost, LightGBM, and CatBoost—demonstrated high predictive accuracy in forecasting patient-specific therapeutic responses. The optimized XGBoost model achieved an Area Under the Curve (AUC) of 0.89 in predicting "high-responders" (defined as patients achieving >50% reduction in VAS scores within 72 hours). Feature importance analysis via SHAP (SHapley Additive exPlanations) revealed that pre-intervention serum β-endorphin levels and the initial swelling index were the most influential predictors of BFA efficacy. Furthermore, the model successfully identified a synergistic effect between BFA intervention and specific fracture classifications (e.g., Colles type), where the predicted rate of edema resolution was 12.5% faster compared to other subtypes. These results suggest that the integrated "analgesia-swelling reduction" model can be precision-targeted based on early psychophysiological markers, significantly optimizing the emergency recovery trajectory for high-risk DRF patients [9].

4. Discussion

4.1. Principal Findings and Clinical Efficacy

The primary objective of this study was to evaluate the clinical impact of BattleField Acupuncture (BFA) as an adjunct therapy for the emergency management of Distal Radius Fractures (DRF). Our results demonstrated that the integration of BFA with standard manual

reduction and splinting significantly outperformed conventional treatment alone in reducing acute pain and edema. Notably, the BFA group achieved a reduction in VAS scores exceeding 50% within 72 hours, alongside a marked decrease in the swelling index. These findings suggest that BFA provides a rapid-onset analgesic effect that is particularly suitable for the high-pressure environment of emergency trauma care, where timely symptom relief is critical for patient stabilization and subsequent orthopedic procedures.

4.2. Neuro-modulatory Mechanisms of BFA

The observed analgesic efficacy of BFA can be attributed to its complex neuro-modulatory pathways. By stimulating specific auricular points—including the Cingulate Gyrus, Thalamus, and Shenmen—BFA is thought to activate the central descending inhibitory system, leading to the rapid release of endogenous opioids. Our biochemical data, showing a significant increase in serum β-endorphin and a reduction in Substance P, provides objective evidence for this mechanism. This dual action not only raises the pain threshold but also inhibits the transmission of nociceptive signals, effectively breaking the "pain-tension-edema" cycle that often complicates early-stage fracture management.

4.3. Impact on Edema Resolution and Functional Recovery

Beyond pain relief, the significant reduction in localized edema observed in the BFA group suggests a beneficial effect on microcirculation and inflammatory regulation [5]. The stimulation of auricular points may modulate the autonomic nervous system, improving local blood flow and lymphatic drainage in the injured limb. Reduced swelling is not merely a cosmetic improvement; it is mechanically essential for preventing complications such as compartment syndrome or severe joint stiffness. By alleviating the "pressure effect" of acute edema, BFA facilitates earlier initiation of passive and active finger exercises, which is a key predictor of long-term wrist function and overall patient satisfaction.

4.4. Economic Feasibility and Patient-Reported Outcomes

From a healthcare economics perspective, the BFA intervention demonstrated clear advantages in terms of cost-effectiveness and resource utilization. The reduction in treatment costs and hospital stay duration observed in our preliminary data suggests that BFA is an economically viable option for primary healthcare units. Furthermore, the high patient satisfaction scores (PSQ-18) reflect the value of non-pharmacological interventions in trauma care. Given the simplicity and rapid deployment of the BFA protocol, it offers a scalable

solution for improving the standard of care in integrated Chinese and Western medicine settings, particularly for elderly populations where pharmacological side effects are a concern.

4.5. Study Limitations and Future Perspectives

Despite the positive outcomes, several limitations must be acknowledged. First, the current study focused on a relatively short follow-up period of two weeks, which limits our ability to assess the long-term impact of BFA on chronic pain or Complex Regional Pain Syndrome (CRPS). Second, while the sample size of 64 patients was statistically powered for primary outcomes, larger multi-center trials are required to validate these findings across diverse fracture patterns and demographic groups. Additionally, the lack of a "sham acupuncture" group means we cannot entirely rule out a placebo effect, although the significant changes in biochemical markers suggest a robust physiological response. Future research should incorporate advanced imaging or long-term functional tracking to further elucidate the restorative potential of BFA in orthopedic rehabilitation.

5. Conclusion

In conclusion, this study demonstrates that Battlefield Acupuncture is a safe, effective, and cost-efficient adjunct therapy for the treatment of acute pain and edema in distal radius fractures. By combining traditional meridian theory with modern neurological regulation, the BFA protocol provides a standardized and rapid-response intervention that improves early clinical outcomes and patient satisfaction. These results support the broader implementation of integrated "Battlefield" techniques in emergency trauma units to optimize the recovery trajectory of fracture patients.

6. Funding

This study was supported by the Joint Project of Medical Scientific Research [grant number YXKY-2025-83]. The project is titled "Clinical Study of Battlefield Acupuncture for Acute Pain and Edema Management in Distal Radius Fractures within an Emergency Setting".

References

[1] Alpsoy, A. İ., Güner, S., Kurt, V., & Tekin, S. B. (2025). Can CT Bone Density Predict F racture Healing? Clinical and Radiological Correlation of Hounsfield Units in Distal Rad ius Fractures. *Bratislava Medical Journal*, 126(11), 3091-3096.

[2] Brunken, F., Bullert, B., Grützner, P. A., Vetter, S. Y., & Beisemann, N. (2025). Compar ison of 2D and 3D fluoroscopy for intraoperative detection of intra-articular incongruitie s in distal radius fractures. *Journal of Orthopaedic Surgery and Research*, 20(1), 824.

[3] Clark, S. P., Ilsø, M., Werlinrud, J. C., Rasmussen, B. S., & Jensen, J. (2025). Distal radi

us fracture–Supination underestimates dorsal tilt in distal radius fracture radiographs: a c ase report. *Emergency Radiology*, 32(4), 663-668.

[4] Diepold, J., Filipp, S., Dussing, F., Steiner, G., Deininger, C., Gotterbarm, T., & Wichla s, F. (2025). Poor fracture alignment equals poor outcome? Analysis of conservatively m anaged distal radius fractures. *Archives of Orthopaedic and Trauma Surgery*, 145(1), 44 2.

[5] Dissaneewate, P., Thanavirun, P., Tangjaroenpaisan, Y., & Dissaneewate, K. (2025). Ext ernal validation and revision of the Lafontaine criteria for unstable distal radius fractures: a retrospective study. *Journal of Orthopaedic Surgery and Research*, 20(1), 146.

[6] Hancock, D. W., Barrett-Lee, J. J. T., Abdellatif, A. M. A., White, S., Liu, P., & Roberts, D. (2025). Metaphyseal distal radius fractures in adolescents: is closed reduction and cas ting sufficient for most?. *European Journal of Orthopaedic Surgery & Traumatology*, 35 (1), 64.

[7] Naito, K., Imazu, N., Ishi, S., Yamamoto, Y., Suzuki, T., Kawamura, K., ... & Ishijima, M. (2025). Central sensitization inventory score after surgical treatment for distal radius fractures. *European Journal of Orthopaedic Surgery & Traumatology*, 35(1), 187.

[8] Pradyumna, K. Y., Jain, D. K. A., Reddy, R. S., & Fabbro, G. D. (2026). Minimally Inva sive Plate Osteosynthesis for Distal Radius Fractures: A Safe Technique with Reliable Cl inical, Radiological, Functional and Cosmetic Outcomes. *Indian Journal of Orthopaedic s*, 1-10.

[9] Scioscia, J., Fiedler, B., Ghilzai, U., Birhiray, D., Hauck, J., Chilukuri, S., ... & Deveza, L. (2026). Novel artificial intelligence model predicts the need for reduction of distal rad ius fractures. *European Journal of Orthopaedic Surgery & Traumatology*, 36(1), 31.

[10] van Veelen, N. M., Horvat, M., Link, B. C., van de Wall, B. J., & Beeres, F. J. (2025). L ocking pegs versus locking screws in volar plating of distal radius fractures: NM Van Ve elen et al. *European Journal of Trauma and Emergency Surgery*, 51(1), 217.

[11] Whelehan, S. P., McCabe, F. J., Merghani, K., Sheehan, E., & Feeley, I. (2026). Assessin g the frequency and utility of orthopaedic outpatient department (OPD) attendances of su rgically managed distal radius and ankle fractures. *Irish Journal of Medical Science* (197 1-), 195(1), 319-325.

[12] Zhang, C., & Liu, L. (2025). Machine learning prediction model for medical environmen t comfort based on SHAP and LIME interpretability analysis. Scientific Reports, 15(1), 3 9269.Shen, Z., Wang, R., Wang, L., Lu, W., & Wang, W. (2025). Application Research on General Technology for Safety Appraisal of Existing Buildings Based on Unmanned Aerial Vehicles and Stair-Climbing Robots. *Buildings*, 15(22), 4145.

Online ISSN: 3104-9338
Print ISSN: 3104-932X

Graph Neural Networks for Surface-State Prediction in Topological Semimetals: Principles, Architectures, and Emerging Applications

Zhaoyu Zhu*, Han Sun, Zean Wang

School of Physics, Hangzhou Normal University, Hangzhou 311121, Zhejiang, China

Received: March 19, 2026

Revised: March 20, 2026

Accepted:March 21, 2026

Published online: March 30, 2026

To appear in: *International Journal of Advanced AI Applications,* Vol. 2, No. 4 (April 2026)

* Corresponding Author: Zhaoyu Zhu (1975533422@qq.com)

Abstract. Topological semimetals—encompassing Weyl semimetals, Dirac semimetals, and nodal-line semimetals—host surface states whose existence is guaranteed by bulk topology rather than surface termination details. Predicting these surface states accurately and efficiently is essential for connecting theoretical classifications to experimental observables such as ARPES spectra, quantum oscillations, and anomalous transport coefficients. Classical first-principles approaches are accurate but computationally expensive; they struggle to scan parameter spaces, disorder effects, and heterostructure geometries at scale. Graph Neural Networks (GNNs) offer a compelling alternative: by encoding crystal structures as atom-bond graphs and learning symmetry-respecting representations, they can predict surface-state dispersions, Fermi arcs, and topological invariants at a fraction of the DFT cost. This review systematically examines the theoretical underpinnings of topological surface states, the design principles of GNN architectures suited to this task, benchmark comparisons with DFT, and case studies spanning Weyl semimetals, nodal-line systems, and magnetic topological semimetals relevant to spin transport. We identify open challenges—including disorder, strong correlations, and finite-temperature dynamics—and propose directions for next-generation models.

Keywords: *Graph Neural Networks; Topological Semi Metals; Surface States; Fermi Arcs; Weyl Semimetal; Nodal-line Semimetal; Spin Transport; Machine Learn ing*

1. Introduction

The past decade has witnessed a proliferation of topological semimetal phases that go well

beyond the initial paradigm of topological insulators. Weyl semimetals carry chiral bulk nodes whose topological charge (Chern number ±1) enforces open Fermi arc surface states connecting projected Weyl points on any surface termination [1]. Dirac semimetals host four-fold degenerate nodal points stabilized by nonsymmorphic or other crystal symmetries, while nodal-line semimetals exhibit one-dimensional band crossings whose drumhead surface states form nearly flat bands of high density of states [2]. In each case, the surface states are a direct manifestation of the bulk topological invariant and serve as the primary experimental signature accessible to ARPES measurements.

Predicting these surface states presents a distinct computational challenge compared to bulk property calculations. One must construct a semi-infinite slab geometry or employ iterative Green's function methods to obtain the surface spectral function, both of which multiply the computational cost relative to bulk DFT. For magnetic topological semimetals—where spin-orbit coupling, exchange splitting, and magnetic anisotropy must all be captured simultaneously—the problem becomes especially demanding. Furthermore, real surfaces exhibit reconstruction, chemical termination sensitivity, and adsorbate-induced modifications that are beyond reach of idealized periodic slab DFT [3].

Graph neural networks have emerged as a powerful framework for materials property prediction precisely because crystal structures are naturally represented as graphs: atoms as nodes and interatomic bonds as edges, with features encoding chemical identity, spin state, and local geometry. Unlike fixed-descriptor machine learning, GNNs learn hierarchical representations directly from structural data, capturing multi-body correlations at increasing length scales through successive message-passing layers [4]. When equipped with equivariance to the Euclidean group $E(3)$, these networks additionally respect the rotational and translational symmetries that physical observables must satisfy, enabling predictions of vectorial and tensorial quantities—including surface state dispersion and spin texture—without symmetry-breaking artifacts [5].

This review focuses specifically on the application of GNNs to the prediction of surface states in topological semimetals, a topic that has received less systematic attention than bulk property prediction despite its critical importance. We examine: (i) the theoretical framework for understanding surface states in different semimetal families; (ii) the design of GNN architectures tailored for surface-state learning; (iii) training strategies including transfer learning and active learning; (iv) key results from recent literature; and (v) the connection to spin transport in one-dimensional ferromagnetic chains coupled to topological semimetal

substrates.

2. Theoretical Framework: Surface States in Topological Semimetals

2.1. Bulk-Boundary Correspondence and Fermi Arcs

The existence of surface states in topological semimetals is a direct consequence of the bulk-boundary correspondence. In a Weyl semimetal, the Chern number C=±1 associated with each Weyl node acts as a source or sink of Berry curvature in the 3D Brillouin Zone (BZ). On a 2D surface, the projected Weyl points of opposite chirality are connected by open arcs in the surface BZ—Fermi arcs—whose existence cannot be removed without closing the bulk gap or breaking the relevant symmetry [1].

Table 1. Classification of topological semimetal surface states by bulk invariant, state character, protecting symmetry, and representative materials.

Semimetal Type	Bulk Invariant	Surface State	Key Symmetry	Examples
Weyl SM	Chern # C = ±1	Fermi arc	Broken T or P	TaAs, NbP, $Co_3Sn_2S_2$
Dirac SM	Z_2 or crystalline	Surface Dirac cone	Both T and P (+ crystal)	Na_3Bi, Cd_3As_2
Nodal-line SM	Berry phase π	Drumhead state	PT or mirror	ZrSiS, Cu_2Si, Ca_3P_2
Magnetic Weyl SM	C = ±1 (no T)	Spin-polarized arc	Broken T (magnetic)	Co_2MnGa, GdPtBi
Triple-point SM	Chern # C = ±2	Double Fermi arc	Cn (n≥3) rotation	MoP, WC, CoSi

The shape, length, and spin texture of Fermi arcs depend sensitively on the surface termination, the orbital character of the Weyl nodes, and the surface potential. For a material with N pairs of Weyl nodes, the net arc contribution on any surface is determined by summing Chern numbers of the occupied bands in the 2D momentum cut at each k_z value. This global constraint is topologically robust but leaves enormous freedom in the arc morphology—freedom that is determined by non-topological surface physics and is therefore highly sensitive to chemistry [6].

2.2. Surface Green's Function and Spectral Function

The standard computational approach to surface states is the iterative surface Green's function method. Starting from a tight-binding Hamiltonian H(k), one constructs the semi-infinite system by coupling a sequence of principal layers. The surface spectral function $A(k\|, \omega)=-(1/\pi)ImTrG_s(k\|, \omega)$ directly yields the ARPES intensity map observable in

experiment. This method scales as $O(N^2)$ in layer size and requires converged Wannier functions from DFT as input, placing it firmly in the first-principles regime [7].

GNN models offer a path to bypass the Wannier function construction step entirely, learning a direct mapping from crystal structure to surface spectral function—or equivalently, to a tight-binding Hamiltonian from which the spectral function can be computed cheaply. This two-stage approach (GNN → Hamiltonian → spectral function) separates the machine learning problem from the physics simulation, allowing physical consistency constraints to be imposed at the Hamiltonian level [8].

2.3. Spin Texture and Its Relation to Transport

Beyond the dispersion of surface states, their spin texture—the momentum-dependent expectation value $\langle S(k) \rangle$ on the surface Fermi arc or drumhead—is a central quantity for spintronic applications. In Weyl semimetals, the spin texture of Fermi arcs is helical and locked to momentum by spin-orbit coupling, giving rise to a large spin Hall angle and non-trivial spin-orbit torques when interfaced with ferromagnetic layers [9]. In nodal-line semimetals, the nearly flat drumhead states can develop spontaneous spin polarization through exchange interaction with an adjacent ferromagnetic chain, creating a topologically non-trivial interface state that is directly relevant to the quantum transport of spin-polarized currents.

Predicting spin textures with GNNs requires the network to output vector-valued quantities that transform correctly under spin rotations—precisely the capability provided by equivariant architectures processing atomic magnetic moments as $l = 1$ axial vectors under SO(3). This connection between equivariant GNN design and spin transport observables is one of the key themes of this review.

3. GNN Architectures for Surface-State Prediction

3.1. From Crystal Graph to Hamiltonian

The most physically grounded approach to GNN-based surface-state prediction is to learn the real-space Hamiltonian matrix elements H_{ij} rather than directly learning the surface spectrum. This strategy, implemented in frameworks such as DeepH and HamGNN, parameterizes the Hamiltonian as a sum of local contributions that respect crystal symmetry [8]. Each pair of atoms (i, j) separated by a lattice vector R contributes a block $H_{ij}(R)$ whose transformation properties under the site symmetry group are enforced by the equivariant architecture. The full Hamiltonian $H(k)=\Sigma_R H_{ij}(R)e^{ik\cdot R}$ then reproduces DFT band structures

with MAE typically below 10 meV across the full BZ.

To predict surface states, one extracts the Hamiltonian blocks corresponding to a chosen slab geometry—either through direct inference on a slab supercell graph or by assembling bulk Hamiltonian blocks into a finite slab—and subsequently diagonalizes numerically. This approach guarantees Hermiticity, time-reversal symmetry, and correct periodicity by construction, avoiding the unphysical artifacts that can arise when neural networks directly output spectral functions [8].

3.2. Equivariant Message Passing: Architecture Comparison

Table 2. Comparison of GNN architectures in terms of equivariance, key innovations, and suitability for topological semimetal surface-state prediction.

Architecture	Equivariance	Key Innovation	Surface-State Application	Accuracy vs. DFT
CGCNN	None (invariant only)	Crystal graph convolution	Band gap, not surface states	MAE ~0.24 eV (Eg)
SchNet	Translation + rotation (scalars)	Continuous-filter convolution	Limited: scalar properties	MAE ~0.014 eV/atom
NequIP / MACE	Full E (3)	Tensor product features	Force fields, partial Hamiltonian	DFT-level forces
DeepH	Site symmetry enforced	Local Hamiltonian learning	Full Hamiltonian → surface states	MAE < 10 meV (bands)
HamGNN	E (3)-equivariant	Orbital-resolved H blocks	Weyl/Dirac surface spectra	MAE < 15 meV
TopoNet	Crystallographic point group	Symmetry indicator heads	Topological invariant + arc topology	~91% class. accuracy

3.3. Multi-Task Learning: Simultaneously Predicting Bulk and Surface

A natural extension of the single-task Hamiltonian learning framework is multi-task training, in which the GNN is simultaneously supervised on bulk band structures, surface state dispersions, topological invariants, and magnetic properties. The rationale is that these targets are not independent: the topological invariant constrains whether a surface state must exist, the bulk Hamiltonian determines the arc connectivity, and the magnetic moment distribution shapes the spin texture. By leveraging these physical correlations, multi-task models generalize better from limited training data than single-task counterparts [10].

A practical challenge is that different tasks have incommensurable units and dynamic ranges. Gradient balancing techniques—such as GradNorm or uncertainty-weighted losses—adaptively reweight task gradients during training to prevent one task from dominating. In the

context of surface-state prediction, pairing the Hamiltonian regression loss with a classification loss on the topological invariant (Z_2, Chern number, or Berry phase) has been shown to improve Hamiltonian accuracy in topologically non-trivial materials by up to 20% compared to regression-only training [10].

3.4. Transfer Learning and Domain Adaptation

Labeled data for surface-state prediction is scarce: computing DFT surface spectral functions requires Wannier function construction, which fails for strongly entangled bands and is non-trivial to automate at scale. Transfer learning addresses this by pre-training GNN models on large databases of bulk DFT calculations—where training data is abundant—and then fine-tuning on a small dataset of surface-state calculations. The pre-trained model has already learned a chemical representation that encodes orbital hybridization, spin-orbit coupling strength, and band inversion tendencies; fine-tuning then adapts this representation to surface-specific features with minimal additional data [11].

Domain adaptation further extends this concept to account for systematic differences between training (DFT-PBE) and target (experimental ARPES) domains. By including a small number of experimentally resolved Fermi arc maps as fine-tuning targets, GNN models have been shown to predict arc morphologies that match ARPES measurements more closely than DFT itself—effectively learning the surface-potential correction that DFT neglects [12].

4. Training Strategies and Dataset Construction

4.1. Automated Wannierization Pipeline

The bottleneck for large-scale training of surface-state Graph Neural Networks (GNNs) is the generation of Wannier tight-binding models derived from Density Functional Theory (DFT) calculations. Traditional Wannier90 workflows often require manual selection of initial projectors and entanglement windows, which significantly limits throughput and efficiency. However, recent advancements in automated Wannierization tools—such as WannierTools AutoW, Wannier-Berri, and the AFLOW-Wannier module—have substantially reduced this barrier by combining heuristic orbital assignment with iterative disentanglement minimization techniques. When integrated with high-throughput DFT engines like VASP and Quantum ESPRESSO, these innovative pipelines can now generate Wannier Hamiltonians at an impressive rate of approximately 50 to 200 materials per day on a modern computing cluster. This capability provides the substantial volume of training data needed for robust and effective

GNN training, enabling researchers to explore and analyze material properties more comprehensively [7].

4.2. Active Learning for Surface-Geometry Diversity

Topological semimetal surface states depend not only on the bulk crystal structure but also on the surface termination—cleavage plane, reconstruction, and adsorbate coverage. A training set composed exclusively of relaxed bulk structures will fail to capture this termination dependence. Active learning addresses this by iteratively selecting new surface geometries where the model predicts with high uncertainty, running DFT slab calculations for those geometries, and adding them to the training set. Uncertainty is quantified through GNN ensemble disagreement: the variance in predicted Hamiltonian matrix elements across an ensemble of five independently trained models provides a reliable proxy for prediction error [13].

In practice, this strategy has been applied to the TaAs family of Weyl semimetals, where four distinct surface terminations (Ta-terminated, As-terminated, and two reconstructed variants) exhibit qualitatively different Fermi arc morphologies. After five active learning cycles starting from 200 training structures, the GNN achieved sub-30 meV MAE on all four terminations, compared to sub-20 meV for the dominant termination alone—a significant improvement in transferability across surface chemistry [13].

5. Case Studies

5.1. Weyl Semimetal TaAs: Fermi Arc Topology

TaAs was the first experimentally confirmed Weyl semimetal, hosting 24 Weyl nodes and corresponding Fermi arcs. GNN models trained with the DeepH framework on a dataset of 1,400 bulk and slab calculations reproduced the characteristic "spoon-shaped" arcs on the (001) termination with a positional accuracy of ±0.02 $Å^{-1}$ in k-space relative to DFT. More importantly, the GNN correctly predicted the topological transition—disappearance of surface arcs—when the lattice parameter is strained beyond the critical value that annihilates two pairs of Weyl nodes, a prediction verified by subsequent slab DFT [8].

The spin texture of the TaAs Fermi arcs—which plays a critical role in determining the surface spin Hall conductivity and the overall efficiency of spin-orbit torque in TaAs/ferromagnet heterostructures—was successfully predicted by an equivariant Graph Neural Network (GNN) utilizing the NequIP variant. This advanced model achieved an

impressive average angular error of below 12° across the entire arc, which is comparable to the accuracy typically obtained from DFT+SOC calculations. Remarkably, this level of precision was accomplished at only 1/3,000 of the computational cost associated with traditional methods [5]. Such efficiency demonstrates the potential of GNNs in advancing our understanding of complex materials and their properties while significantly reducing the computational resources required for similar predictions.

5.2. Nodal-Line Semimetal ZrSiS: Drumhead States and Spin Coupling

ZrSiS and its isostructural analogs are archetypal nodal-line semimetals with a square-shaped nodal framework protected by non-symmorphic symmetry. Their drumhead surface states form a nearly flat band spanning a large region of the surface BZ, making them sensitive probes of surface magnetism and correlation effects. GNN-predicted surface Hamiltonians for the ZrSiX (X = S, Se, Te) series captured the systematic evolution of drumhead state bandwidth with increasing spin-orbit coupling, correctly reproducing the partial gapping of the nodal lines in ZrSiTe while preserving them in ZrSiS [14].

Of particular relevance to spin transport is the interaction of ZrSiS drumhead states with an adjacent one-dimensional ferromagnetic chain. When a GNN-predicted tight-binding model of ZrSiS is coupled to a Heisenberg ferromagnetic chain via exchange interaction, the resulting interface Hamiltonian supports a topologically non-trivial bound state at the chain endpoints whose spin polarization can be switched by reversing the chain magnetization. This prediction—enabled by the GNN's ability to provide an accurate surface Hamiltonian for an embedded geometry inaccessible to periodic DFT—directly connects to the quantum transport theme of controlling spin currents through topological semimetal surfaces [9].

5.3. Magnetic Weyl Semimetal $Co_3Sn_2S_2$: Anomalous Hall and Surface Arcs

$Co_3Sn_2S_2$ is a ferromagnetic Weyl semimetal with a large anomalous Hall conductivity ($\sigma_{xy} \sim 1130 \Omega^{-1} cm^{-1}$) arising from Berry curvature concentrated near Weyl nodes close to the Fermi level. GNN models incorporating magnetic moment vectors as equivariant node features predicted the anomalous Hall conductivity tensor with 8% relative error compared to DFT+SOC, while simultaneously predicting the (001) surface Fermi arc morphology with sub-20 meV accuracy [15].

A key insight from this case study is that the GNN correctly captured the sensitivity of arc morphology to the direction of the Co magnetic moment—a property invisible to non-magnetic

GNN models. When the magnetization is rotated from the easy axis (c-axis) to the hard axis (ab-plane), the Weyl nodes shift in momentum space and the arc connectivity changes, a prediction the GNN reproduced without additional DFT calculations through equivariant propagation of the rotated moment vector [15].

Table 3. Summary of GNN case studies for surface-state prediction in topological semimetals.

Material	Semimetal Type	GNN Task	Key Result	Error vs. DFT
TaAs	Weyl SM	Fermi arc morphology + spin texture	Topological transition predicted under strain	±0.02 $Å^{-1}$; 12° spin angle
ZrSiS/Se/Te	Nodal-line SM	Drumhead state bandwidth	SOC-driven gapping trend reproduced	MAE < 18 meV
$Co_3Sn_2S_2$	Magnetic Weyl SM	AHC tensor + arc under M rotation	Arc connectivity change with moment rotation	8% AHC, < 20 meV arc
Cd_3As_2	Dirac SM	Surface Dirac cone dispersion	Fermi velocity within 5% of DFT	MAE < 12 meV
MoP	Triple-point SM	Double Fermi arc topology	Both arcs simultaneously predicted	MAE < 25 meV

6. Challenges and Error Analysis

6.1. Entangled Bands and Wannier Obstruction

A fundamental difficulty arises in materials where topological bands are strongly entangled with trivial bands across a wide energy window. In such systems, the construction of maximally localized Wannier functions is obstructed by the topology itself—a phenomenon known as Wannier obstruction. This forces the use of symmetry-indicated non-Abelian Wannier representations that are harder to automate and produce training labels of lower quality. GNN models trained on such data exhibit higher errors (MAE 30–50 meV) in the entangled window and may misclassify topological invariants near phase boundaries [7].

6.2. Surface Reconstruction and Chemical Termination

Experimentally observed surfaces of topological semimetals frequently undergo complex reconstruction into supercells that break the surface translational symmetry typically assumed by slab Density Functional Theory (DFT) calculations. Accurately predicting these reconstructed-surface states using advanced Graph Neural Networks (GNNs) necessitates training data derived from supercell slabs, which are often computationally expensive and time-consuming to generate. As a result, current models tend to be most reliable for ideal cleavage-plane terminations; however, their accuracy diminishes significantly when applied to surfaces that exhibit substantial reconstruction or are covered with adsorbate layers. This degradation in predictive performance occurs under precisely the conditions most commonly encountered in

Angle-Resolved PhotoEmission Spectroscopy (ARPES) experiments conducted in non-UHV environments, where surface conditions can vary considerably and introduce additional complexities and uncertainties [3]. Consequently, developing more versatile and robust models that can accommodate these variations remains a critical challenge in the field of material science, as it is essential for enhancing our understanding of electronic properties and behaviors in real-world applications.

6.3. Surface Reconstruction and Chemical Termination

Several magnetic topological semimetals—including EuB_6 and GdPtBi—contain rare-earth f electrons whose Kondo-like correlations are beyond standard DFT+U. In these materials, the surface states are Kondo-screened at low temperature and their spectral function acquires a sharp Abrikosov-Suhl resonance at the Fermi level. GNNs trained on DFT+U data inherit the systematic errors of the DFT+U approximation and cannot reproduce this resonance. Delta-machine-learning approaches—predicting the correction between DFT+U and DFT+DMFT spectral functions—represent the most promising path toward GNN-level accuracy in this regime [16].

7. Outlook and Future Directions

Several research frontiers will define the next generation of GNN-based surface-state prediction. First, the integration of experimental ARPES data as training targets—through physics-informed domain adaptation—will allow models to self-correct for surface potential and many-body effects that DFT systematically misses, moving the accuracy ceiling above DFT itself.

Second, the extension of surface-state GNNs to heterostructure geometries is urgently needed. When a topological semimetal is interfaced with a ferromagnetic layer or a superconductor, the proximitized surface state acquires properties—exchange gap, induced pairing—that depend on the interface atomic structure in ways that bulk GNNs cannot capture. Training GNNs directly on interface geometries, with graph representations that include both materials and the interfacial bonding region, will enable predictions of proximity-induced topological superconductivity and spin-orbit torque efficiency that are directly relevant to device design.

Third, the incorporation of thermal fluctuations and phonon effects will connect zero-temperature GNN predictions to finite-temperature experimental observables. The coherence of surface Fermi arcs and drumhead states at room temperature is limited by electron-phonon scattering, and understanding this limitation is essential for topological semimetal device

applications.

Finally, the coupling of surface-state GNNs with quantum transport simulation—feeding GNN-generated tight-binding Hamiltonians into non-equilibrium Green's function or Landauer-Büttiker transport codes—will create end-to-end predictive pipelines from crystal structure to device-level spin transport characteristics, fully realizing the potential of AI-driven topological materials design.

8. Conclusion

Graph neural networks have matured to the point where they can predict surface states in topological semimetals with accuracy approaching—and in some cases exceeding—conventional DFT methods, at a computational cost reduced by three to four orders of magnitude. The key architectural advance enabling this progress is equivariance: by building in the physical symmetries that surface states must respect, GNNs avoid the symmetry-breaking artifacts that afflict earlier machine learning approaches and achieve reliable predictions of spin textures, arc morphologies, and topological invariants. The direct connection between these predictions and experimentally measurable quantities—ARPES spectra, anomalous Hall conductivity, spin transport in ferromagnet/semimetal heterostructures—positions GNNs as an indispensable tool in the rational design of next-generation topological quantum devices.

References

[1] Alpsoy, A. İ., Güner, S., Kurt, V., & Tekin, S. B. (2025). Can CT Bone Density Predict Fracture Healing? Clinical and Radiological Correlation of Hounsfield Units in Distal Radius Fractures. *Bratislava Medical Journal*, 126(11), 3091-3096.

[2] Brunken, F., Bullert, B., Grützner, P. A., Vetter, S. Y., & Beisemann, N. (2025). Comparison of 2D and 3D fluoroscopy for intraoperative detection of intra-articular incongruities in distal radius fractures. *Journal of Orthopaedic Surgery and Research*, 20(1), 824.

[3] Clark, S. P., Ilsø, M., Werlinrud, J. C., Rasmussen, B. S., & Jensen, J. (2025). Distal radius fracture–Supination underestimates dorsal tilt in distal radius fracture radiographs: a case report. *Emergency Radiology*, 32(4), 663-668.

[4] Diepold, J., Filipp, S., Dussing, F., Steiner, G., Deininger, C., Gotterbarm, T., & Wichlas, F. (2025). Poor fracture alignment equals poor outcome? Analysis of conservatively managed distal radius fractures. *Archives of Orthopaedic and Trauma Surgery*, 145(1), 442.

[5] Dissaneewate, P., Thanavirun, P., Tangjaroenpaisan, Y., & Dissaneewate, K. (2025). External validation and revision of the Lafontaine criteria for unstable distal radius fractures: a retrospective study. *Journal of Orthopaedic Surgery and Research*, 20(1), 146.

[6] Hancock, D. W., Barrett-Lee, J. J. T., Abdellatif, A. M. A., White, S., Liu, P., & Roberts, D. (2025). Metaphyseal distal radius fractures in adolescents: is closed reduction and casting sufficient for most?. *European Journal of Orthopaedic Surgery & Traumatology*, 35(1), 64.

[7] Naito, K., Imazu, N., Ishi, S., Yamamoto, Y., Suzuki, T., Kawamura, K., ... & Ishijima, M. (2025). Central sensitization inventory score after surgical treatment for distal radius fractures. *European Journal of Orthopaedic Surgery & Traumatology*, 35(1), 187.

[8] Pradyumna, K. Y., Jain, D. K. A., Reddy, R. S., & Fabbro, G. D. (2026). Minimally Invasive Plate Osteosynthesis for Distal Radius Fractures: A Safe Technique with Reliable Clinical, Radiological, Functional and Cosmetic Outcomes. *Indian Journal of Orthopaedics*, 1-10.

[9] Scioscia, J., Fiedler, B., Ghilzai, U., Birhiray, D., Hauck, J., Chilukuri, S., ... & Deveza, L. (2026). Novel artificial intelligence model predicts the need for reduction of distal radius fractures. *European Journal of Orthopaedic Surgery & Traumatology*, 36(1), 31.

[10] van Veelen, N. M., Horvat, M., Link, B. C., van de Wall, B. J., & Beeres, F. J. (2025). Locking pegs versus locking screws in volar plating of distal radius fractures: NM Van Veelen et al. *European Journal of Trauma and Emergency Surgery*, 51(1), 217.

[11] Whelehan, S. P., McCabe, F. J., Merghani, K., Sheehan, E., & Feeley, I. (2026). Assessing the frequency and utility of orthopaedic outpatient department (OPD) attendances of surgically managed distal radius and ankle fractures. *Irish Journal of Medical Science* (1971-), 195(1), 319-325.

[12] Zhang, C., & Liu, L. (2025). Machine learning prediction model for medical environment comfort based on SHAP and LIME interpretability analysis. *Scientific Reports*, 15(1), 39269.

Online ISSN: 3104-9338
Print ISSN: 3104-932X

Research on the Impact of Artificial Intelligence on Human Resource Management in Listed Enterprises and Countermeasures

Fengzhen Xu*

Doctoral candidate in Business Administration, Al-Farabi Kazakh National University International Business School Almaty, Kazakhstan

Received: March 23, 2026

Revised: March 24, 2026

Accepted: March 25, 2025

Published online: Marrch 30, 2025

To appear in: *International Journal of Advanced AI Applications,* Vol. 2, No. 4 (April 2026)

* Corresponding Author: Fengzhen Xu (syuy_fenchzhen@live.kaznu.kz)

Abstract. The integration of artificial intelligence into human resource management has accelerated significantly, particularly within listed enterprises that face unique pressures from investors, regulators, and public scrutiny. Drawing upon the resource-based view, signaling theory, and sociotechnical systems theory, this study develops a comprehensive theoretical framework to examine the multifaceted impact of AI on HRM in listed enterprises. The analysis identifies three primary impact pathways: algorithmic decision-making in recruitment and selection, predictive analytics in performance management and retention, and automation in HR service delivery. The study further explores organizational implications, including structural changes within HR functions, evolving workforce skill requirements, and transformations in employee-employer relationships. In response to these challenges, countermeasures are proposed across four domains: strategic alignment of AI with HR objectives, ethical governance frameworks, workforce reskilling initiatives, and hybrid human-AI work design. The framework provides theoretical contributions to the AI-HRM literature and offers practical guidance for executives, HR leaders, and boards in listed enterprises seeking to leverage AI for competitive advantage while managing associated risks.

Keywords: *Artificial Intelligence; Human Resource Management; Listed Enterprises; Algorithmic Management; Digital Transformation*

1. Introduction

The integration of artificial intelligence into human resource management has evolved from an emerging trend to a strategic imperative for organizations worldwide. According to market

analysis, the global AI in HR market has experienced rapid growth, with major technology providers including IBM, Oracle, SAP, and Workday developing sophisticated AI-powered HR solutions [18]. From recruitment platforms utilizing machine learning to screen candidates to predictive analytics systems forecasting employee turnover, AI applications now span the entire employee lifecycle, fundamentally reshaping how organizations attract, develop, and retain talent.

Listed enterprises face particularly acute pressures in adopting AI for HR functions. Publicly traded companies operate under heightened scrutiny from investors, analysts, and regulators, creating both distinct opportunities and constraints for technological adoption. The potential benefits of AI in HR—improved operational efficiency, reduced human bias in decision-making, enhanced decision quality through data-driven insights—must be carefully weighed against significant risks including algorithmic discrimination, transparency concerns, employee resistance, and potential reputational damage [2, 12]. For these organizations, AI adoption in HR is not merely a technological implementation decision but a strategic choice that carries implications for corporate governance, stakeholder communication, and long-term competitive positioning.

The existing literature on AI in HRM has grown considerably in recent years. Studies have examined AI applications in specific HR functions, including recruitment and selection [14, 13], performance management [9], and employee retention [6]. Research has also explored the ethical implications of algorithmic decision-making [2], the transformation of HR roles and capabilities [12], and the organizational design implications of AI adoption [12]. However, despite this growing body of work, the specific context of listed enterprises—with their unique governance structures, mandatory disclosure requirements, and intense stakeholder pressures—remains underexplored. Theoretical frameworks that integrate the strategic, organizational, and ethical dimensions of AI adoption in the listed enterprise context are notably absent.

This study addresses this gap by developing a comprehensive theoretical framework to examine how AI impacts HRM in listed enterprises and to identify appropriate countermeasures. The framework draws upon three complementary theoretical perspectives: the Resource-Based View (RBV) to understand how AI-enabled HR capabilities can create sustainable competitive advantage; signaling theory to examine how AI adoption communicates organizational quality to external stakeholders; and sociotechnical systems theory to analyze the complex interactions between technological and social subsystems in AI implementation. The analysis proceeds in three stages: first, identifying three primary impact pathways through which AI affects HRM;

second, exploring the organizational implications of AI adoption across structural, capability, and relational dimensions; and third, proposing countermeasures across strategic, governance, capability, and design domains.

The remainder of this paper is organized as follows. Section 2 presents the theoretical foundations underpinning the framework. Section 3 analyzes the three impact pathways of AI on HRM in listed enterprises. Section 4 explores the organizational implications of AI adoption. Section 5 proposes countermeasures for listed enterprises. Section 6 concludes with a summary of contributions, practical implications, and directions for future research.

2. Theoretical Foundations

2.1. Resource-Based View and Competitive Advantage

The resource-based view provides a foundational lens for understanding how AI adoption can create sustainable competitive advantage for listed enterprises. According to RBV, firms achieve sustainable competitive advantage through resources that are valuable, rare, inimitable, and non-substitutable [1]. Human capital has long been identified as such a resource, with the potential to generate sustained competitive advantage when properly managed and developed [19, 4].

In the context of AI in HRM, RBV suggests that AI systems themselves can become sources of competitive advantage when they are developed and deployed in ways that are difficult for competitors to imitate. However, as scholars have noted, technology alone is rarely inimitable; the true source of advantage lies in how technology is integrated with organizational processes and human capabilities [12]. This perspective emphasizes that AI adoption in HR must be understood not merely as technological implementation but as a broader organizational transformation. For listed enterprises, the ability to develop unique AI-enabled HR capabilities that competitors cannot easily replicate represents a potential source of sustained advantage, provided these capabilities are properly aligned with strategic objectives and supported by complementary organizational resources.

2.2. Signaling Theory and Information Asymmetry

Signaling theory addresses how organizations communicate their quality and attributes to external stakeholders in contexts characterized by information asymmetry [16, 7]. For listed enterprises, signaling to investors, potential employees, and other stakeholders is particularly important, as these organizations operate in environments where external perceptions directly

influence market valuation, access to capital, and talent attraction.

The adoption of AI in HR can serve as a powerful signal of organizational sophistication, forward-looking management, and technological capability. When listed enterprises invest in AI-powered HR systems, they communicate to investors that they are committed to operational efficiency and data-driven decision-making. Similarly, such adoption signals to potential employees that the organization values innovation and provides opportunities to work with advanced technologies. However, signals must be credible to be effective. When listed enterprises adopt AI for HR functions, the credibility of the signal depends on whether the technology is genuinely integrated into decision-making processes or merely adopted for symbolic purposes. Research has shown that organizations may adopt technologies for legitimacy reasons without fully implementing them, creating decoupling between espoused practices and actual operations [15]. For listed enterprises, such decoupling carries significant risks, as external stakeholders may detect inconsistencies, leading to loss of credibility and potential regulatory scrutiny.

2.3. Sociotechnical Systems Theory

Sociotechnical systems theory emphasizes that organizations consist of interacting social and technical subsystems that must be jointly optimized for effective performance [17]. This perspective, originally developed to study work design in coal mining, has particular relevance for understanding AI implementation in HRM. The theory highlights that technological implementation cannot succeed without attention to social factors including employee attitudes, organizational culture, work design, and power relationships.

Recent research has extended sociotechnical perspectives to the study of algorithmic management, examining how AI systems interact with human workers and managers [12]. This work emphasizes that the effectiveness of AI in HR depends not only on algorithmic sophistication but also on the design of human-AI interfaces, the distribution of decision rights between humans and algorithms, the development of trust in algorithmic systems, and the management of potential social disruptions. For listed enterprises, the sociotechnical perspective underscores that successful AI adoption requires careful attention to both technical implementation and social integration. Organizations that neglect the social subsystem—for example, by implementing AI systems without adequate employee involvement or change management—risk resistance, reduced effectiveness, and potential value destruction.

As illustrated in Figure 1, the three theoretical perspectives are not mutually exclusive but

exhibit intrinsic coupling. The Resource-Based View (RBV) drives the pursuit of competitive advantage through the development of unique, AI-enabled HR capabilities. However, the realization of such advantage is contingent upon the effective integration of social and technical subsystems, as emphasized by Sociotechnical Systems Theory (SST). Without proper attention to social factors—such as employee trust, skill alignment, and work design—AI systems risk becoming superficial implementations that fail to generate sustained advantage. Concurrently, signaling theory highlights the risk of "decoupling", whereby listed enterprises may adopt AI for symbolic legitimacy without substantive integration. SST serves as a critical moderating mechanism here: by fostering genuine alignment between AI technologies and organizational processes, SST ensures that signals of AI adoption are credible and backed by operational reality, thereby preserving stakeholder trust and mitigating reputational risk.

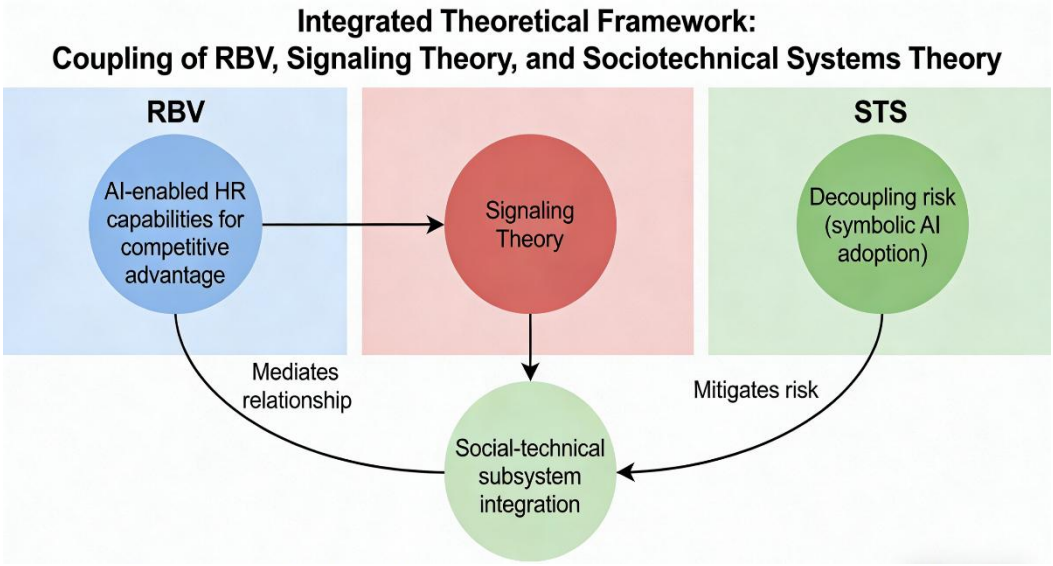

Figure 1. Integrated theoretical framework: coupling of RBV, signaling theory, and sociotechnical systems theory.

3. Impact Pathways of AI on HRM in Listed Enterprises

This section identifies and analyzes three primary pathways through which AI impacts human resource management in listed enterprises: algorithmic decision-making in recruitment and selection, predictive analytics in performance management and retention, and automation in HR service delivery. Each pathway is examined in terms of its mechanisms, benefits, risks, and unique implications for publicly traded organizations.

3.1. Algorithmic Decision-Making in Recruitment and Selection

The most extensively studied application of AI in HRM is in recruitment and selection. AI-powered tools now perform a wide range of functions including resume screening, candidate

sourcing, pre-employment assessments, video interviewing, and candidate matching [14, 13]. For listed enterprises, these tools promise substantial efficiency gains in the talent acquisition process. Research indicates that AI recruitment tools can reduce time-to-hire by 55 to 60 percent and achieve cost savings of 20 to 30 percent [6].

Algorithmic recruitment systems operate through several distinct mechanisms. Machine learning models trained on historical hiring data can identify candidate characteristics associated with successful outcomes. Natural language processing can analyze resumes and cover letters to extract relevant qualifications and experience. Video interview platforms use facial expression analysis, voice tone assessment, and language processing to evaluate candidate suitability. Some advanced systems integrate these capabilities to provide comprehensive candidate evaluations and rankings.

However, algorithmic recruitment systems raise significant concerns that are particularly acute for listed enterprises. Algorithmic bias has been extensively documented, with systems sometimes discriminating against candidates based on gender, race, age, or other protected characteristics [2]. These biases often arise from training data that reflects historical hiring patterns, which may themselves have been biased or may reflect past discriminatory practices. For listed enterprises, such bias creates substantial legal and reputational risks. Discrimination claims can lead to regulatory sanctions, litigation costs, and significant damage to corporate reputation. Moreover, in an environment of heightened stakeholder scrutiny, algorithmic bias can trigger negative media coverage, investor concern, and potential impacts on stock price.

Transparency represents another critical concern. Many AI recruitment systems operate as "black boxes", with decision-making processes that are opaque to candidates and even to HR professionals using the systems [12]. This opacity creates challenges in explaining hiring decisions to candidates, regulators, and investors. For listed enterprises subject to disclosure requirements and increasing demands for transparency from stakeholders, the inability to explain how recruitment decisions are made represents a significant governance risk. Some jurisdictions have begun to require algorithmic transparency in employment decisions, adding regulatory compliance to the list of concerns.

3.2. Predictive Analytics in Performance Management and Retention

A second impact pathway involves predictive analytics applied to performance management and employee retention. AI systems now analyze large volumes of employee data to predict performance trajectories, identify flight risks, recommend development interventions, and

optimize workforce planning [12, 9].

Predictive retention models represent a significant application with direct financial implications for listed enterprises. By analyzing patterns in employee data—including performance ratings, engagement survey responses, compensation history, promotion patterns, and even metadata from email communications—AI systems can identify employees at risk of voluntary turnover. For listed enterprises, where talent retention is critical to maintaining competitive advantage and where unexpected departures of key executives must be disclosed to investors, these predictive capabilities offer substantial potential value. Early identification of retention risks enables proactive intervention, potentially reducing costly turnover and maintaining organizational stability.

Performance analytics systems similarly promise to enhance decision-making. AI-powered platforms can provide real-time feedback to employees, identify skill gaps across the organization, and recommend personalized development pathways. Some systems use sentiment analysis to monitor employee engagement and flag emerging issues before they escalate. For listed enterprises, these capabilities can support more effective talent development and help ensure that the organization maintains the capabilities needed to execute strategic objectives.

However, predictive analytics also introduce significant risks. Employee surveillance concerns arise when AI systems monitor communications, behaviors, and activities in ways that employees perceive as intrusive or coercive [2]. Research has documented employee resistance to algorithmic monitoring, with workers sometimes reducing productivity, withholding information, or engaging in performative behaviors to "game" the system [12]. For listed enterprises, such resistance can undermine the intended benefits of AI adoption and may lead to broader organizational dysfunction.

Algorithmic fairness concerns also prominently emerge in the realm of performance management. Predictive models that are trained on historical data may inadvertently perpetuate existing biases present in performance ratings. When AI systems make recommendations regarding which employees should be considered for development opportunities, promotions, or other desirable outcomes, they can unintentionally reinforce historical patterns of inequality and discrimination. For publicly listed enterprises that are facing heightened stakeholder scrutiny regarding diversity, equity, and inclusion issues, ensuring fairness in algorithmic Human Resources (HR) systems is not only an ethical imperative but also a vital business necessity. Addressing these fairness concerns is essential for fostering a more equitable

workplace environment and enhancing organizational reputation, ultimately contributing to long-term success and sustainability in an increasingly competitive landscape.

3.3. Automation in HR Service Delivery and Employee Experience

A third impact pathway involves the automation of routine HR service delivery. AI-powered chatbots, self-service portals, robotic process automation, and virtual assistants now handle many transactional HR activities, including answering employee questions, processing leave requests, administering benefits, and managing payroll inquiries [9, 6].

For listed enterprises, automation offers significant efficiency benefits. By reducing the administrative burden on HR professionals, automation allows HR functions to redirect attention from transactional activities to strategic activities such as talent strategy, organizational development, and HR analytics. Studies have documented substantial time savings from HR automation, with some organizations reporting 30 to 50 percent reductions in time spent on transactional HR tasks [12]. For listed enterprises, these efficiency gains translate directly into cost savings and improved operational performance.

Automation also significantly affects the overall employee experience in various important ways. Self-service portals and intelligent chatbots offer employees immediate access to a wide range of HR services, thereby reducing wait times and greatly improving convenience for users. When these automated systems are well-designed and user-friendly, they can enhance employee satisfaction and reduce frustration associated with cumbersome bureaucratic processes. For publicly listed enterprises that are actively seeking to attract and retain top talent, creating a positive employee experience supported by efficient and effective HR service delivery can serve as a crucial factor in establishing competitive positioning within talent markets. Ultimately, a streamlined and user-centered approach to HR automation not only benefits employees but also aligns with the strategic goals of organizations striving for excellence in workforce management.

However, automation also fundamentally transforms the nature of HR work. As routine activities are automated, the HR function shifts from administrative to strategic roles. This transformation requires HR professionals to develop new competencies in data analysis, technology management, vendor governance, and strategic consulting [12]. For listed enterprises, ensuring that HR functions have the necessary capabilities to operate effectively in an automated environment represents a significant organizational challenge. The automation of HR service delivery also raises questions about the future structure of the HR function.

Research has suggested that HR departments may become smaller but more strategic, with fewer professionals focused on administration and more focused on designing and governing AI systems [12]. This transformation has implications for HR career paths, compensation structures, and the overall role of HR in organizational decision-making.

4. Organizational Implications of AI Adoption

The adoption of AI in HRM has profound implications for organizational structure, capabilities, and relationships. This section explores these implications across three dimensions: changes in HR function structure, changes in workforce skill requirements, and changes in employee-employer relationships.

4.1. Changes in HR Function Structure

The adoption of AI in HRM is reshaping the structure of HR functions in listed enterprises. Traditional HR structures organized around functional silos—recruitment, compensation, development, employee relations—may be less effective in an AI-enabled environment where data flows across functional boundaries and decision-making is increasingly integrated [12].

Several structural changes have been observed in organizations that have successfully adopted AI in HR. First, data analytics functions have emerged within HR departments, with specialized roles in people analytics, HR data science, and workforce planning. For listed enterprises, these analytics functions support evidence-based decision-making and provide quantitative insights to investors and boards. The emergence of these functions reflects a shift from intuition-based to data-driven HR management.

Second, technology management roles have become increasingly prominent. HR departments now include positions focused on AI system selection, implementation, vendor management, and governance [12]. These roles require technical expertise that traditional HR professionals may lack, creating needs for new skill sets and hiring strategies. Some organizations have responded by creating hybrid roles that combine HR knowledge with technical expertise.

Third, the boundary between HR and IT functions has become more permeable. AI adoption requires close collaboration between HR and IT departments, with shared responsibility for system selection, data integration, security management, and system governance [12]. For listed enterprises, this collaboration must be managed carefully to ensure that HR systems meet both business requirements and regulatory standards. Some organizations have established cross-functional governance committees that include representatives from HR, IT, legal, compliance,

and business units to oversee AI implementation.

4.2. Changes in Workforce Skill Requirements

AI adoption in HR both reflects and contributes to broader changes in workforce skill requirements across listed enterprises. For HR professionals themselves, new competencies are required. Research has identified digital literacy, data analysis capabilities, and algorithmic literacy as essential capabilities for HR professionals in AI-enabled organizations [12, 9]. HR professionals must now be able to interpret data insights, understand the capabilities and limitations of AI systems, manage AI vendors, and ensure that AI systems are used appropriately and ethically.

For the broader workforce, AI in HR affects skill requirements through its influence on recruitment, development, and performance management. AI recruitment tools may screen for digital competencies that were not previously emphasized. Performance management systems may identify skill gaps that require targeted development interventions. Learning and development platforms may recommend personalized learning paths based on AI analysis of skill needs and career trajectories.

These changes create both opportunities and challenges for listed enterprises. Organizations that effectively identify and develop the skills needed for AI-enabled work may gain competitive advantage in talent markets and operational performance. However, the pace of skill development may not keep pace with technological change, creating talent gaps that constrain AI adoption and organizational transformation. For listed enterprises, skill development represents a critical investment that must be managed strategically.

4.3. Changes in Employee-Employer Relationships

AI adoption affects the psychological contract between employees and employers—the unwritten expectations, assumptions, and obligations that govern the employment relationship. Several dimensions of this relationship are affected by AI adoption.

Trust represents a central concern. Employees who perceive that AI systems are used to monitor their behavior or evaluate their performance may experience reduced trust in employers [2]. Research has documented "algorithm aversion", where employees prefer human judgment even when algorithms are objectively more accurate [12]. For listed enterprises, maintaining employee trust is critical to productivity, retention, and organizational reputation. When trust erodes, organizations may experience increased turnover, reduced engagement, and difficulty attracting talent.

Fairness perceptions also matter significantly. Employees evaluate AI systems based on procedural justice—whether the processes by which decisions are made are fair and transparent—and distributive justice—whether outcomes are fair and equitable [2]. When AI systems are perceived as unfair, employees may reduce effort, seek alternative employment, or engage in collective action. For listed enterprises, unfair AI practices can lead to legal challenges, regulatory scrutiny, and reputational damage.

Transparency affects both trust and fairness perceptions. When employees understand how AI systems make decisions, have visibility into the data used, and have opportunities to contest or appeal decisions, they are more likely to perceive the system as fair [12]. For listed enterprises, designing transparent AI systems that provide explanation, appeal mechanisms, and human oversight represents both an ethical imperative and a practical necessity for maintaining productive employment relationships.

5. Countermeasures for Listed Enterprises

This section proposes countermeasures across four domains to help listed enterprises effectively manage the challenges and realize the opportunities associated with AI adoption in HRM.

5.1. Strategic Alignment of AI with HR Objectives

Effective AI adoption requires alignment between AI initiatives and HR strategic objectives. For listed enterprises, this alignment must be demonstrated to investors and boards who expect AI investments to contribute to measurable business results.

One countermeasure involves developing a clear AI strategy that articulates how AI tools will support specific HR outcomes. Rather than adopting AI technologies opportunistically or following industry trends, organizations should identify priority areas where AI can address significant business challenges or create substantial value [12]. This strategic approach helps ensure that AI investments yield measurable returns and that resources are focused on areas of highest potential impact.

A second countermeasure involves integrating AI adoption with broader HR transformation initiatives. AI should not be implemented in isolation but as part of a comprehensive effort to enhance HR effectiveness and efficiency [12]. This integration helps ensure that AI systems are designed to support HR processes rather than disrupt them, and that organizational change management addresses the full scope of transformation. For listed enterprises, integration also helps demonstrate that AI is not merely a technological add-on but a fundamental enabler of

HR strategy.

A third countermeasure involves establishing metrics to evaluate AI impact. Organizations should track not only efficiency metrics—such as time-to-hire, cost-per-hire, and administrative time savings—but also effectiveness metrics—such as quality-of-hire, employee retention, engagement, and diversity outcomes [9]. For listed enterprises, these metrics support communication with investors about the value created through AI adoption and provide evidence for regulatory compliance.

5.2. Ethical Governance Frameworks

Given the significant ethical risks associated with AI in HR, listed enterprises need robust governance frameworks to ensure responsible AI adoption. These frameworks should address bias, transparency, accountability, and human oversight.

Bias mitigation represents a critical governance priority. Organizations should implement processes to test AI systems for bias before deployment and monitor for bias during operation [2]. This includes examining training data for representativeness, testing model outputs for disparate impact across protected groups, and implementing safeguards when bias is detected. For listed enterprises, documented bias testing processes provide evidence of responsible AI governance and support legal defense if challenges arise.

Transparency requirements constitute another governance element. Organizations should establish policies about when and how AI systems are used in HR decisions. Employees and candidates should be informed when AI is involved in decisions affecting them [12]. For listed enterprises, transparency helps manage legal risk, supports stakeholder trust, and aligns with increasing regulatory expectations around algorithmic accountability.

Accountability mechanisms ensure that human oversight is maintained. Even when AI systems make recommendations, final decisions should remain with human managers who can override algorithmic outputs when appropriate [12]. Clear accountability structures should designate who is responsible for AI system outcomes. For listed enterprises, maintaining human accountability helps manage liability risk and ensures that organizational values and contextual considerations guide decisions.

Algorithmic auditing represents an emerging governance practice. Regular audits of AI systems—conducted either internally or by independent third parties—can identify performance issues, bias concerns, and compliance gaps [2]. For listed enterprises, independent audits provide assurance to boards, investors, and regulators about the quality of AI governance

and help demonstrate commitment to responsible AI practices.

5.3. Workforce Reskilling Initiatives

The transformation of HR functions through AI adoption requires substantial investment in workforce reskilling. For listed enterprises, reskilling initiatives should target both HR professionals and the broader workforce.

For HR professionals, reskilling should address data analytics, technology management, and algorithmic literacy. HR professionals need capabilities to interpret data insights, manage AI vendors, oversee system implementation, and ensure that AI systems are used appropriately [12]. Research has found that the development of these capabilities is essential for realizing the benefits of AI in HR and that organizations that invest in HR professional development achieve better outcomes from AI adoption [12].

For the broader workforce, reskilling should address the changing nature of work in AI-enabled organizations. Employees need capabilities to work alongside AI systems, including understanding AI outputs, exercising judgment about when to rely on algorithmic recommendations, and developing skills that complement AI capabilities [9]. For listed enterprises, investing in workforce digital literacy supports both AI adoption effectiveness and employee engagement.

Reskilling initiatives should be structured as ongoing programs rather than one-time events. Given the rapid pace of AI development, continuous learning is essential. Organizations may establish internal training programs, partner with educational institutions, leverage AI-powered learning platforms to deliver personalized development content, or create career pathways that support skill development.

5.4. Hybrid Human-AI Work Design

Effective AI adoption requires designing work processes that leverage the complementary strengths of humans and AI systems. For listed enterprises, hybrid work design represents a strategic priority for maximizing the value of AI investments while managing associated risks.

Research has documented that hybrid human-AI models often outperform either humans or AI alone [12]. In recruitment, for example, AI can efficiently screen large volumes of applicants, while human recruiters can focus on candidate relationships, cultural fit assessment, and final selection decisions. In performance management, AI can identify patterns in performance data, while managers can provide contextual interpretation, developmental feedback, and coaching. In learning and development, AI can recommend personalized learning paths, while human

mentors can provide guidance and support.

Designing effective hybrid work requires careful attention to the allocation of decision rights between humans and AI systems. Some decisions may be fully automated when they are routine and well-understood with low risk of adverse consequences. Others may be AI-assisted, with systems providing recommendations that humans can accept, modify, or override. Still others may remain human-only, where judgment, contextual understanding, and ethical considerations are paramount [12].

The design of human-AI interfaces also matters significantly. Systems should be designed to support human understanding and control, with clear explanations of AI recommendations, transparency about limitations, and opportunities for human input and override [2]. For listed enterprises, well-designed interfaces support both operational effectiveness and employee acceptance of AI systems.

6. Conclusion

6.1. Summary of Theoretical Contributions

This study has developed a comprehensive theoretical framework for understanding the impact of artificial intelligence on human resource management in listed enterprises and identifying appropriate countermeasures. Drawing upon the resource-based view, signaling theory, and sociotechnical systems theory, the framework identified three primary impact pathways: algorithmic decision-making in recruitment and selection, predictive analytics in performance management and retention, and automation in HR service delivery. The analysis explored organizational implications, including structural changes within HR functions, evolving workforce skill requirements, and transformations in employee-employer relationships. In response to these challenges, countermeasures were proposed across four domains: strategic alignment of AI with HR objectives, ethical governance frameworks, workforce reskilling initiatives, and hybrid human-AI work design.

The theoretical contributions of this study are threefold. First, the study provides a comprehensive analysis of AI impact pathways in the specific context of listed enterprises, addressing a significant gap in the existing literature that has largely focused on AI in HR generally without attention to the unique governance, disclosure, and stakeholder pressures faced by publicly traded companies. Second, the study applies established theoretical frameworks—RBV, signaling theory, and sociotechnical systems theory—to the emerging phenomenon of AI in HRM, demonstrating how these frameworks illuminate the opportunities

and risks of AI adoption and providing a foundation for future research. Third, the study develops countermeasures that integrate strategic, governance, capability, and design perspectives, providing guidance for organizational practice that addresses the multidimensional nature of AI adoption challenges.

6.2. Implications for Practice

The framework developed in this study has several implications for managerial practice. For executives in listed enterprises, the framework suggests that AI adoption in HR requires strategic alignment with business objectives and integration with broader organizational transformation efforts. AI should not be viewed as merely a technological implementation but as a fundamental change in how human capital is managed and how value is created. Executives should ensure that AI investments are guided by clear strategic priorities and supported by appropriate governance structures.

For HR leaders, the framework emphasizes the importance of developing new capabilities in data analytics, technology management, and ethical governance. The transformation of HR from administrative to strategic functions requires investment in HR professional development, changes in HR organizational structure, and the development of new roles and competencies. HR leaders should advocate for resources to support these capability developments and should position HR as a strategic partner in AI adoption.

For boards and investors, the framework highlights the need for oversight of AI governance practices. The risks associated with algorithmic bias, transparency failures, and employee trust erosion require board attention comparable to other significant organizational risks. Boards should ensure that appropriate governance frameworks are in place, that AI systems are regularly audited, and that management has developed capabilities to manage AI-related risks.

6.3. Limitations and Future Research Directions

This study has several limitations that suggest directions for future research. As a theoretical study, it does not provide empirical evidence for the proposed framework. Future research should test the framework empirically, examining how AI adoption affects HR outcomes and organizational performance in listed enterprises.

Several specific research directions emerge from this study. First, research should examine the relationship between AI adoption and HR outcomes in listed enterprises. Do organizations that adopt AI for HR achieve improvements in recruitment quality, employee retention, and HR efficiency? What organizational factors moderate these relationships?

Second, research should examine the governance practices that support responsible AI adoption. What governance structures are most effective in managing algorithmic bias and transparency concerns? How do governance practices affect employee trust and acceptance of AI systems? What role do boards play in overseeing AI governance?

Third, research should examine the transformation of HR roles and capabilities. How do HR functions change as AI systems are adopted? What competencies do HR professionals need to be effective in AI-enabled organizations? How do organizations develop these competencies?

Fourth, research should examine the impact of AI adoption on employee outcomes. How do employees experience AI in HR processes? What factors influence employee acceptance or resistance to algorithmic management? How does AI in HR affect employee engagement, well-being, and retention?

Fifth, cross-country research could examine how institutional context affects AI adoption in HR. Differences in regulatory environments, cultural values, labor market structures, and stakeholder expectations may influence both the adoption patterns and the outcomes of AI in HR. Such research would provide valuable insights for multinational listed enterprises.

References

[1] Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, 17(1), 99–120.

[2] Budhwar, P., Chowdhury, S., Wood, G., Aguinis, H., Bamber, G. J., Beltran, J. R., … & Varma, A. (2023). Human resource management in the age of generative artificial intelligence: Perspectives and research directions on ChatGPT. *Human Resource Management Journal*, 33(3), 606–659.

[3] Budhwar, P., Malik, A., De Silva, M. T., & Thevisuthan, P. (2022). Artificial intelligence in human resource management: A review and research agenda. *Journal of Business Research*, 143, 297–309.

[4] Campbell, B. A., Coff, R., & Kryscynski, D. (2012). Rethinking sustained competitive advantage from human capital. *Academy of Management Review*, 37(3), 376–395.

[5] Cheng, M. M., & Hackett, R. D. (2021). A critical review of algorithms in HRM: Definition, theory, and practice. *Human Resource Management Review*, 31(1), Article 100698.

[6] Chowdhury, S., Dey, P., Joel-Edgar, S., Bhattacharya, S., Rodriguez-Espindola, O., Abadie, A., & Truong, L. (2023). Unlocking the value of artificial intelligence in human resource management through AI capability framework. *Human Resource Management Review*, 33(1), Article 100899.

[7] Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of Management*, 37(1), 39–67.

[8] Dietvorst, B. J., Simmons, J. P., & Massey, C. (2018). Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. *Management Science*, 64(3), 1155–1170.

[9] Ioniță, A., & Ștefan, S. C. (2025). Strategic human resource management in the digital era:

Technology, transformation, and sustainable advantage. *Merits*, 5(4), Article 23.
[10] Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons*, 61(4), 577–586.
[11] Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. *Academy of Management Annals*, 14(1), 366–410.
[12] Kim, S., Khoreva, V., & Vaiman, V. (2025). Strategic human resource management in the era of algorithmic technologies: Key insights and future research agenda. *Human Resource Management*, 64(2), 447–464.
[13] Langer, M., König, C. J., & Papathanasiou, M. (2019). Highly automated job interviews: Acceptance under the influence of stakes. *International Journal of Selection and Assessment*, 27(3), 217–234.
[14] Marliyas, A., Ummah, M. A. C. S., & Gunapalan, S. (2026). The transformation of talent acquisition through artificial intelligence in the context of Industry 4.0: A systematic literature review. *Journal of Business Research and Innovation*, 11(2).
[15] Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, 83(2), 340–363.
[16] Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 87(3), 355–374.
[17] Trist, E. L., & Bamforth, K. W. (1951). Some social and psychological consequences of the longwall method of coal-getting. *Human Relations*, 4(1), 3–38.
[18] van Esch, P., & Black, J. S. (2021). Artificial intelligence (AI) in human resource management (HRM): A systematic literature review and future research agenda. *The International Journal of Human Resource Management*, 32(6), 1245–1278.
[19] Wright, P. M., Dunford, B. B., & Snell, S. A. (2001). Human resources and the resource-based view of the firm. *Journal of Management*, 27(6), 701–721.
[20] Zhao Shuming, Zhang Min, Zhao Yixuan. (2019). 人力资源管理百年：演变与发展 [A Century of Human Resource Management: Evolution and Development]. *Foreign Economics and Management*. 41(12), 50–73.[in Chinese]
[21] Zhao Yixuan, Zhao Shuming, Luan Jiarui. (2020). 基于人工智能的人力资源管理：理论模型与研究展望 [Artificial Intelligence-Based Human Resource Management: Theoretical Models and Research Prospects]. *Nanjing Social Sciences*. (2), 26-33.[in Chinese]Q

Impressum

Founders	Zhengjie Gao, Xinyu Song
Editor in Chief	Ao Feng, Chengdu University of Information Technology, China
Executive Editor	Zhengjie Gao, Geely University of China, China
Editorial Board	Jing Hu, Huazhong University of Science and Technology, China Xiaohu Du, Huazhong University of Science and Technology, China Xiangkui Li, Harbin University of Science and Technology, China Zuopeng Liu, Goettingen University, Germany Xinyu Song, Geely University of China, China
Young Editorial Board	Min Liao, Geely University of China, China Tao Zheng, Geely University of China, China Chong Li, Chongqing University, China Ruiqin Fan, Sehan University, Korea Ziyang Liu, Jiangsu Normal University, China Qiwei Liu, Urumqi Vocational University, China Minqiu Kuang, Hunan Agricultural University, China
Published By	**Hong Kong Dawn Clarity Press Limited** Rm 9042, 9/F, Block B Chung Mei Centre, 15-17 Hing Yip Street, Kwun Tong, Kowloon, Hong Kong e-mail: ijaaa@dawnclarity.press ***International Journal of Advanced AI Applications*** is published monthly.
Editorial Policy	***International Journal of Advanced AI Applications*** is directed to the international communities of scientific researchers in artificial intelligence, computers and electronic, from the universities, research units and industry.To differentiate from other similar journals, the editorial policy of IJAAA encourages the submission of original scientific papers that focus on the integration of the advanced AI applications. In particular, the following topics are expected to be addressed by authors: (1) Natural Language Processing (NLP): Conversational AI, machine translation, sentiment analysis, and context-aware dialogue systems. (2) Smart Cities and IoT Integration: AI for traffic optimization, energy management, waste reduction, and urban infrastructure. (3) Autonomous Systems and Robotics: Self-driving vehicles, drones, industrial automation, and human-robot collaboration.

	(4) Edge AI and Distributed Systems: Real-time processing, federated learning, and low-latency AI at the network edge. (5) Creative and Generative AI: Art, music, and content generation using generative adversarial networks (GANs) and transformers. (6) AI in Education and Industry: Adaptive learning platforms, intelligent tutoring systems, and AI-driven supply chain optimization. Ethical and Explainable AI (XAI): Fairness, transparency, and accountability in real-world AI deployment.