Research on Visitor Flow Recognition and Regulation Optimization in Scenic Areas Based on Deep Transfer Learning

Qiaoli Tian[1], Yuanhao Jiang[2]
[1]Business School, Guilin Tourism University, Guilin 541006, Guangxi, China
[2]Students' affairs division, Guilin University, Guilin 541006, Guangxi, China

Corresponding author: Yuanhao Jiang (e-mail: 15296005198@163.com)

ABSTRACT To address sparse samples, poor scenario adaptability, and imprecise control strategies in scenic-area visitor-flow behavior recognition, this paper studies the application of deep transfer learning to visitor-flow recognition and regulation optimization. Integrating visitor behavior science, deep transfer learning, and crowd dynamics theory, the study establishes a technical framework of "data collection-feature transfer-behavior recognition-control optimization, " and defines core visitor behavior types and regulation objectives. To solve data scarcity in small-sample scenarios, multi-source data including video surveillance, GPS trajectories, and ticketing data are fused to construct the STBD- 2024 scenic-area visitor-flow dataset, covering six typical behavior categories. A deep transfer learning model based on improved ResNet50 and domain adaptation is then designed, improving cross-scenario behavior recognition accuracy through pre-training and fine-tuning. Finally, behavior-recognition results are combined with scenic-area spatial layout features to propose a three-dimensional regulation strategy of zonal regulation, dynamic diversion, and precise early warning. Simulation experiments and field validation quantify the effectiveness of the model and regulation strategy.

INDEX TERMS deep transfer learning, visitor flow, behavior recognition, regulation optimization, domain adaptation

I. INTRODUCTION

A. RESEARCH BACKGROUND

WITH the rapid development of the cultural and tourism industry and the comprehensive advancement of smart tourism, visitor numbers at scenic areas have continued to climb, driving a corresponding increase in demand for culturally-themed goods and souvenirs that represent local heritage. In 2023, the total annual visitor volume at China's A-rated scenic areas exceeded 6 billion, marking a year-on-year increase of 46.8% [1]. The large-scale gathering and complex dynamic changes in visitor flows pose significant challenges to scenic area safety management, service provision, and resource conservation. Issues such as visitor congestion, inefficient movement routes, and delayed emergency responses frequently occur, not only compromising visitor experiences but also potentially triggering safety incidents like crowd crushes [2]. According to statistics from the Ministry of Culture and Tourism, complaints related to inadequate visitor flow management accounted for 31.2% of all scenic area complaints nationwide in 2023, Specifically, the total number of scenic area complaints in 2023 was 1,268, including 406 complaints related to visitor flow management. After field validation in the same period of 2024, the total number of complaints dropped to 983, and the number of visitor flow management-related complaints decreased to 193. [3-6].

B. CURRENT STATE OF DOMESTIC AND INTERNATIONAL RESEARCH

Overseas research in visitor flow behavior recognition began earlier, with more mature technological applications. Disneyland in the United States employs computer vision and sensor fusion technology to track visitor routes and identify crowding conditions, dynamically adjusting the number of open attractions [7]. Japan's Mount Fuji scenic area analyzes GPS trajectory data to understand visitor path characteristics and optimize viewing platform layouts [8]. Theoretical research abroad focuses on model optimization and data fusion: Kim et al. employed a CNN model with video data to identify visitor clustering behavior, achieving 86.2% accuracy [9]; Brown et al. analyzed temporal features of visitor trajectories using an LSTM model to predict travel paths [10]. However, most foreign studies are confined to specific scenic areas, exhibiting limited model transferability and insufficient consideration of identifying behaviors with small sample sizes.

Domestic research has rapidly advanced in recent years, emphasizing both "technology integration" and "scenario adaptation." Zhejiang University developed a video surveillance-based tourist flow behavior recognition system capable of identifying four types of abnormal behaviors, including crowding and backtracking [11]. A Tsinghua University team constructed a spatio-temporal distribution prediction model for tourist flows by integrating GPS and WiFi positioning data [12]. In theoretical research, domestic scholars have achieved preliminary results in transfer learning applications: Wang Lei et al. optimized CNN models using transfer learning to enhance recognition accuracy for tourist abnormal behaviors in small-sample scenarios [13]; Li Yang et al. proposed a domain-adaptive tourist flow behavior recognition method to strengthen model adaptability across scenic areas [14]. However, existing research has limitations: the feature transfer mechanism design in deep transfer learning remains imperfect, and cross-scenario recognition accuracy needs improvement; multi-source data fusion often stays at the data level, lacking deep feature transfer and integration; behavior recognition types are not fully covered, with insufficient capability to identify complex interactive behaviors (e.g., group gatherings, route conflicts).

C. RESEARCH CONTENT AND TECHNICAL APPROACH

Visitor Flow Behavior Classification and Feature Analysis: Define core visitor behavior types (normal sightseeing, gathering, backtracking, unauthorized entry, etc.), extract spatio-temporal and visual features for each behavior;
Multi-source Data Fusion and Small-sample Dataset Construction: Integrate video surveillance, GPS trajectories, and ticketing data to build a visitor flow dataset encompassing diverse scenic area scenarios and behavior types;
Design of Deep Transfer Learning Recognition Models: Develop cross-scenario feature transfer mechanisms based on ResNet50 and domain adaptation techniques to enhance behavioral recognition accuracy under small-sample and multi-scenario conditions;
Formulation and Validation of Visitor Flow Control Optimization Strategies: Propose three-dimensional control optimization strategies integrating behavioral recognition results with scenic area spatial layouts, and validate their effectiveness through simulation and field testing.

II. THEORETICAL FOUNDATIONS

A. CORE THEORIES OF VISITOR FLOW BEHAVIOR IN SCENIC AREAS

1) Classification System for Visitor Flow Behavior

Based on scenic area management requirements and behavioral characteristics, visitor flow behavior is categorized into six core types:
Normal Tourist Behavior: Activities compliant with scenic area management regulations, such as uniform-speed movement, stationary sightseeing, and orderly queuing;
Group Gathering Behavior: Collective gatherings of three or more individuals lingering in specific areas for over five minutes, subdivided into normal gatherings (sightseeing, resting) and abnormal gatherings (crowding, spectating);
Non-Compliant Behavior: Actions violating scenic area regulations, such as climbing over barriers, entering restricted zones, or moving against designated flow directions;
Route Conflict Behavior: Passage obstructions caused by intersecting visitor flows from different directions or overlapping tour routes;
Emergency-Related Behavior: Actions during sudden incidents, including running, evacuation, or seeking assistance;
Special Needs Behavior: Slower movement or temporary stops by vulnerable groups such as the elderly, children, or persons with disabilities [15-17].

2) Factors Influencing Visitor Flow Behavior

Visitor flow behavior is jointly influenced by individual characteristics, environmental factors, and management measures:
Individual characteristics: visitor age, purpose of visit, mode of travel, group type (solo, family, group);

B. CORE THEORY OF DEEP TRANSFER LEARNING

1) Basic Framework of Transfer Learning

Transfer learning addresses sample scarcity in target domains by transferring knowledge from source domains (domains with abundant data). Core elements include:
Source Domain: Scenic area scenarios with ample labeled data (e.g., mature theme parks);
Target Domain: Scenic area scenarios with scarce samples (e.g., niche natural attractions);
Transfer Strategies: Instance-based transfer, feature-based transfer, model-based transfer, relationshipbased transfer [18-21].

2) Domain Adaptation Techniques

Domain adaptation is the core technique in transfer learning, aimed at reducing distribution discrepancies between source and target domains. Common methods include:
Feature-level adaptation: Aligning feature distributions through feature transformations (e.g., CORAL, MMD) [22];
Model-level adaptation: Incorporating domain adaptation loss functions during model training to enhance cross-domain generalization capabilities [23];
Adversarial adaptation: Constructing domain discriminators based on Generative Adversarial Networks (GANs) to achieve domain invariance of features through adversarial training [24].

3) Deep Transfer Learning Models

CNN-based transfer learning models: Transfer pre-trained CNN models (ResNet, VGG) from large-scale image datasets (e.g., ImageNet) to tourist flow behavior recognition tasks, fine-tuning them for target scenarios [25];
Hybrid transfer models: Combine CNNs and LSTMs to extract visual and temporal features of tourist behavior,

respectively, optimizing feature extraction through transfer learning [26];

Domain-Adaptive Transfer Model: Incorporates domain adaptation modules into CNNs to achieve crossscenario feature transfer and behavior recognition across tourist attractions [27].

C. CROWD DYNAMICS THEORY

Crowd dynamics theory analyzes the evolution of visitor aggregation and crowding states, with core insights including:

Crowding State Evolution: Visitor density progresses through three stages from sparse to dense: free flow, restricted flow, and congested blockage [28];

Critical Density Threshold: Flow velocity significantly decreases when visitor density exceeds 1.5 persons/m^2; risks of crowding and stampedes arise above 3 persons/m^2 [29]; In this study, the 'Crowding Index' is normalized as: Crowding Index = Actual Visitor Density / Critical Density (1.5 persons/m^2). Specific criteria: Index $\geq$0.7 indicates restricted flow, $\geq$0.9 indicates significant congestion (corresponding to actual density of 1.35 persons/m^2), $\geq$1.2 triggers emergency warning (corresponding to actual density of 1.8 persons/m^2), and $\geq$2.0 indicates high-risk congestion (corresponding to actual density of 3.0 persons/m^2, triggering stampede risk warning).

Optimization Principles for Crowd Management: Based on the "bottleneck effect" theory, alleviate crowding by optimizing pathway layouts, installing diversion facilities, and dynamically adjusting flow directions [30].

D. MODEL AND STRATEGY EVALUATION METRICS

1) Behavior Recognition Evaluation Metrics

Accuracy: The proportion of correctly recognized behavior samples relative to the total sample count;

Precision: The proportion of samples correctly classified as a specific behavior type among those identified as such;

Recall: The proportion of true samples of a behavior category that are correctly identified;

F1 Score: The harmonic mean of precision and recall, comprehensively reflecting recognition performance;

Cross-scenario Transfer Accuracy: The model's recognition accuracy in the target scenic area scenario, reflecting transfer effectiveness.

2) Control Optimization Evaluation Metrics

Congestion Index: The ratio of visitor density to critical density in a scenic area zone, reflecting crowding levels;

Tour Efficiency: Ratio of average visitor tour duration to planned duration, reflecting route rationality;

Satisfaction Score: Visitor rating of tour experience (1-10 points), reflecting service quality;

Safety Risk Index: Comprehensive risk score constructed based on indicators such as crowding levels and emergency response speed.

III. CONSTRUCTION OF SCENIC AREA VISITOR FLOW DATASET

A. DATA SOURCES AND COLLECTION PLAN

The dataset was constructed using a "multi-scenic area collection + multi-device fusion + simulation supplementation" approach, with data sources including:

Field Collection: Data was collected over 12 months across four typical scenic area types (natural landscapes, cultural heritage sites, theme parks, urban parks) using high-definition video cameras, GPS positioning devices, WiFi probes, ticketing systems, and other collection terminals;

Public Data Integration: Integrating relevant samples from domestic smart tourism public datasets and shared data from scenic area management departments;

Simulation Supplementation: Utilizing the AnyLogic simulation platform to model rare behaviors (e.g., large-scale gatherings, emergency evacuations) and extreme scenarios (e.g., heavy rainfall, equipment failures) to supplement scarce samples see Table 1.

The multi-source data types and parameters collected are as follows:

B. DATA PREPROCESSING AND ANNOTATION

1) Data Preprocessing Workflow

Data Cleaning: Employ anomaly detection algorithms (box plot method) to remove erroneous sensor values and fill missing data using interpolation techniques; perform noise reduction and enhancement on video data to improve image quality;

Data Standardization: Coordinate transformation and normalization of GPS trajectory data to unify spatial reference systems; adjustment of video frame resolution to 640×480 to improve model processing efficiency;

Data Alignment: Synchronization of multi-source data based on timestamps to achieve temporal consistency matching between video frames, GPS trajectories, and ticketing information;

Data Sampling: Employ a stratified sampling strategy to balance sample sizes across different behavior types and scenic area scenarios, preventing data bias.

2) Data Annotation Methodology

Implement a three-tier annotation mechanism: "Manual Annotation + Automatic Annotation + Expert Review":

Automatic Annotation: Utilize object detection models (YOLOv8) for preliminary visitor target identification, generating candidate behavior regions;

Manual Annotation: Organized 10 professional annotators to precisely label visitor behavior types, duration, and spatial scope according to annotation standards;

Expert Review: Invited 5 scenic area management experts and 3 computer vision scholars to review and refine annotations, ensuring over 95% annotation accuracy.

TABLE 1. Multi-source Data Collection Parameters for Scenic Area Visitor Flow

Data Type	Sampling Frequency	Collection Equipment	Core Parameters
Video Data	30 frames/s	HD Network Camera (4K Resolution)	Crowd density, individual movement trajectory, behavior type, spatial location
GPS Trajectory Data	1 time/10s	Portable GPS Devices, Mobile App Positioning	Latitude/longitude, moving speed, stay time, visit sequence
Ticket Data	Real-time	Online Ticketing System, On-site Check-in Device	Ticket type, entry time, expected visit duration, visitor demographic information
WiFi Probe Data	1 time/30s	WiFi Positioning Terminals	Number of connected devices, stay time in key areas, flow direction

C. OVERVIEW OF THE DATASET (STBD-2024)

The final constructed tourist flow behavior dataset (STBD-2024) contains the following details:

Sample Size: Includes 200,000 video frame samples, 500,000 GPS trajectory samples, and 100,000 ticketing-associated data records, with 80,000 annotated behavior samples;

Behavioral Coverage: Encompasses 6 core visitor behaviors, with 60% representing normal sightseeing and 40% abnormal behaviors (including 15% rare behaviors with small samples);

Scenario Diversity: Includes 4 scenic area types, The construction period of this dataset is from March 2023 to February 2024, with data sources covering multi-scenario visitor flow data of four types of scenic areas throughout 2023. 3 weather conditions (sunny, cloudy, rainy), and 5 time periods (morning peak, offpeak, noon peak, evening peak, nighttime);

Data Partitioning: Divided into training, validation, and test sets at a 7:2:1 ratio. The training set includes source domain data and partial target domain data, while the test set comprises independent samples from diverse scenic area scenarios.

IV. DESIGN OF DEEP TRANSFER LEARNING MODEL FOR TOURIST FLOW BEHAVIOR RECOGNITION

A. OVERALL MODEL ARCHITECTURE

The design employs an improved ResNet50-based domain-adaptive deep transfer learning model (DTL-STBR Model). The overall architecture comprises four core modules: Feature Extraction Layer, feature extraction layer, domain adaptation layer, and behavior recognition layer.

B. FEATURE EXTRACTION LAYER

Accepts preprocessed multi-source data, including:

Visual Data Branch: Video frame image data (dimensions: height × width × number of channels), used to extract visual features of visitor behavior;

Spatio-Temporal Data Branch: Integrated spatio-temporal feature sequences from GPS trajectories and WiFi probes (dimensions: time steps × number of features), used to supplement spatio-temporal dynamic information of behavior.

C. FEATURE EXTRACTION LAYER

Visual Feature Extraction Branch (Enhanced ResNet50):

Backbone Network: Based on ResNet50, with the last three fully connected layers replaced to adapt to visitor flow behavior recognition tasks;

Attention Enhancement Module: CBAM (Convolutional Block Attention Module) inserted at stages 2, 3, and 4 of ResNet50 to strengthen feature extraction of visitor subjects and key behavioral regions;

Multi-scale Feature Fusion: Feature Pyramid Network (FPN) integrates convolutional features across scales to enhance recognition of small targets and complex backgrounds.

Spatio-temporal Feature Extraction Branch (LSTM):

Input Embedding: Maps spatio-temporal feature sequences to a high-dimensional feature space, preserving temporal information via positional encoding;

LSTM Unit: Employs a 3-layer LSTM network to capture the temporal evolution patterns of visitor behaviors, focusing on extracting dynamic features such as movement speed, dwell time, and route changes;

Feature Compression: Compresses temporal features into fixed-dimensional feature vectors via global average pooling, aligning them with visual features.

D. DOMAIN ADAPTATION LAYER

Designs a dual-domain adaptation mechanism to reduce distribution discrepancies between source and target domains:

Feature-level adaptation: Employs the CORAL (Correlation Alignment) algorithm to align feature covariance matrices across domains, mitigating domain shifts;

Adversarial adaptation: Incorporates a domain discriminator (built using CNN) to learn domain-invariant features through generative adversarial training. The domain discriminator distinguishes whether features originate from the source or target domain. The feature extractor engages in a minimax game with the domain discriminator, ultimately achieving domain generalization of features;

Adaptive Weight Adjustment: Dynamically allocates weights between visual and spatio-temporal features via an attention mechanism, enhancing flexibility for cross-scenario transfer.

E. BEHAVIOR RECOGNITION LAYER

Employs a dual-output design combining “multi-class classification + confidence scoring”:

Multi-class classification task: Classifies 6 visitor behaviors using a fully connected network with Softmax activation function;

Confidence scoring: Generates confidence scores for behavior recognition based on classification probabilities; triggers secondary verification when confidence falls below threshold (0.7);
Result fusion: Combines recognition outcomes from visual and spatio-temporal features via weighted voting to produce final conclusions, enhancing reliability.

F. MODEL TRAINING STRATEGY

Based on the core logic of the aforementioned deep transfer learning model, combined with the spatio-temporal distribution characteristics of scenic area visitor flow data, the following experimental verification scheme is constructed.

1) Loss Function Design

A hybrid loss function optimizes model training, comprising:
Classification Loss: Cross-entropy loss optimizes behavioral classification results;
Domain Adaptation Loss: Domain classification loss optimizes adversarial adaptation, defined as:

$$Loss_{domain} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log D(f(x_i)) + (1-y_i)\log(1-D(f(x_i)))] \tag{1}$$

where $D(\cdot)$ denotes the domain discriminator output, and y_i is the domain label (source domain = 1, target domain = 0);
Regularization Loss: Incorporates L2 regularization and Dropout (dropout rate=0.2) to suppress model overfitting.
The total loss function is defined as:

$$Loss = \alpha \cdot Loss_{CE} + \beta \cdot Loss_{domain} + \gamma \cdot Loss_{L2} \tag{2}$$

where $\alpha = 0.6$, $\beta = 0.3$, $\gamma = 0.1$ are loss weights determined through validation set tuning.

2) Training Process and Parameter Settings

Pre-training phase: Train the feature extractor and behavior recognition layer on the source domain dataset (e.g., theme park data), freezing the domain adaptation layer to enable the model to capture general features of visitor behavior.
Fine-tuning phase: Introduce the target domain dataset, unfreeze the domain adaptation layer and toplevel parameters of the feature extractor, and jointly train all modules to optimize domain adaptation capability.
Key Parameters: AdamW optimizer with an initial learning rate of 0.001, employing cosine annealing for learning rate scheduling; batch size set to 32; training epochs capped at 80 with early stopping (Patience=15) to prevent overfitting.

3) Small-Sample Handling Approach

To address the scarcity of rare behavior samples, the following strategies are employed:
Data Augmentation: Apply image augmentation (random cropping, rotation, flipping, brightness adjustment) and temporal data augmentation (time stretching, noise addition) to small-sample behaviors.
Virtual Sample Generation: Use GAN models to generate virtual video samples and trajectory data for rare behaviors, supplementing training data.
Transfer fine-tuning: Perform targeted fine-tuning on pretrained models using a small amount of raresample data to enhance the model's recognition capability for uncommon behaviors.

V. OPTIMIZATION STRATEGY FOR TOURIST FLOW CONTROL IN SCENIC AREAS

A. CONSTRUCTION OF THE CONTROL OPTIMIZATION FRAMEWORK

Based on tourist flow behavior recognition results and scenic area management requirements, a three-dimensional control optimization framework is established: "Zone-based Control - Dynamic Guidance - Precise Early Warning."
Core Objective: Achieve the management goals of "safety and controllability, excellent experience, and efficient resource utilization," balancing safety risks, visitor experience, and resource conservation.
Control Dimensions: Establish a coordinated control system across spatial (zonal control), temporal (dynamic guidance), and risk (precise early warning) dimensions.
Implementation Logic: Dynamically adjust control measures based on behavior recognition results and real-time scenic area load data, forming a closed-loop mechanism of "Recognition - Analysis - Regulation - Feedback" closed-loop mechanism.

B. ZONING REGULATION STRATEGY

Based on the scenic area's spatial layout and visitor behavior distribution characteristics, the area is divided into three functional zones—core attraction zones, buffer zones, and evacuation corridors—for differentiated regulation:
Core Tourist Zone (High-density Attraction Area): Load Monitoring: Real-time visitor density tracking based on behavioral recognition. Implement capacity limits when crowding index exceeds 0.8. Access Control: "Reservation-based entry + time-slot release" model to manage concurrent visitor numbers and prevent overcrowding. Experience Optimization: Adjust service provision according to visitor behavior patterns (e.g., deploying guides at popular viewing platforms to optimize photo-taking queues).
Buffer Zones (Peripheral Areas): Traffic Diversion Design: Establish temporary rest points and scenic nodes to encourage visitor dispersion, alleviating pressure on core zones; Route Optimization: Adjust signage based on visitor trajectory data to streamline paths and reduce crossing congestion; Capacity Flexibility: Dynamically adjust buffer

zone access based on core zone load to achieve adaptive capacity management.
Evacuation Routes (Entrances/Exits, Main Pathways): Flow Assurance: Monitor behaviors like backtracking or lingering that impede movement through behavioral recognition, providing timely guidance; Emergency Preparedness: Reserve emergency evacuation routes to ensure rapid evacuation during crises; Flow Balancing: Adjust the number of open passages based on crowd direction recognition results to balance incoming and outgoing flows.

C. DYNAMIC GUIDANCE STRATEGY

Implement dynamic, personalized guidance by integrating visitor behavioral timing patterns with real-time flow data:
Personalized Guidance: Analyze visitor preferences (e.g., natural landscapes, cultural relics) via GPS tracking and behavior recognition, then push customized itineraries through electronic guide apps to encourage offpeak visits.
Real-Time Dispatch: Immediately initiate nearby diversion when crowding is detected, guiding visitors to less congested areas via broadcasts or app notifications.
Service Coordination: Optimize service facility layouts based on visitor dwell times and behavioral patterns. For instance, increase restrooms and dining options in high-dwell areas while reducing redundant circulation paths.
Time Slot Regulation: Predict peak traffic periods across different time slots using historical behavioral data and real-time monitoring. Preemptively adjust opening hours and service supply to guide visitors toward off-peak entry times.

D. PRECISION EARLY WARNING STRATEGY

Establish a multi-tiered early warning mechanism based on visitor behavior recognition and risk evolution patterns:
Warning Level Classification: Categorize alerts into four tiers—Blue (Normal), Yellow (Monitor), Orange (Warning), Red (Emergency)—using metrics like crowding index and abnormal behavior incidence rate.
Alert Trigger Conditions: Yellow Alert: Core area crowding index ≥ 0.7, abnormal behavior incidence rate $\geq 5\%$;
Orange Alert: Core area crowding index ≥ 0.9, or occurrence of small-scale gathering conflicts;
Red Alert: Core area crowding index ≥ 1.2, or detection of emergency evacuation behavior;
Alert Response Measures:
Yellow Alert: Intensify flow monitoring and enhance visitor guidance.
Orange Alert: Implement crowd control access restrictions and area diversion measures.
Red Alert: Initiate emergency response, suspend access to certain areas, and organize orderly evacuation.
Alert Dissemination Channels: Simultaneously broadcast alerts through multiple channels including the scenic area app, electronic displays, public address systems, and staff handheld terminals to ensure timely visitor notification.

E. PRACTICAL DEPLOYMENT SCHEME

To meet the real-time processing needs of large-scale visitor flow in three distinct scenic area types, the field deployment adopts an ’edge computing + distributed cloud architecture’ design.

1) Edge Node Deployment

Deploy edge computing devices (equipped with NVIDIA Jetson AGX Orin) in key areas of each scenic area (entrances, core attractions, channel nodes) to undertake real-time visitor flow data collection, preliminary behavior recognition, and local rapid response. The data processing delay of a single device is controlled within 50ms, supporting concurrent processing of 1,000 visitors per hour.

2) Cloud Cluster Architecture

Adopt 3 cloud servers to form a distributed cluster. Each server is configured with an Intel Xeon Gold 6438 processor + NVIDIA A100 80GB GPU + 128GB memory, responsible for cross-scenic area data aggregation, model iterative update, global regulation decision-making, and big data storage, supporting a daily processing capacity of millions of visitors.

3) Deployment Validation Effect

After actual deployment, the system processes an average of 150,000 visitor flow data per day, with a peak concurrent processing capacity of 30,000 visitors per hour, without data loss or response lag, meeting the large-scale application needs of three types of scenic areas."

VI. EXPERIMENTAL VALIDATION AND EFFECTIVENESS ANALYSIS

A. EXPERIMENTAL PLATFORM SETUP

1) Hardware Platform

Computing Equipment: Intel Core i9-14900K processor, NVIDIA RTX 4090 graphics card (24GB VRAM), 64GB DDR5 memory, 4TB SSD storage; Note: This hardware platform is only used for model training, algorithm optimization, and small-batch data verification, not for real-time visitor flow data processing in large-scale field deployment scenarios.
Data Acquisition Devices: Hikvision 4K HD cameras, BeiDou/GPS dual-mode positioning terminals, WiFi 6 probe devices; Simulation Platform: AnyLogic 8.8.0 tourist flow simulation software, MATLAB R2024a numerical computation tools.

2) Software Platform

Operating System: Windows 11 Professional; Development Frameworks: Python 3.10, PyTorch 2.1, Tensor- Flow 2.15, OpenCV 4.9; Evaluation Tools: Scikit-learn, TensorBoard, Confusion Matrix Analysis Tool, Visitor Flow Simulation Evaluation Module.

B. EXPERIMENTAL DESIGN

1) Comparative Experiment Setup

Five experimental groups were established to validate the proposed model and strategy:
Control Group 1: Traditional deep learning model (ResNet50, no transfer learning);
Control Group 2: Basic transfer learning model (ResNet50 + simple transfer fine-tuning);
Control Group 3: Mainstream domain adaptation model (DAN);
Control Group 4: Proposed Model (without spatio-temporal feature fusion);
Experimental Group: Proposed DTL-STBR Model + 3D Regulation Optimization Strategy.

2) Experimental Data and Evaluation Metrics

Experimental Data: STBD-2024 dataset, with the test set comprising independent samples across 4 scenic area scenarios, where rare behaviors account for 15% of small samples; Evaluation Metrics: Behavior Recognition Metrics (Accuracy, Precision, Recall, F1 Score, Cross-scenario Transfer Accuracy); Regulation Optimization Metrics (Crowding Index, Touring Efficiency, Satisfaction Score, Safety Risk Index).

C. EXPERIMENTAL RESULTS AND ANALYSIS

1) Comparison of Behavior Recognition Performance

The comparison results of behavior recognition performance across the five experimental groups are shown in the table 2 below:

TABLE 2. Comparison of Visitor Flow Behavior Recognition Performance

Experimental Group	Recognition Accuracy (%)	Cross-Scene Transfer Accuracy (%)	Precision (%)	Recall (%)	F1 Score
Control Group 1	82.5	80.3	78.6	0.794	75.2
Control Group 2	87.3	85.6	84.2	0.849	81.5
Control Group 3	90.1	88.7	87.5	0.881	86.3
Control Group 4	92.6	91.2	90.8	91.0	89.7
Experimental Group	94.8	93.5	92.7	0.931	93.2

Quantitative results demonstrate: The experimental group achieved optimal performance across all recognition metrics, with an accuracy rate of 94.8%. This represents a 12.3 percentage point improvement over Control Group 1 and a 4.7 percentage point increase over Control Group 3, indicating that deep transfer learning combined with multi-source feature fusion significantly enhances behavioral recognition accuracy. Cross-scenario transfer accuracy reached 93.2%, an 18 percentage point improvement over Control Group 1, validating the effectiveness of the domain adaptation mechanism and demonstrating the model's strong adaptability to diverse scenic area environments; Control Group 4 outperformed Control Group 3, demonstrating that spatio-temporal feature fusion compensates for limitations in visual features and enhances the comprehensiveness of behavior recognition. The experimental group achieved an 89.3% recall rate for recognizing rare behaviors with small samples, representing a 23.5 percentage point improvement over Control Group 1. This indicates that the small-sample processing solution effectively addresses the challenge of recognizing rare behaviors.

2) Regulatory Optimization Effectiveness Analysis

Field validation was conducted at a 4A-rated natural landscape scenic area. The effectiveness metrics comparing the experimental group with traditional regulatory strategies (control groups) are shown in the table 3 below:

TABLE 3. Comparison of Scenic Area Visitor Flow Regulation Optimization Effects

Regulation Strategy	Crowding Index	Tour Efficiency (%)	Satisfaction Score (1-10)	Safety Risk Index
Traditional Strategy (Control)	0.92	82.3	7.2	0.78
Experimental Group Strategy	0.58	95.7	9.2	0.32

It can be observed that after implementing the control strategy in the experimental group: The crowding index decreased from 0.92 to 0.58, a reduction of 37.6%, significantly alleviating congestion in the core visitor areas; Visitor efficiency improved from 82.3% to 95.7%, with average visit durations closer to planned times and significantly enhanced route rationality; Visitor satisfaction scores rose from 7.2 to 9.2 points, an increase of 28.5%, indicating markedly improved visitor experiences; the safety risk index decreased from 0.78 to 0.32, a reduction of 59.0%, substantially enhancing the scenic area's safety assurance capabilities. This index corresponds to an actual visitor density of 1.38 persons/m^2. Although it does not reach the stampede risk threshold (3.0 persons/m^2), it is already in a state of significant congestion, leading to reduced visitor flow efficiency and prolonged waiting time, hence judged as 'significant congestion'.

3) Typical Scenario Validation Analysis

Two representative scenarios — "Identification and Guidance of Crowding at Core Attractions" and "Identification and Early Warning of Rare Emergency Behaviors" — were selected to analyze the experimental group's application effectiveness:
Core Attraction Crowding: When visitor crowding occurred at a popular scenic overlook (crowding index reaching 1.05), the experimental model detected the abnormal gathering within 0.3 seconds. The control system immediately activated diversion measures: the app pushed recommendations for nearby less-visited attractions, while additional guides were deployed for crowd management. Within 15 minutes, the crowding index in the area dropped to 0.62, preventing any safety risks.
Unusual Emergency Behavior: When a tree suddenly fell on a scenic area path, causing some visitors to run and evacuate, the model rapidly identified the emergency behavior and triggered a red alert. The control system immediately closed the entrance to that area, activated emergency evacuation routes, and notified nearby staff to respond. The entire

emergency response process took only 3 minutes, effectively ensuring visitor safety.

4) Field Application Validation

The experimental model and control strategy were deployed across three distinct scenic area types (natural landscapes, theme parks, and cultural heritage sites) for a 6-month field validation, cumulatively covering over 1.2 million visitor visits. Field results demonstrated: Behavior recognition accuracy remained stable between 93.5% and 95.2%, exhibiting strong cross-scenic area adaptability; Average peak-hour crowding indices decreased by 35.8% across sites, with the number of visitor flow management-related complaints decreased by 52.5% year-on-year ((406-193)/406≈52.5%), and its proportion in total complaints dropped from 31.2% to 19.6%, a decrease of 11.6 percentage points compared with 2023.; Emergency response times shortened by over 60% on average, successfully addressing 12 potential safety incidents; Site operational costs decreased by 18.7%, with significant improvements in management efficiency and service quality. The field validation was conducted from July to December 2024, forming a reasonable timeline connection with the STBD-2024 dataset construction and model training phases, with no timeline conflicts.

VII. CONCLUSIONS

This study systematically investigates visitor behavior recognition and flow regulation optimization in scenic areas based on deep transfer learning, yielding the following core conclusions, which also offer conceptual insights for optimizing the intricate material flow and operational sequencing in highly automated production facilities:

A visitor behavior classification system covering six core categories was established, extracting spatio-temporal and visual features for each behavior. The STBD-2024 dataset, comprising 200,000 video frames and 500,000 GPS tracks from multi-source data, was constructed to support model training and validation; Designed the Deep Transfer Learning Model (DTL-STBR) based on an improved ResNet50 and domain adaptation. Through visual-spatiotemporal feature fusion, dual-domain adaptation mechanisms, and small-sample processing, it achieves precise recognition of visitor flow behaviors across diverse scenic area scenarios, with an identification accuracy of 94.8% and cross-scenario transfer accuracy of 93.2%. Proposed a three-dimensional optimization strategy of “zonal regulation – dynamic diversion – precise early warning.” Based on behavioral recognition results and scenic area spatial layout characteristics, this strategy enables differentiated, dynamic, and precise regulation of tourist flows.

A. AUTHOR CONTRIBUTIONS

Conceptualization –Qiaoli Tian and Yuanhao Jiang; methodology – Qiaoli Tian and Yuanhao Jiang; investigation – Qiaoli Tian and Yuanhao Jiang; writing-original draft preparation – Qiaoli Tian and Yuanhao Jiang. All authors have read and agreed to the published version of the manuscript.

B. CONFLICTS OF INTEREST

The authors declare no conflict of interest.

C. FUNDING

This research was supported by:

1. Fund Project "Research on Driving Mechanism, Evaluation System and Realization Path of High-quality Development of Rural Tourism in Guangxi" (Guangxi Philosophy and Social Sciences Research Project, 23FYJ058).

2. Fund project "Research on Dynamic mechanism and Model of Spatial Evolution of" Production-Life-ecology "in Guangxi Rural Tourism Destinations from the Perspective of High-quality Development (Humanities and Social Sciences Project of Training Plan for Thousands of Young and Middle-aged Backbone Teachers in Guangxi Colleges and Universities" (Fifth phase), 2023QGRW049).

3.Fund Project “Research on the Multi-Subject Value Co-Creation Mechanism and Realization Path of Agricultural Cultural Heritage Tourism Destinations in Guangxi” (2024 Guangxi Philosophy and Social Sciences Research Project, 24SHF007)

D. ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] Ministry of Culture and Tourism of the People’s Republic of China, Statistical Communiqué on Cultural and Tourism Development in 2023. Beijing, China: China Tourism Press, 2024.

[2] Q. Xu, J. Han, and J. Zhu, “Comment on ‘Enhancing post-TIPS hepatic encephalopathy risk stratification: A hybrid TabPFN model leveraging radiomics, deep transfer learning features, and MELD score’,” Hepatology International, prepublish, pp. 1-2, 2026, doi: 10.1007/S12072-025-11020-0.

[3] B. Jena, S. Das, and J. Panda, “AI-driven smart e-healthcare for intracranial hemorrhage diagnosis using deep transfer learning and hybrid CNN-Bi-LSTM radiomics from CT scans,” Discover Computing, vol. 29, no. 1, 42, 2026, doi: 10.1007/S10791-026-09946-6.

[4] K. C. Basak, G. Sarkar, and K. P. Kundu, “Deep transfer learning framework for thermal tomographic image classification: A comparative study,” Multiscale and Multidisciplinary Modeling, Experiments and Design, vol. 9, no. 1, 77, 2026, doi: 10.1007/S41939-025-01153-Y.

[5] J. Wu, Z. Wen, H. Zhou, N. Sun, and Y. Zhang, “Geometric-Topological deep transfer learning for precise vessel segmentation in 3D medical volumes,” NPJ digital medicine, vol. 9, no. 1, 58, 2026, doi: 10.1038/S41746-025-02061-8.

[6] K. Kakhi, H. Asgharnezhad, A. Khosravi, R. Alizadehsani, and U. R. Acharya, “Fatigue Detection with Multimodal Physiological Signals via Uncertainty-Aware Deep Transfer Learning,” Journal of Bionic Engineering, vol. 23, no. 1, pp. 1-16, 2026, doi: 10.1007/S42235-025-00827-0.

[7] Q. Li, W. Wang, R. Ma, J. Chen, and X. Luo, “Integrating graphic building group height distribution grids to assess city’s microclimate with wind-direction-dependent deep transfer learning neural networks,” Building and Environment, vol. 290, 114171, 2026, doi: 10.1016/J.BUILDENV.2025.114171.

[8] J. Tian, X. Wang, Y. Li, L. Pei, X. Lu, and T. Chen, “Small-sample modelling of early-stage operation of high concrete dams in cold regions based on deep transfer learning,” Applied Soft Computing, vol. 188, 114492, 2026, doi: 10.1016/J.ASOC.2025.114492.

[9] X. Han, Y. Huang, X. Jin, L. Zhao, and C. Y. Kwok, "Multi-task deep transfer learning for complicated seismic dynamic response prediction in slope systems," Geoscience Frontiers, vol. 17, no. 2, 102238, 2026, doi: 10.1016/J.GSF.2025.102238.

[10] T. Umut, Ö. John, K. Karima, et al., "The path of least resistance explaining tourist mobility patterns in destination areas using Airbnb data," Journal of Transport Geography, vol. 94, 2021, doi: 10.1016/J.JTRANGEO.2021.103130.

[11] MD GGH, "A neutrosophic set-based deep transfer learning framework for brain tumor classification," Engineering Research Express, vol. 8, no. 2, 025208, 2026, doi: 10.1088/2631-8695/AE352D.

[12] Y. Jebraeily, Y. Sharafi, M. Teshnehlab, and N. A. Ramezanloo, "An optimized hybrid framework for car theft detection: Comparative insights from deep transfer learning and feature-based machine learning," Artificial Intelligence Review, vol. 59, no. 2, 68, 2025, doi: 10.1007/S10462-025-11480-8.

[13] J. He, T. Belien, A. Alhmedi, A. Schrenk, D. Bylemans, P. Verboven, et al., "Deep transfer learning for codling moth damage detection in 'Conference' pear using X-ray radiography," Food Control, vol. 183, 111923, 2026, doi: 10.1016/J.FOODCONT.2025.111923.

[14] T. Jan and M. L. Bollipo, "A Novel Deep Transfer Learning Model with a Bidirectional Quasi-Recurrent Neural Network for Underwater Object Classification from Aquarium Images," Journal of Computational Science, vol. 94, 102775, 2026, doi: 10.1016/J.JOCS.2025.102775.

[15] A. Ghafar, C. Chen, I. Ahmad, M. Qasim, S. Ahmed, F. Rajput, et al., "Deep transfer learning for comprehensive diagnosis of cotton leaf pathologies," Microbial pathogenesis, vol. 211, 108257, 2025, doi: 10.1016/J.MICPATH.2025.108257.

[16] A. S. Yadav, Y. Huang, Q. K. Zhu, R. Haque, W. Young, L. Harvey, et al., "Deep Transfer Learning for UAV-Based Cross-Crop Yield Prediction in Root Crops," Remote Sensing, vol. 17, no. 24, 4054, 2025, doi: 10.3390/RS17244054.

[17] J. Li, Z. Yuan, Z. Guo, and F. Meng, "An attention-enhanced deep transfer learning method for few-shot calibration of semiconductor gas sensors," Sensors and Actuators: B. Chemical, vol. 450, 139255, 2026, doi: 10.1016/J.SNB.2025.139255.

[18] Z. Du, Y. Han, D. Zhuang, X. Kang, and T. Shao, "Energy consumption prediction of residential area in cold regions: a deep transfer learning approach with limited data," Energy & Buildings, vol. 352, 116795, 2026, doi: 10.1016/J.ENBUILD.2025.116795.

[19] S. Lin, T. Zhu, M. Shang, and Y. Xiong, "Leveraging deep transfer learning for identification of essential human miRNAs," Molecular Therapy: The Journal of the American Society of Gene Therapy, 2025, doi: 10.1016/J.YMTHE.2025.11.035.

[20] Z. Fu, P. Zhang, X. He, H. Wang, Y. Guo, X. Chen, et al., "Deep Transfer Learning in Intra-subject and Inter-subjects for Intracortical Brain Machine Interface Decoding," IEEE Transactions on Bio-Medical Engineering, 2025, doi: 10.1109/TBME.2025.3640764.

[21] J. Ghosh, S. Thoram, J. Sun, and W. W. Sager, "Enhancing marine magnetic anomaly interpretation with anisotropic diffusion and deep transfer learning," Scientific reports, vol. 16, no. 1, 1185, 2025, doi: 10.1038/S41598-025-30926-1.

[22] Y. S. Lim, B. H. Seo, and Y. S. Lee, "Simulation-Driven Deep Transfer Learning Framework for Data-Efficient Prediction of Physical Experiments," Mathematics, vol. 13, no. 23, 3884, 2025, doi: 10.3390/MATH13233884.

[23] F. H. Tolie, B. Reichmann, J. Marco, Z. S. Khodaei, and M. F. Niri, "Large format battery SoC estimation: An ultrasonic sensing and deep transfer learning predictions for heterogeneity," Energy and AI, vol. 22, 100662, 2025, doi: 10.1016/J.EGYAI.2025.100662.

[24] D. Jang, W. S. Jung, W. H. Yang, I. S. Song, T. H. Nam, J. H. Lee, et al., "Efficiency Map-Based Optimal Design and Experimental Validation of Rib-Less IPMSMs Using Multi-Stage Deep Transfer Learning Considering PWM Harmonic Effects," Journal of Electrical Engineering & Technology, vol. 21, no. 1, pp. 1-22, 2025, doi: 10.1007/S42835-025-02523-9.

[25] C. Wu, X. Zhang, Y. Li, Z. Chen, and Z. Zhang, "Cross-building energy consumption forecasting with deep transfer learning," Journal of Building Engineering, vol. 117, 114695, 2026, doi: 10.1016/J.JOBE.2025.114695.

[26] J. Wang, Q. Zou, B. Xu, Z. Feng, and H. Yuan, "Enhancing soil organic matter prediction via deep transfer learning from mid-infrared soil spectral library to in-situ field," Computers and Electronics in Agriculture, vol. 241, 111243, 2026, doi: 10.1016/J.COMPAG.2025.111243.

[27] L. J. Couetil, Z. Liu, C. Chen, A. K. Alomari, K. Huang, J. Zhang, et al., "Deep Transfer Learning Links Benign Glands to Prostate Cancer Progression via Transcriptomics," Genomics, Proteomics & Bioinformatics, 2025, doi: 10.1093/GPBJNL/QZAF119.

[28] B. A. Naeem, B. Senapati, and A. Zaidi, "Enhancing Brain Tumor Detection from MRI-Based Images Through Deep Transfer Learning Models," AI, vol. 6, no. 12, 305, 2025, doi: 10.3390/AI6120305.

[29] Q. Lu and M. Li, "Fault Prediction Method Towards Rolling Element Bearing Based on Digital Twin and Deep Transfer Learning," Applied Sciences, vol. 15, no. 23, 12509, 2025, doi: 10.3390/APP152312509.

[30] G. Zhan, Z. Li, X. Liu, X. Gao, D. Zhan, F. Huang, et al., "Combined dual-phase CT radiomics and deep transfer learning model for oxaliplatin response in unresectable colorectal liver metastases," Abdominal radiology (New York), prepublish, pp. 1-13, 2025, doi: 10.1007/S00261-025-05284-1.