International Journal of Geo-Information

MDPI

Article

Trajectory-Driven Road Network Extraction via Coupled Multi-Level Grid Semantics

Yunfei Zhang [1], Hongjie Zhu [1], Baifa Wu [2], Naisi Sun [1], Cuifeng Zhang [3], Tianyu Zhong [1] and Chaoyang Shi [4,5,*]

[1] School of Aeronautical Engineering, Changsha University of Science & Technology, Changsha 410114, China; zhang.yunfei@csust.edu.cn (Y.Z.); hongjie@stu.csust.edu.cn (H.Z.); 23101020015@csust.edu.cn (N.S.); zhongtianyu@stu.csust.edu.cn (T.Z.)
[2] The First Surveying and Mapping Institute of Hunan Province, Changsha 410114, China; wubaifa@xcyy.net.cn
[3] Changsha Investigation, Survey & Design Research Institute, Changsha 410007, China; zhangcuifeng111@foxmail.com
[4] School of Civil and Hydraulic Engineering, Huazhong University of Science and Technology, Wuhan 430074, China
[5] Engineering Laboratory of Spatial Information Technology of Highway Geological Disaster Early Warning in Hunan Province, Changsha University of Science & Technology, Changsha 410114, China
* Correspondence: cyshi@hust.edu.cn

Abstract

Road network extraction and updating are crucial for urban development, map updating, and mobility applications. Existing trajectory-based methods often underutilize grid-level semantic information and neighborhood context, thereby limiting their robustness to noisy, heterogeneous, and cross-city trajectory conditions. This study proposes a supervised framework for trajectory-driven road network extraction by coupling intra-grid movement semantics with inter-grid neighborhood context. Multi-level features, including convex-hull shape descriptors, directional clustering, DTW-based (Dynamic Time Warping) heterogeneity, and neighborhood density differences, are used to train a Random Forest classifier for key-grid detection. The detected key grids are further processed through morphology-aware thinning and Kalman smoothing to generate a topology-preserving and vectorization-ready road skeleton. The model is trained on pedestrian trajectories from Shenzhen and directly transferred to vehicle trajectories in Wuhan and Changsha under a zero-shot setting. Experimental results show that the proposed method achieves longer correctly extracted road length and competitive length-based precision compared with raster-based reference methods, while feature-importance and ablation analyses confirm the complementary role of neighborhood context. The proposed pipeline is scalable, interpretable, and transferable, supporting trajectory-based road map updating and urban network analysis.

Keywords: road network extraction; trajectory data; grid; morphology; random forest

Check for updates

Academic Editors: Wolfgang Kainz and Hartwig H. Hochmair

Received: 25 March 2026
Revised: 2 June 2026
Accepted: 4 June 2026
Published: 7 June 2026

Copyright: © 2026 by the authors. Published by MDPI on behalf of the International Society for Photogrammetry and Remote Sensing. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

The accelerated pace of urbanization and the continuous growth of population have rendered road network data increasingly critical for the construction of integrated transportation systems and the enhancement of infrastructure connectivity in metropolitan regions [1]. As fundamental components of geographic information systems (GIS) and intelligent transportation systems (ITS), road networks contribute significantly to improving the quality of urban life and facilitating urban–rural integration [2,3]. Moreover, accurate, high-quality road networks serve as a foundational reference for urban spatial analysis,

transportation planning, and location-based services [4,5]. Accordingly, the automatic construction and timely updating of road networks have become central tasks in transportation geographic information science. Nevertheless, trajectory-driven road map updating remains challenging due to noisy and irregular sampling, heterogeneous data modalities (e.g., pedestrian vs. vehicle), and the need to preserve network topology at the city scale.

With the widespread availability of GPS-enabled mobile devices, crowdsourced trajectory data—such as pedestrian or vehicle traces—have emerged as a rich source of spatial–temporal information. Compared to static surveying data and remote-sensing imagery, these trajectories exhibit denser spatial coverage, richer semantic context, and fine-grained temporal characteristics [6,7]. Consequently, an increasing number of studies have sought to extract road networks directly from such data by leveraging their inherent mobility patterns and structural features. Recent advancements have expanded the scope from one-shot geometry extraction to spatial–semantic road map construction using trajectories and geo-tagged content [8], and to topology-aware representations that explicitly embed trajectories in road-network space [9]. Turning raw trajectories into topology-ready networks, however, requires models that capture both localized movement organization and broader spatial context while remaining robust across cities and modalities. In this study, we focus on generating a topology-preserving, vectorization-ready road skeleton from trajectories, rather than attribute enrichment or trajectory–network alignment.

Various methods have been proposed for road network extraction using GNSS trajectory data. For example, Cao and Sun [10] and Yang et al. [11] utilized Delaunay triangulation to derive road centerlines and boundaries, demonstrating strong adaptability to varying trajectory densities [12]. Zhou et al. [13] constructed pedestrian density maps from walking traces and applied discrete Morse theory [14] to extract "ridge lines" as indicative of pedestrian pathways. Guo et al. [15] enhanced the compactness of the extracted network by adjusting trajectory distributions. Yang et al. [16] further proposed a multi-scale fractal analysis combined with connected-component filtering to differentiate between random and goal-directed walking behaviors. Further, Yang et al. [17] developed the human-flow probability field (HFPF) approach and combined it with hydrological modeling to effectively delineate both primary and secondary pedestrian paths. Complementing these density- and field-based strategies, hybrid incremental pipelines that fuse spatial density with temporal continuity have also been proposed to progressively update road networks from spatio-temporal trajectories [18]. In recent studies, Tang et al. [19,20] proposed an outdoor hiking road network construction method that uses trajectory density stratification and kernel density estimation to generate a 2D road network, fuses elevation data for 3D extension, and employs rasterized density maps, direction-constrained density clustering, and adaptive thresholding to extract pedestrian areas and intersections, building a complete 3D outdoor pedestrian road network framework. Meanwhile, Dal Poz and Morceli [21] focused on smartphone-based GPS trajectories, using Kernel Density Estimation, morphological skeletonization, and Voronoi diagrams to extract roads in mixed urban–rural environments. These density- and field-based approaches can reveal corridor-like structures from global statistics, but they may under-utilize neighborhood interactions and often require additional regularization to ensure topological consistency. Specifically, the classification of a grid cell should be constrained by its neighbors; for instance, a cell with moderate trajectory density is more likely to be part of a road if its adjacent cells exhibit consistent heading directions, forming a coherent linear flow rather than isolated noise.

In addition to global density-based methods, clustering techniques have also been widely used. For example, fuzzy C-means clustering with velocity constraints was used to distinguish adjacent road segments [22], while B-spline curves were applied to smooth road geometries, albeit with high computational costs and sensitivity to local anomalies [23].

Kasemsuppakorn and Karimi [24] extracted azimuth and speed-based key points for simplified road generation via Partitioning Around Medoids (PAM), a method also adopted by Wu et al. [25]. Automatic road network construction from massive GPS trajectory data has also been demonstrated using scalable clustering/inference pipelines [26]. Furthermore, Stanojević et al. [27] proposed a two-phase method that clusters trajectory points for intersection detection and groups segments based on directional and speed similarities. Xie et al. [28] introduced a density-based clustering method that integrates similarity metrics across position, speed, direction, and angle for pedestrian network inference. Recently, Buchin et al. [29] automatically constructed high-precision road network maps from chaotic GPS trajectory data by clustering and optimizing paths using multi-width trajectory clustering. Such clustering approaches have proven effective even for modality-specific analyses, for example, mining spatial patterns and road-type preferences from crowdsourced cycling traces [30]. Clustering- and similarity-based methods capture local movement organization but can be sensitive to noise and parameterization, and typically need explicit mechanisms to enforce connectivity and suppress spurious branches.

Intersection-aware approaches have also drawn significant attention. Karagiorgou and Pfoser [31] identified intersections via abrupt changes in speed and direction and then connected them into a coherent network. Direction-ratio statistics have been exploited explicitly for intersection detection from trajectories [32], and Yuan et al. [33] advanced this line of work by detecting intersections and lane geometry changes, using principal curves for lane centerlines and Gaussian mixture models for topology inference. Shen et al. [34] incorporated cycling behavior to locate intersections and applied shape-aware curve fitting for turn path reconstruction. Zhang et al. [5] proposed a virtual representative point and CFDP clustering framework combined with Delaunay triangulation to improve intersection extraction. Lyu et al. [35] developed the Motion-Aware Map Construction (MAMC) approach by clustering turning points and trajectory segments, while Wang et al. [36] decomposed large interchanges into smaller intersections to simplify road network generation. Furthermore, Jiao et al. [37] employed forward and backward trajectory tracking mechanisms to identify divergence and convergence points in interchanges, effectively addressing the challenge of detecting false intersections in multi-layer complex road networks. While these intersection- and lane-level models improve junction fidelity, their cross-modal and cross-city generalization remains difficult without hierarchical semantics and neighborhood-aware context.

Despite the diversity and technical sophistication of these methods, three limitations persist in current trajectory-based road network extraction techniques. First, many studies emphasize point-level semantics or global density fields, while the hierarchical coupling of intra-cell semantics and inter-cell context remains under-explored. Second, generalization across modalities and cities is often limited when sampling density, noise, and movement behaviors change substantially. Third, topology artifacts (e.g., block-like clusters and spurious short branches) can degrade the connectivity and vectorization-readiness of extracted networks, and reproducible end-to-end evaluation protocols are not always clearly documented.

To address these challenges, this study proposes a novel framework for trajectory-based road network extraction that couples multi-level grid features with supervised learning. Specifically, we transform trajectory data into a structured grid space, enabling the construction of both intra-grid and inter-grid semantic indicators, such as convex-hull density, direction clustering, and neighborhood density differences. A Random Forest classifier is trained to identify key grids that most likely correspond to road segments. Finally, an improved morphological thinning algorithm is applied to extract a topologically coherent single-pixel road network, followed by structural refinement to eliminate noise and

discontinuities. In this way, the pipeline combines hierarchical grid semantics, supervised key-grid detection, and topology-aware refinement to provide a scalable and interpretable route from raw trajectories to vectorization-ready networks.

The main contributions of this work are as follows:

(1) We design a hierarchical grid-based representation that couples intra-grid movement structure (dispersion, directionality, and segment heterogeneity) with inter-grid neighborhood continuity (density gradients), enabling trajectory semantics to be modeled beyond point-level features and purely global density fields.

(2) We formulate candidate road-region discovery as a key-grid binary classification problem and train a Random Forest model on pedestrian trajectories, which is then directly transferred to vehicle trajectories in other cities to assess cross-modal and cross-city generalization under heterogeneous sampling and urban layouts.

(3) We develop a topology-oriented reconstruction pipeline that combines morphological closing, four-neighborhood thinning with artifact correction, and Kalman smoothing to produce a single-pixel, vectorization-ready road skeleton, reducing grid-induced staircasing while preserving junction structures for downstream GIS network analysis.

(4) We provide a reproducible end-to-end evaluation setup based on buffer matching against OSM centerlines and report both key-grid detection performance (Precision/Recall/F1) and network extraction quality (correctly extracted length and length-based precision) across multiple zones, facilitating fair comparison and replication.

The remainder of this paper is organized as follows. Section 2 describes the proposed framework, including trajectory preprocessing, multi-level grid feature construction, supervised key-grid detection, and morphology-based reconstruction with smoothing. Section 3 reports experimental settings, evaluation protocols, and results on within-city testing and cross-city/cross-modal transfer. Section 4 discusses the implications, limitations, and potential extensions. Section 5 concludes the paper.

2. Methods

This section presents a supervised-learning framework for trajectory-based road network extraction. The overall pipeline is summarized in Algorithm 1, which consists of four main stages: (i) trajectory preprocessing and grid indexing (Steps 1–2), (ii) coupled multi-level grid feature construction (Steps 3–5), (iii) key-grid detection via Random Forest classification (Steps 6), and (iv) morphology-aware reconstruction for a topology-preserving road skeleton (Steps 7–10). The detailed implementation of each stage is elaborated in the following subsections.

2.1. Trajectory Data Preprocessing and Grid Indexing

Pedestrian trajectories, recorded by GPS-enabled devices, contain timestamped geographic coordinates that reflect human movement paths. However, GPS signals are often affected by multipath effects and signal delays in dense urban or indoor environments. Therefore, it is essential to preprocess raw trajectory data to ensure positional consistency and reduce noise.

Each GPS trajectory segment is defined as:

$$Tr_j = \left\{ P_{j,i} = \left(x_{j,i}, y_{j,i}, t_{j,i} \right) \middle| i = 1, 2, \ldots, I_j \right\}, j = 1, 2, \ldots, J \quad (1)$$

where $P_{j,i}$ denotes the ith point of the jth trajectory segment. $x_{j,i}$, $y_{j,i}$ and $t_{j,i}$ denote longitude, latitude, and timestamp, respectively. I_j is the number of points in segment j, and J is the total number of trajectory segments.

Algorithm 1: Coupled multi-level grid semantics for trajectory-driven road network extraction

Input:
Raw trajectory points $P = \{P_i = (x_i, y_i, t_i) | i = 1, 2, \ldots, n\}$
Spatial extent: $[X_{min},\ X_{max},\ Y_{min},\ Y_{max}]$
Grid size: 200×200
DBSCAN parameters: anglular threshold $\theta = 30°$, $MinPts = 5$
Label file containing grid-type annotations {sequence index, type}

Output:
Refined road network skeleton in raster and vector formats

Step 1: Preprocess raw trajectories and construct trajectory segments
  Remove duplicate points and filter out outliers;
  Construct trajectory segments $Tr_j = \{P_{j,i} = (x_{j,i}, y_{j,i}, t_{j,i}) | i = 1, 2, \ldots, I_j\}, j = 1, 2, \ldots, J$;
  Compute direction vectors $\vec{v}_{j,i}$ for consecutive trajectory points.
Step 2: Partition the study area into regular grids $G = \{g_n | n = 1, 2, \ldots, N\}$, and assign each trajectory point $P_{j,i}$ to its corresponding grid cell g_n
Step 3: Compute intra-grid features
For each grid $g_n \in G$
  F_1 : number of trajectory points, $N_g = |P_{g_n}|$
  if $N_g > 3$
    Compute convex hull area $S_{hull}(g_n)$ and the minimum bounding circle area $S_{mincir}(g_n)$
    F_2 : HMC Index, $HMC = S_{hull}(g_n)/S_{mincir}(g_n)$
    F_3 : convex hull area $S_{hull}(g_n)$
  Otherwise
    Set $F_2 = 0$ and $F_3 = 0$
  End if
  F_4: number of direction cluster identified by DBSCAN $(\theta, MinPts)$
  F_5: trajectory points density, $\rho_{hull}(g_n) = N_g/S_{hull}(g_n)$
  F_6: DTW-based heterogeneity index computed from pairwise DTW distances between trajectory groups
End for
Step 4: Compute neighborhood features
For each grid $g_n \in G$
  Δden8: density differences between g_n and its eight neighborhood cells
  Δdir4: 4 directional density difference in four principal directions
End for
Step 5: Construct 18-dimensional grid-level feature vector
  $F_{grid} = [F_1,\ F_2,\ F_3, F_4, F_5, F_6,\ \Delta den8,\ \Delta dir4]$
Step 6: Apply Random Forest classification to identify key and non-key grid cells
Step 7: Convert the classified key grids into a binary raster image
Step 8: Perform morphological-aware optimization
  Remove small connected components with fewer than 100 pixels
  Apply morphological closing using 9-pixel cross-shaped structuring element
  Extract the road skeleton using an improved 4-neighborhood thinning algorithm
Step 9: Smooth the skeleton using the Kalman filter
Step 10: Transform the skeleton back to geographic coordinates and output final road network

The preprocessing stage involves three sequential steps to prepare the raw trajectory data for subsequent analysis. First, redundant points and abnormal trajectory segments—characterized by unrealistic distance or speed values—are removed through the construction of Euclidean distance and velocity matrices. Second, direction vectors

are computed for each trajectory point P_i based on the displacement between temporally adjacent observations, capturing local movement orientation. Finally, the study area is partitioned into uniform square grids $G = \{g_n | n = 1, 2, \ldots, N\}$, where N is the total number of uniform square grids. Each grid cell g_n thus serves as a localized spatial unit for further multi-level feature extraction. The subset of points contained within g_n is denoted as $P_{g_n} = \{p_{n,j,i} | j \in \{1, \ldots, J\}, i \in \{1, \ldots, I_{n,j}\}\}$, where $p_{n,j,i}$ represents the ith trajectory point of the jth segment located within grid g_n, organized according to their respective trajectory segment identifiers.

2.2. Coupled Multi-Level Grid Feature Construction

Trajectory semantics are inherently hierarchical: points interact within a cell, reflecting local dispersion and directionality, while road continuity emerges from neighborhood context. We therefore constructed a coupled feature set that integrates internal grid features, which describe within-cell geometry and movement organization, with neighborhood features, which characterize between-cell density continuity. For each grid cell, 18 features were computed, including six intra-grid features and twelve neighborhood features. The six intra-grid features include the number of trajectory points, the HMC index, convex-hull area, the number of directional clusters, trajectory point density, and the DTW-based heterogeneity index. The twelve neighborhood features consist of eight neighborhood density-difference features and four directional density-difference features.

2.2.1. Internal Grid Feature Indices

To capture the intrinsic geometric and semantic characteristics of trajectory distribution within each grid cell, a series of internal grid feature indices is computed. These indices reflect both spatial dispersion and directional complexity of movement patterns.

First, the convex-hull geometry is computed from the set of trajectory points in each grid. The hull area $S_{hull}(g_n)$, hull perimeter $L_{hull}(g_n)$, and hull-based point density $\rho_{hull}(g_n)$ (i.e., number of points per unit hull area) were used to quantify spatial compactness and dispersion.

$$\rho_{hull}(g_n) = \frac{|P_{g_n}|}{S_{hull}(g_n)} \tag{2}$$

In addition, two standard geometric enclosures are derived: the minimum bounding circle, characterized by its area $S_{mincir}(g_n)$ and circumference $L_{mincir}(g_n)$, and the minimum bounding rectangle, represented by its area $S_{minrect}(g_n)$. These structures help evaluate the shape and alignment regularity of trajectory clusters.

To enhance feature differentiation, two shape ratio indices are constructed, named HMR (Hull-to-Rectangle Ratio) and HMC (Hull-to-Circle Ratio), respectively. The calculation formulas are as follows:

$$HMR(g_n) = \frac{S_{hull}(g_n)}{S_{minrect}(g_n)} \tag{3}$$

$$HMC(g_n) = \frac{S_{hull}(g_n)}{S_{mincir}(g_n)} \tag{4}$$

These normalized ratios help assess how tightly the trajectory cluster conforms to different bounding shapes, indicating movement constraints or structural forms.

To quantify directional coherence, an improved DBSCAN clustering algorithm is applied to the direction vectors of trajectory points, with angular difference (rather than

Euclidean distance) as the distance metric. Specifically, the angular distance between two direction vectors $\vec{d}_a$ and $\vec{d}_b$ is defined as

$$\triangle\theta(\vec{d}_a, \vec{d}_b) = arccos\left(\frac{\vec{d}_a \cdot \vec{d}_b}{\left\|\vec{d}_a\right\|\left\|\vec{d}_b\right\|}\right) \tag{5}$$

As illustrated in Figure 1, each direction vector is mapped to a point on the unit circle, where clusters (e.g., orange, green, blue) represent movement along distinct dominant directions. A point is considered a core object if it has at least *MinPts* neighboring vectors with $\triangle\theta \leq \varepsilon$. Following the empirical evidence in existing trajectory-driven road-extraction studies, the angular threshold is set to $\varepsilon = 30°$ [5,12,25], which provides an optimal balance between accommodating GPS measurement noise and distinguishing topological road branches. Consequently, the number of direction clusters within the grid serves as a directional dispersion index, reflecting whether movement is well-aligned or chaotic.

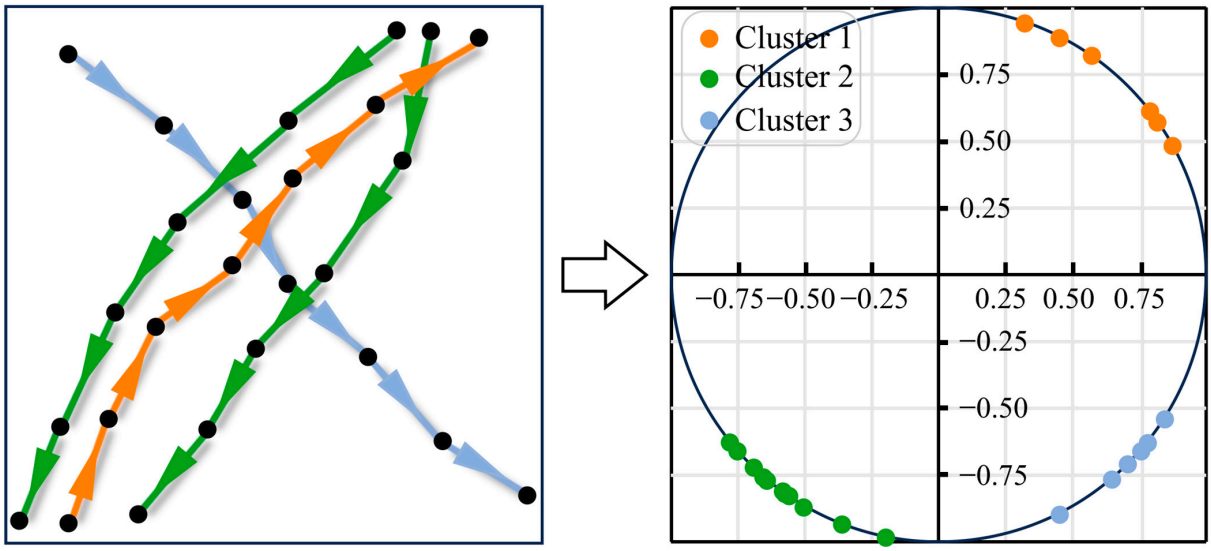


Figure 1. Schematic diagram of directional clustering of trajectory points within the grid. Arrows represent the movement directions of trajectory points. The schematic shows the process of grouping grid-level trajectory points by their movement directions to derive structured road network information from raw trajectory data.

Finally, to evaluate the trajectory segment similarity, the Dynamic Time Warping (DTW) algorithm is used to calculate pairwise distances between all trajectory segments within a grid. For two ordered segments $Tr_a = \{p_i | i = 1, \ldots, m\}$ and $Tr_b = \{q_j | j = 1, \ldots, k\}$, DTW identifies an optimal warping path $W = \{w_1, w_2, \ldots, w_L\}$ that minimizes the cumulative distance:

$$\mathrm{DTW}(Tr_a, Tr_b) = \min_W \sum_{l=1}^{L} d(w_l) \tag{6}$$

where $w_l = (i, j)$ represents the alignment between point p_i and q_j. The path W is subject to boundary conditions $w_1 = (1, 1)$ and $w_L = (m, k)$, as well as monotonicity and continuity constraints.

We then defined a grid-level heterogeneity (chaotic) index as the proportion of segment pairs whose DTW distance exceeded a threshold τ, reflecting how consistently trajectories align within the grid (Figure 2).

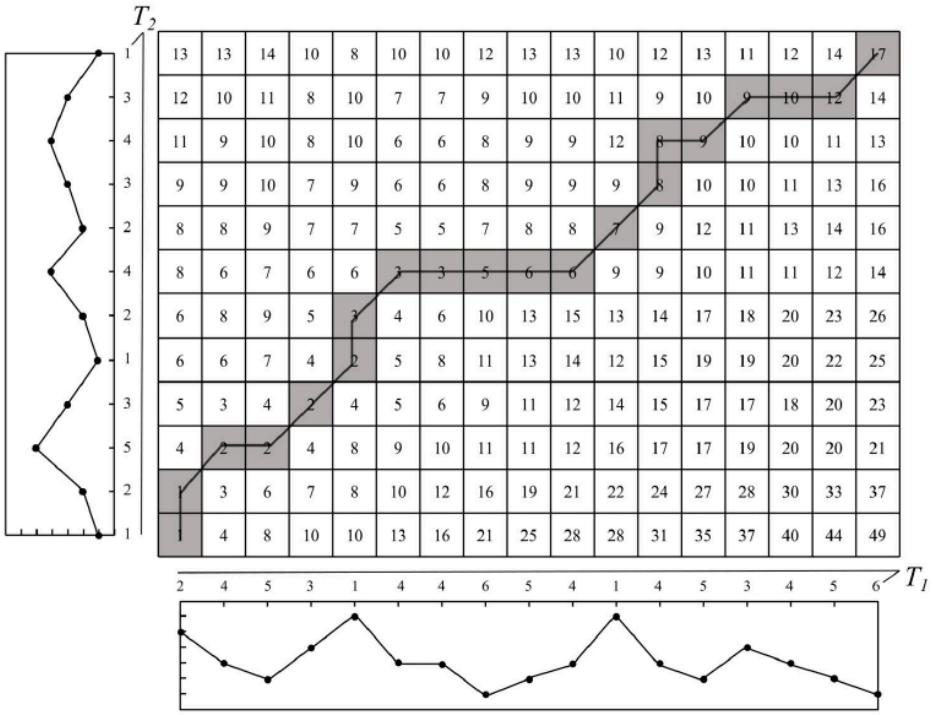


Figure 2. Visualization of the DTW principle. The black polyline denotes the optimal warping path that minimizes accumulated DTW distance under monotonicity and continuity constraints, and gray-shaded grid cells correspond to grid units along the optimal matching path. Marginal plots on the left and bottom depict the raw variation curves of two time-series, respectively.

2.2.2. Neighborhood Grid Feature Indices

Given the spatial continuity and connectivity inherent in road infrastructure, the contextual relationships between adjacent spatial units are crucial for accurate road network inference. To capture such relationships, this study introduces two neighborhood-level feature indices based on grid-level trajectory point densities $\rho(g_n)$, as illustrated in Figure 3.

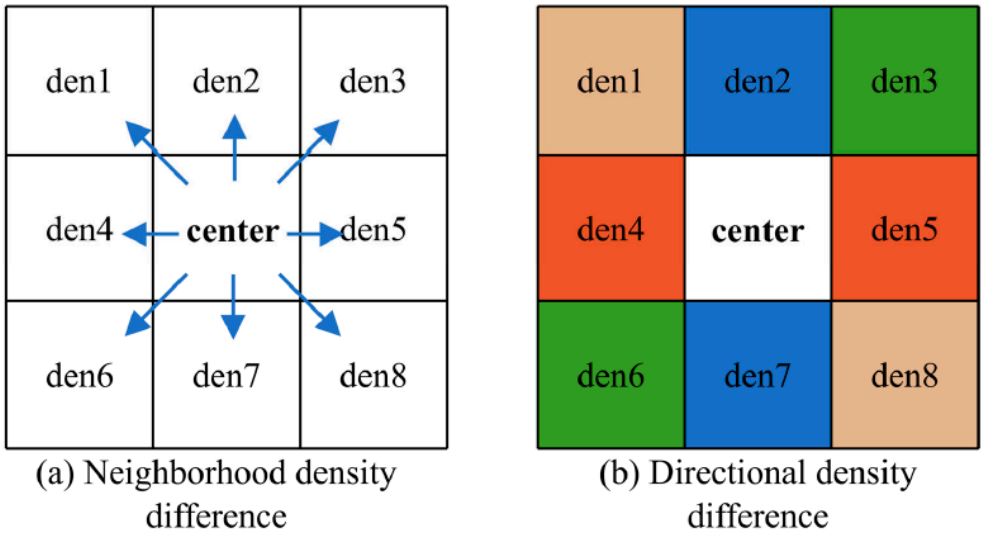


Figure 3. Calculation of neighborhood feature index. (**a**) Neighborhood density difference: the blue arrows denote the calculation directions of density discrepancy between the central grid cell and its eight surrounding adjacent cells. (**b**) Directional density difference: cells filled with identical colors belong to the same directional group, and density differences are computed between the central cell and each corresponding color-coded directional grid set. These two components characterize the spatial and directional heterogeneity of trajectory point distributions in the grid neighborhood.

The Neighborhood Density Difference (NDD) quantifies local density variation between the central grid and each of its eight immediate neighbors (Figure 3a). For every adjacent grid den_1 through den_8, the difference in trajectory point density relative to the central cell is computed. Missing neighbors (e.g., boundaries) were assigned a difference of 0. This produces eight neighborhood density-difference features describing local density gradients.

The Directional Density Difference (DDD) further captures anisotropic spatial patterns by focusing on opposing pairs of grids along the same directional axis—namely, vertical, horizontal, and diagonal directions (Figure 3b). For each axis, the absolute difference in density between the two opposing neighbors (e.g., den_2 and den_7) is calculated and averaged to produce a directional smoothness index. This feature reflects the directional continuity of traffic flows, which is vital for detecting linear road segments and distinguishing between regular road grids and cul-de-sacs or intersections.

Together, these two indices provide a robust description of the spatial context surrounding each grid, enhancing the classifier's ability to identify key road-related regions based on both local density gradients and directional consistency.

2.3. Key-Grid Detection via Supervised Learning

Key-grid detection was formulated as a binary classification task: each grid cell g_i is labeled as a key grid (1) or non-key grid (0) according to the following labeling function:

$$L(g_i) = \begin{cases} 1 & \text{if } P_{g_i}^{CtrPt} \in \Omega_{\text{road}} \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

where $P_{g_i}^{CtrPt}$ denotes the center point of grid index i, and Ω_{road} represents the reference road surface that derived from OpenStreetMap (OSM). The coordinates of the grid center point are calculated as:

$$center_lat = Lat_{min} + (grid_x + 0.5) \cdot \Delta Lat \tag{8}$$

$$center_lon = Lon_{min} + (grid_y + 0.5) \cdot \Delta Lon \tag{9}$$

where Lat_{min} and Lon_{min} are the minimum latitude and longitude of the study area, $grid_x$ and $grid_y$ are the grid indices, and ΔLat and ΔLon are the grid resolutions. Since the coordinates of the grid center point are calculated directly, both latitude and longitude coordinates in the formula require an addition of 0.5 to account for a linear translation equivalent to half the step size.

Label generation is executed through a deterministic, semi-automated workflow to ensure spatial consistency. The primary labeling criterion is based on a geometric point-in-polygon (PIP) relationship: a grid cell is automatically labeled as a road unit if its geometric center is located within the buffer zone of the reference road network (derived from OSM). To account for complex environments such as sidewalks and narrow trails, high-resolution remote-sensing imagery is utilized for manual verification. This manual step serves as a secondary quality control to ensure that only grids with a dominant road presence, specifically those where the road surface coverage exceeds 50% of the grid area, are maintained as positive samples.

A Random Forest (RF) classifier is employed to perform the classification task. As an ensemble learning method based on decision tree aggregation, Random Forest exhibits high tolerance to noise, effective handling of nonlinear feature interactions, and strong generalization capabilities, particularly for moderate-scale spatial datasets. In this application, each grid cell is represented by an 18-dimensional feature vector, which includes the internal semantic indicators (e.g., shape descriptors, directional clustering, trajectory

similarity) and neighborhood-context features (e.g., density gradients). The center point of the grid serves as the spatial reference for classification.

As illustrated in Figure 4, the training dataset is randomly partitioned into multiple subsets, each feeding into a different decision tree using a distinct combination of feature subsets (e.g., $M_1, M_2, \cdots$). Each tree independently outputs a predicted class label for the given grid cell. The final classification result is determined through majority voting across all trees in the ensemble. This voting-based mechanism reduces the risk of overfitting and increases classification robustness across spatially heterogeneous regions.

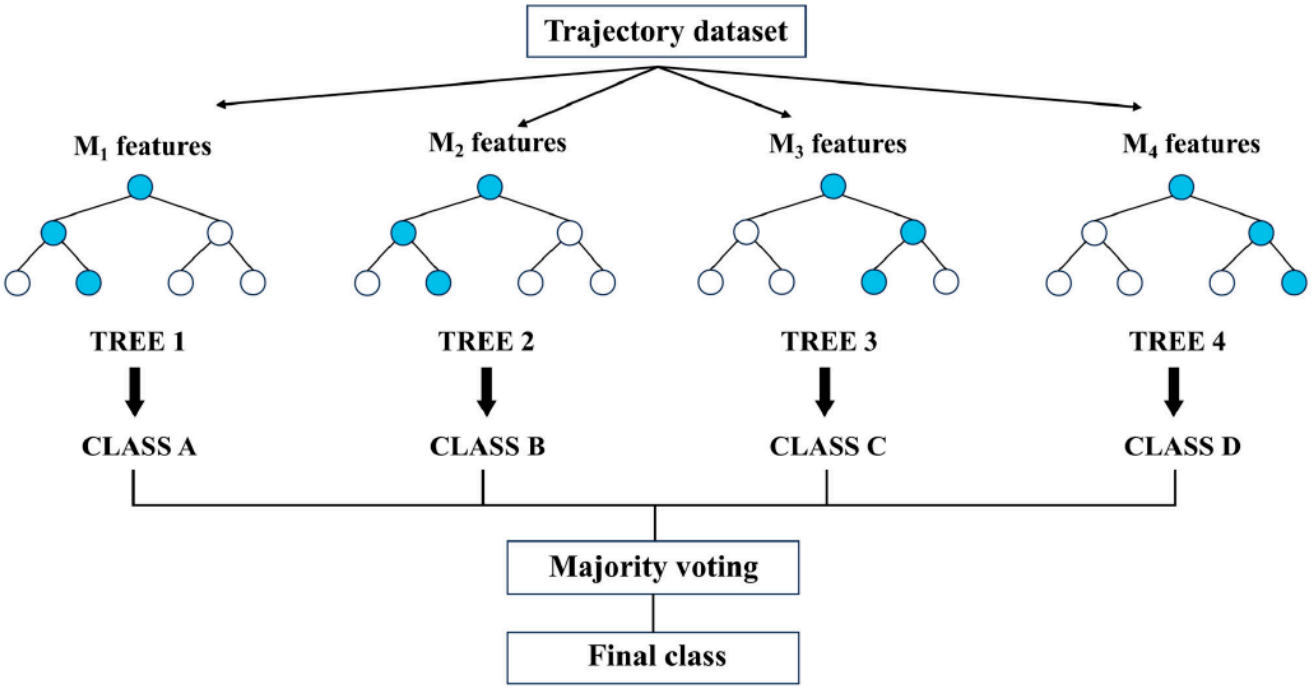


Figure 4. Random Forest principle. This figure illustrates the Random Forest algorithm, which builds multiple decision trees using feature subsets split from the trajectory dataset and aggregates independent predictions to yield robust classification results. Blue solid and white hollow circles denote selected and unselected split nodes within decision trees, respectively; black arrows indicate the data flow from feature grouping, individual tree prediction, to majority voting for final classification.

The model not only provides reliable classification results but also facilitates feature-importance analysis, enabling interpretation of which semantic indicators contribute most to key-grid identification. This aspect is particularly valuable for understanding which spatial characteristics are most predictive of road presence in trajectory-based representations.

2.4. Morphology-Based Road Network Reconstruction

Upon classification of key grids, a binary raster image is generated by mapping the centers of positively identified grid cells into pixel space, thereby forming a coarse representation of potential road regions. To derive a geometrically clean, topologically connected, and single-pixel-wide road skeleton, we implement an enhanced morphological thinning algorithm grounded in four-neighborhood connectivity.

As shown in Figure 5, several types of abnormal structures frequently arise in the initial binary output, such as densely packed square clusters, elongated blocks, and staircase-like patterns. These formations disrupt topological continuity and visual clarity, and must be addressed to produce a vectorization-ready network.

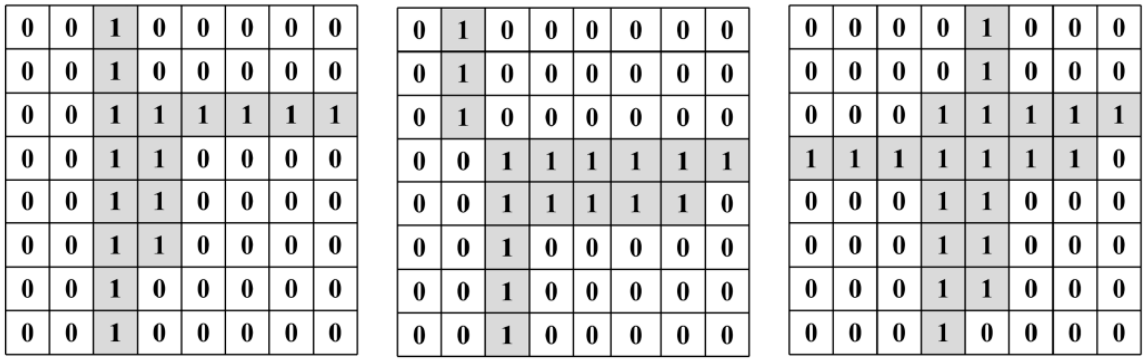


Figure 5. Schematic diagram of an abnormal structure. Grey-shaded cells represent pixels classified as road with a binary value of 1, whereas blank white cells correspond to non-road pixels with value 0. The figure illustrates common artifacts in raw road-extraction results: dense square clusters, elongated blocks, and staircase patterns, which impair topological integrity and require refinement for vectorization.

The proposed refinement procedure targets two key artifact types. First, square and block artifacts—identified by detecting 2×2 or larger high-density pixel regions—are simplified through dimensionality reduction. Specifically, such blocks are converted into transitional triangular patterns, enabling further processing through directional rules. Pixels at the periphery of the blocks with low neighborhood connectivity are iteratively removed, yielding thinner and more centralized road representations. Second, to eliminate residual connected triangle artifacts, we analyze four typical configurations as illustrated in Figure 6. These structures consist of three connected pixels forming an angled "L" shape. For each configuration, the algorithm checks a designated set of outer neighborhood pixels. If the sum of pixel values in this surrounding region is below a certain threshold—indicating isolation or redundancy—the triangle is removed from the network. This ensures that only structurally meaningful segments are retained.

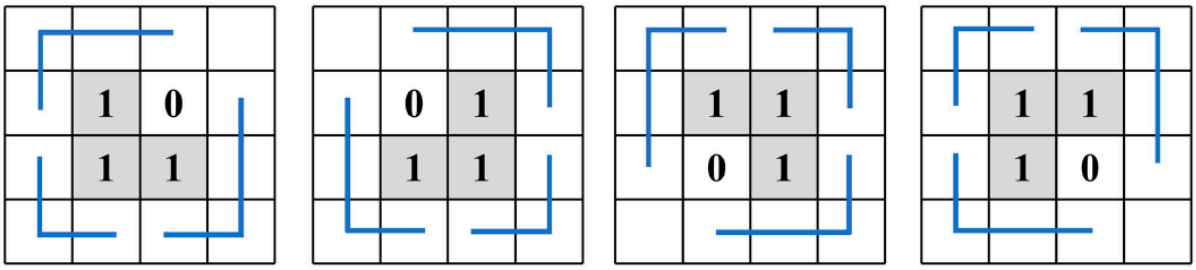


Figure 6. Four different structures of connected triangles. Grey cells represent core road pixels (value = 1) of triangular defects, and blue lines bound the inspected neighboring grids. These L-shaped three-pixel triangles are removed when the cumulative value inside blue-bordered areas falls below a predefined threshold, preserving valid road skeletons.

By combining artifact detection (Figure 5) with structure-aware triangle removal (Figure 6), the proposed morphological process significantly improves the geometric fidelity, topological continuity, and visual smoothness of the extracted road network skeleton. This post-processing step enhances the usability of the output in downstream GIS tasks, such as map matching, vectorization, and spatial topology analysis.

3. Results

3.1. Experimental Data and Study Areas

This study utilizes two types of trajectory data: pedestrian and vehicle, covering three cities, Shenzhen, Wuhan, and Changsha.

Pedestrian trajectory data were collected from the Yuehai Campus of Shenzhen University, covering an area of 1.44 km^2. Characterized by a complex road network, dense vegetation, and numerous hidden footpaths, this area is highly suitable for research on fine-grained pedestrian road extraction. The data were sampled at 1 s intervals and contain longitude, latitude, and timestamp fields, totaling 212 trajectories and 85,802 GPS points.

Vehicle trajectory data were obtained from the urban areas of Wuhan and Changsha. To investigate urban road networks under different topological structures, experimental areas were selected based on road density variations. Wuhan exhibits significant spatial heterogeneity in road distribution; thus, two experimental regions were established: Wuhan experimental zone 1 (30.5377–30.6152° N, 114.2334–114.3320° E) covering 81.3 km^2, and Wuhan experimental zone 2 (30.4538–30.5368° N, 114.2736–114.3639° E) covering 79.8 km^2. Changsha features a regular radial-grid network with high density and connectivity. Two regions were likewise established: Changsha experimental zone 1 (28.1764–28.2232° N, 112.9616–113.0160° E) covering 27.7 km^2, and Changsha experimental zone 2 (28.2181–28.2648° N, 112.9063–112.9607° E) covering 27.6 km^2. Both vehicle datasets include longitude, latitude, and timestamp attributes.

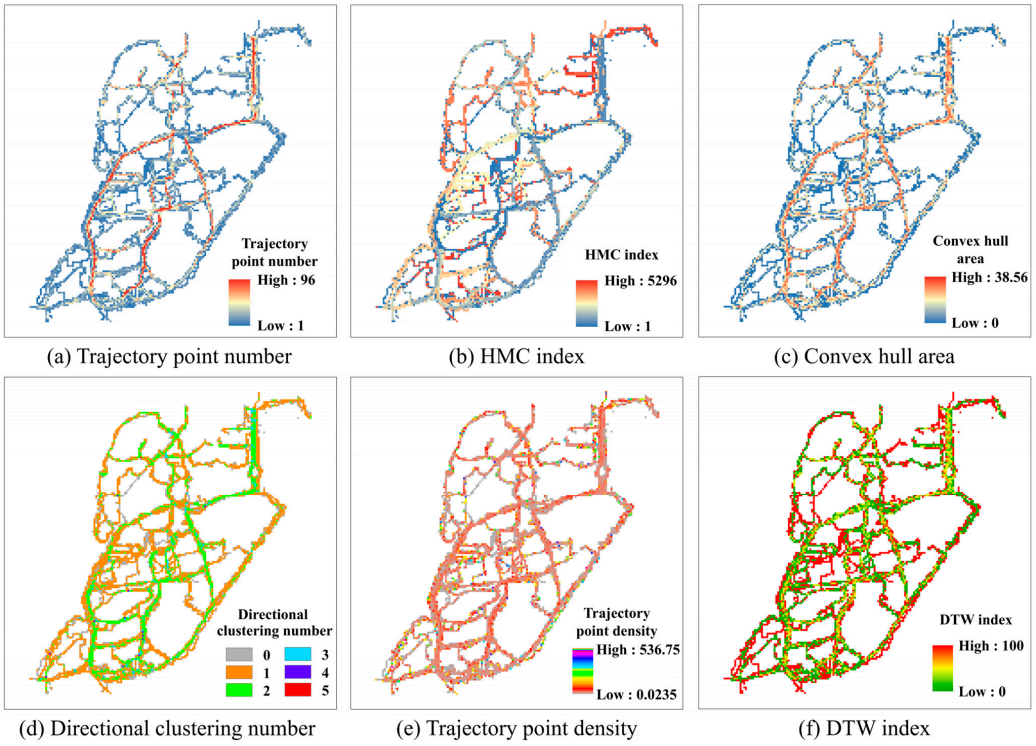
Auxiliary data include Jilin-1 high-resolution satellite imagery with a fused resolution of 0.75 m, used to extract the hidden campus pedestrian network. OpenStreetMap (OSM) road networks for Wuhan and Changsha were selected as the benchmark for experimental accuracy evaluation.

3.2. Experimental Setup and Feature Visualization

We trained the classifier on pedestrian trajectories collected at the Yuehai Campus of Shenzhen University. Each record contains longitude, latitude, and timestamp. To test spatial and modal transferability, we then applied the trained model to vehicle trajectories in Wuhan and Changsha. High-resolution remote-sensing imagery and OSM road centerlines were used as external references. Road-extraction quality was evaluated using a buffer-matching protocol: an extracted segment was counted as correct if it overlapped a 10 m buffer around the reference road centerlines. For the labeled Shenzhen dataset, we report Precision, Recall, and F1-score for binary key-grid detection. For the larger Wuhan and Changsha transfer regions, where exhaustive grid-level labels are unavailable, we evaluate end-to-end road-extraction quality using the length of correctly extracted roads, the total extracted length, and length-based precision under the OSM-buffer matching protocol.

Figure 7 visualizes representative grid-level features, including the number of trajectory points, the HMC index, convex-hull area, directional-cluster count, point density, and the DTW-based chaotic index. These maps reveal coherent spatial patterns along arterial corridors and at intersections, supporting subsequent supervised detection of key grids.

More specifically, the number of trajectory points per grid ranges from 0 to 96. Higher values concentrate along main campus corridors (e.g., Liyan Road, Lide Road, and Ligong Road), indicating denser pedestrian usage. The HMC index shows higher values in smaller pathways where trajectories are tightly clustered, while convex-hull area becomes larger along major roads where movements span a wider within-grid space. Directional cluster counts are predominantly one to two in most grids, but increase to three or four near intersections, consistent with multi-directional turning behaviors. The DTW-based index tends to be lower along structured corridors with aligned movement, and becomes higher in open plaza areas (e.g., “Shiguang Square”) where movements are less constrained and trajectory segments are more heterogeneous.



(a) Trajectory point number (b) HMC index (c) Convex hull area

(d) Directional clustering number (e) Trajectory point density (f) DTW index

Figure 7. Feature visualization. (**a**) Trajectory point number; (**b**) HMC index; (**c**) Convex-hull area; (**d**) Directional clustering number; (**e**) Trajectory point density; (**f**) DTW index. These panels illustrate the multi-dimensional features derived from trajectory data for road network extraction.

To justify the proposed feature set and examine the relative contribution of different semantic dimensions, a Random Forest-based feature-importance analysis was conducted. For clearer presentation, individual neighborhood-related variables were aggregated into their corresponding feature groups. As shown in Figure 8, geometric and movement-structure features, such as hull area, DTW-based heterogeneity, and the HMC index, contribute the most to key-grid detection. Importantly, neighborhood-context features also contribute substantially to the model, indicating that inter-grid density continuity provides complementary information beyond within-grid geometry. These results support the rationale of coupling intra-grid movement semantics with inter-grid contextual features, rather than relying on either feature group alone.

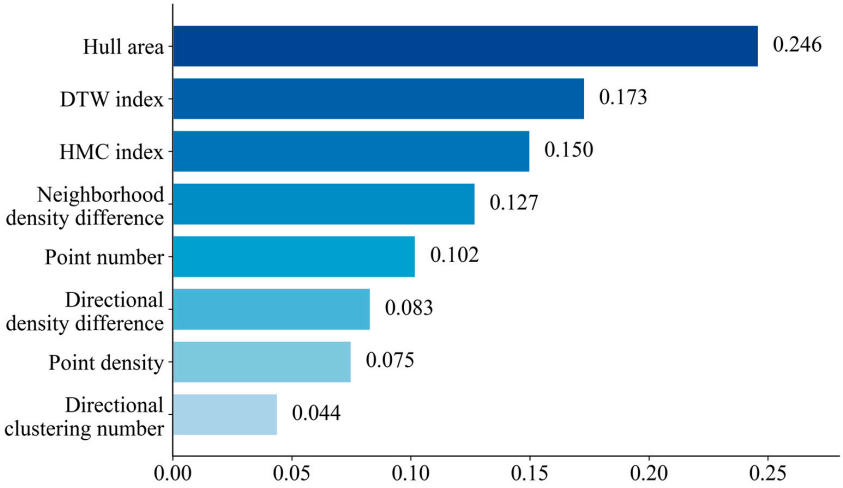


Figure 8. Feature-importance ranking based on the Random Forest model (where directional and neighborhood density differences are aggregated into their respective categories).

3.3. Key-Grid Detection Performance Under Varying Grid Sizes and Classifiers

We assessed sensitivity to grid resolution by testing 100 × 100, 200 × 200, 300 × 300, 400 × 400, and 600 × 600 pixel divisions. Labels were generated using the deterministic point-in-polygon rule described in Section 2.3 and were subsequently quality-checked against high-resolution imagery and campus road references. A Random Forest classifier was trained with an 80/20 train–test split. As summarized in Table 1, the 200 × 200 grid achieved the most balanced result (Precision 83%, Recall 79%, F1 0.81 for key grids).

Table 1. Classification performance under different grid divisions. This table summarizes the grid division configurations, including their actual physical dimensions and corresponding classification categories (key grid/non-key grid) used in the performance evaluation.

Grid Division Size	Actual Size of a Single Grid	Category	Precision	Recall	F1 Score
100 × 100	About 12 m × 12 m	Key grid	71%	62%	0.66
		Non-key grid	78%	85%	0.81
200 × 200	About 6 m × 6 m	Key grid	83%	79%	0.81
		Non-key grid	77%	80%	0.78
300 × 300	About 4 m × 4 m	Key grid	76%	73%	0.74
		Non-key grid	62%	66%	0.64
400 × 400	About 3 m × 3 m	Key grid	74%	75%	0.75
		Non-key grid	55%	53%	0.54
600 × 600	About 2 m × 2 m	Key grid	68%	99%	0.80
		Non-key grid	50%	3%	0.06

The effectiveness of the proposed trajectory-based features is sensitive to grid resolution. As the grid size becomes very small (e.g., 600 × 600), although the total number of samples increases substantially, each grid cell contains significantly fewer trajectory points and segments. This results in dilution of semantic information within individual grids, reduced directional consistency, weaker aggregation signals, amplified noise from sparse points, GPS jitter, and partial road coverage, as well as decreased discriminative power of key features such as trajectory point density, HMR, and HMC indices, direction clustering results, and neighborhood contrasts. Conversely, very large grids enhance statistical robustness within each cell but reduce the number of training samples and lower the overall spatial resolution. This creates a clear trade-off between sample quantity, feature quality, and spatial precision. The 200 × 200 grid resolution provided the best balance among these competing effects, achieving the highest classification performance.

Under the same feature set and the recommended 200 × 200 grid setting, we further compared three supervised classifiers for key-grid detection: Random Forest (RF), Support Vector Machine (SVM), and Extreme Gradient Boosting (XGBoost). Hyperparameters were tuned using grid search on the Shenzhen training data, and performance was evaluated using Precision, Recall, and F1-score. As summarized in Table 2, RF achieved the best overall performance (F1 = 0.79), slightly outperforming SVM (0.78) and XGBoost (0.75). Therefore, RF was adopted as the default classifier in subsequent transfer and end-to-end extraction experiments.

Because the 80/20 grid-level split may still retain spatial dependence among neighboring cells, the within-Shenzhen experiment is mainly used for grid-scale selection and classifier comparison. Generalization is therefore further assessed through the independent zero-shot transfer experiments in Wuhan and Changsha.

Table 2. Comparison of performance indicators of different models. This table compares the performance metrics of three machine learning models (RF, SVM, and XGBoost) used for trajectory-based road network classification, with evaluation indicators presented across columns for quantitative comparison.

Training Evaluation Indicators / Training Model	Precision	Recall	F1 Score
Random forest	80%	79%	79%
SVM	77%	77%	78%
XGboost	75%	75%	75%

3.4. Cross-City and Cross-Modal Transferability of Key-Grid Detection

The identification accuracy of key grids serves as the foundation for constructing road networks. We established an automatic key-grid detection model using pedestrian trajectories from Shenzhen University. Under a grid division of 200 × 200, the training set comprised 5450 samples, and the model achieved a training accuracy of 83% for key grids. The test set contained 1363 samples; under this setting, the model predicted 741 grids as key grids and 622 grids as non-key grids. The spatial distribution of predicted key grids in Shenzhen is shown in Figure 9.

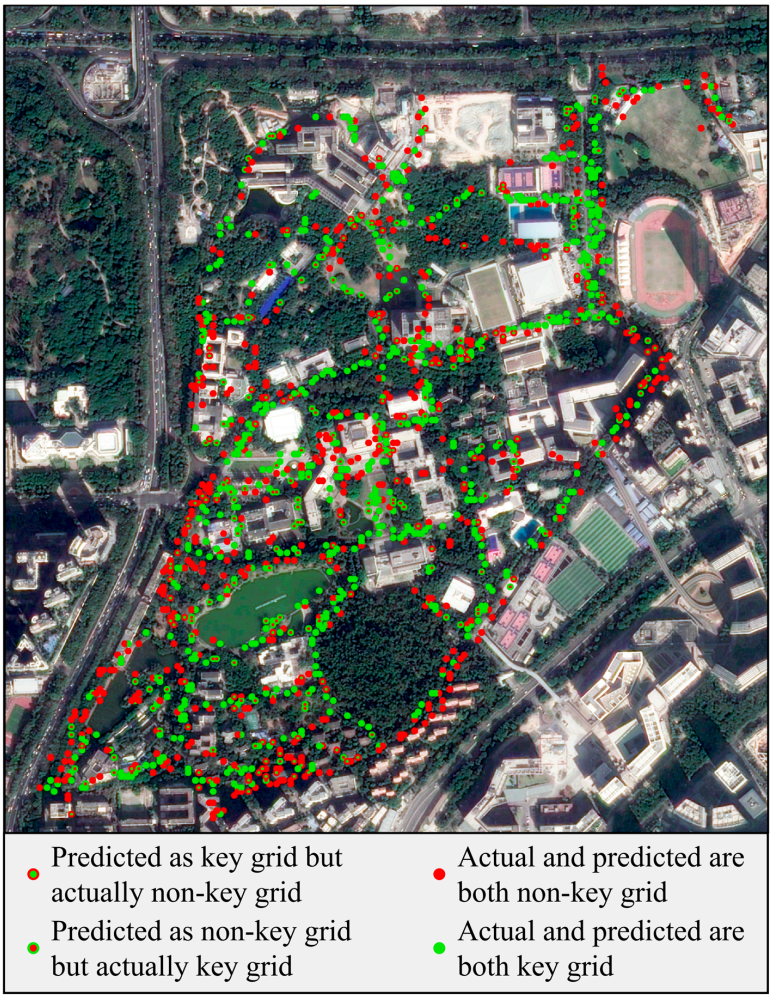


Figure 9. Key-grid identification results in Shenzhen area. Red–green dot: predicted as key grid but actually non-key grid; predicted as non-key grid but actually key grid (false predictions). Red dot: actual and predicted are both non-key grid (true negative). Green dot: actual and predicted are both key grid (true positive). This figure visualizes the spatial distribution of key-grid identification results in the Shenzhen study area.

To verify the generalization capability, this study employs a zero-shot approach to directly transfer the trained model to the Wuhan and Changsha regions. During validation, the model parameters and decision thresholds were retained entirely from the optimal configuration on the training set, without any targeted fine-tuning or parameter adjustments. The identification results are illustrated in Figure 10. For Wuhan experimental zone 1, out of 11,189 sample grids, 7568 were identified as key grids and 3621 as non-key grids; for Wuhan experimental zone 2, among 7922 grids, 4983 were classified as key grids and 2939 as non-key grids. For Changsha experimental zone 1, out of 15,091 grids, 9069 were identified as key grids and 6022 as non-key grids; for Changsha experimental zone 2, among 9265 grids, 5232 were recognized as key grids and 4033 as non-key grids.

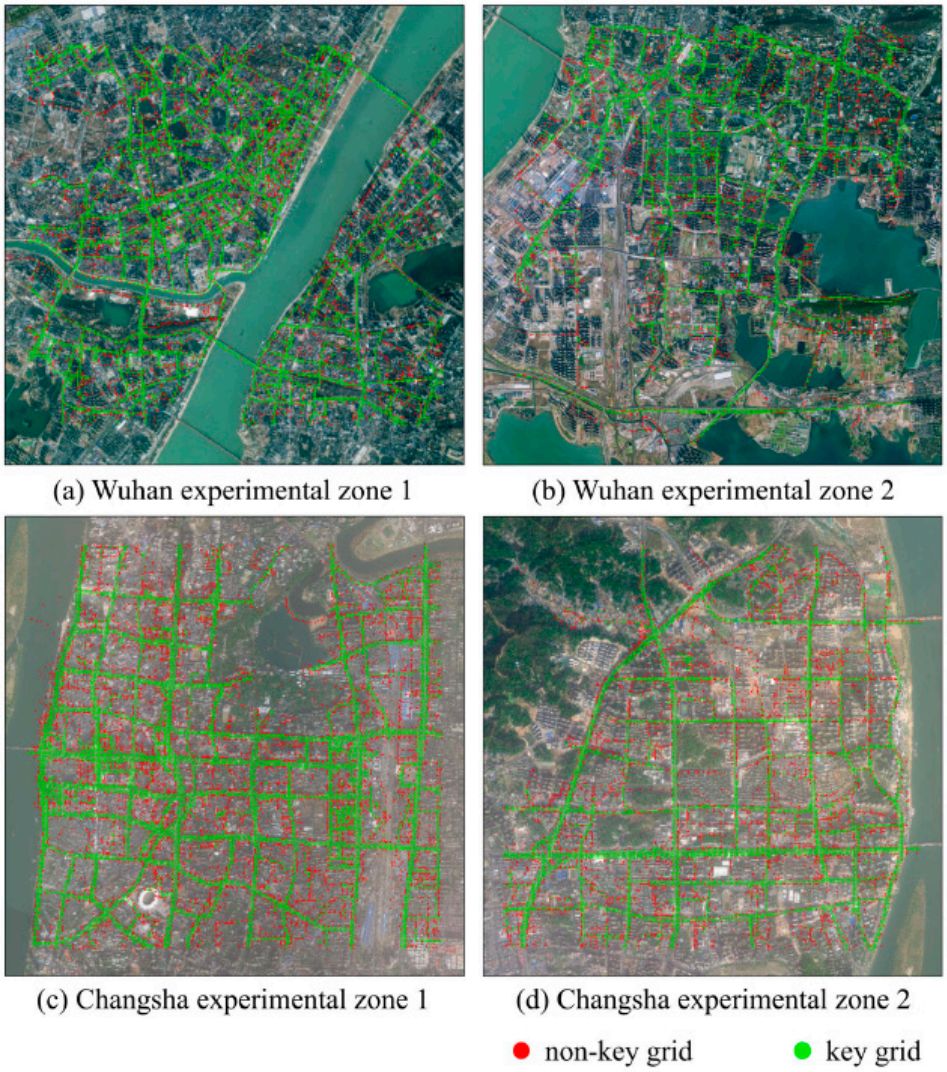


Figure 10. Key-grid identification results of Wuhan and Changsha experimental zones. (**a**) Wuhan experimental zone 1; (**b**) Wuhan experimental zone 2; (**c**) Changsha experimental zone 1; (**d**) Changsha experimental zone 2. Green dots represent identified key grids, and red dots indicate non-key grids across the four study areas.

Overall, the spatial distribution of predicted key grids aligns closely with the apparent road corridors in the imagery overlays, and the model shows consistent behavior across regions with high road network density, sparse road distribution, and heterogeneous layouts (Figure 10).

3.5. Morphological Post-Processing and Curve Smoothing

Starting from the classified key-grid outputs, we converted the predicted key grids into a binary raster representation. We then applied a morphology-aware refinement pipeline to obtain a topology-preserving, vectorization-ready road skeleton. First, isolated components and salt-and-pepper noise were removed using connected-component filtering, which suppressed spurious pixels unlikely to correspond to road structures. We then performed a morphological closing operation with a 9-pixel cross-shaped structuring element to bridge small gaps and enhance local continuity, followed by four-neighborhood thinning to extract a one-pixel-wide skeleton while preserving the main corridor topology.

The effect of these operations is illustrated in Figure 11. Panel (a) shows the trajectory points retained after key-grid screening, which already highlight corridor structures but still contain discontinuities and small isolated fragments. Panel (b) presents the binary image after noise removal, where isolated artifacts were reduced. Panel (c) shows the result after closing, where short breaks between adjacent segments were effectively filled, providing a more continuous support for thinning and skeletonization. Quantitatively, the morphology stage improved structural continuity: in the Wuhan experimental zone 1, against a complex background composed of approximately 3.2 million pixel units, the number of connected components decreased from 144,913 to 4180, corresponding to a 97.1% reduction. The spurious endpoint ratio decreased from 5.42% to 0.2% when computed on the skeleton graph, indicating fewer fragmented segments and fewer dangling branches. This reduction indicates that the initial key-grid output contains many fragmented candidate components, primarily due to sparse trajectory points, GPS drift, and isolated false-positive grids across a large urban raster. Because these components are mostly small and spatially disconnected, they are better interpreted as candidate-level fragmentation rather than failure of the key-grid classifier. The subsequent morphology-aware refinement therefore acts as a spatial-scale filter that removes isolated fragments, bridges short gaps, and preserves the connected structural backbone of the road network.



Figure 11. The process of morphological processing of images. (**a**) Raw trajectory points extracted via key-grid identification; (**b**) Preprocessed trajectory points following noise and outlier filtering; (**c**) Smoothed and connected trajectory points obtained via a closing morphological operation.

To further regularize geometry and improve visual smoothness, we smoothed the extracted centerlines using Kalman filtering applied to sequences of grid-center coordinates (treated as observations). As shown in Figure 12, Kalman smoothing mitigated staircase artifacts and small zigzag oscillations caused by grid discretization while maintaining junction configurations. The zoom-in views demonstrate that the fitted centerlines (red) closely follow the refined skeleton (black) but provide a smoother, more consistent geometry.

This improvement can also be quantified by a reduction in geometric roughness, e.g., the mean turning-angle (or curvature surrogate) enhanced from 96.6883 to 99.6385, and the number of short spurs shorter than 20 m reduced from 25,751 to 2458, which facilitates downstream vectorization and network analysis.



Figure 12. Kalman filter fitting effect diagram. Gray: refined road network; Red: fitted road network (Kalman filter). Blue and green boxes indicate two local magnified areas displayed in the two inset panels for detailed fitting comparison. This figure demonstrates the fitting performance of the Kalman filter on road network data.

3.6. End-to-End Road Extraction and Baseline Comparison

Using the buffer-matching protocol, Table 3 compares the proposed method with two raster-based reference methods [7,38] across four experimental zones. Overall, our method achieves longer correctly extracted road length in most zones while maintaining competitive length-based precision, with the highest precision reaching 83% in Wuhan Zone 1. This indicates that the proposed multi-level grid semantics and neighborhood-context modeling contribute to both corridor completeness and structural continuity. The two reference approaches [7,38] represent raster-based trajectory-to-road-extraction pipelines. Their performance varies across urban scenes, suggesting that fixed rasterization or morphology-based rules can be sensitive to trajectory density, local network complexity, and noise distribution. The advantages of the proposed method are particularly evident in Changsha Zones 1–2, where it extracts substantially longer correct road segments (85.013 km and 60.176 km) than the reference method [38] (53.488 km and 35.942 km), indicating improved coverage of secondary streets and local connections.

For Wuhan zone 2, the reference method [38] achieves higher precision than the proposed method (81% vs. 68%), whereas our method recovers a longer correctly extracted road length (93.512 km vs. 77.960 km). This reflects a precision–coverage trade-off in heterogeneous urban environments: our method preserves more candidate corridors to improve network completeness, but may introduce a small number of false positives in open or mixed-use areas. Reference method [7], which relies on fixed morphological rules without explicit neighborhood-context modeling, tends to over-connect noisy trajectory fragments and produces redundant extracted segments, resulting in a relatively low precision of 54% in this zone. Compared with the raster-based references, the proposed method provides a more balanced extraction outcome by integrating intra-grid movement structures with

inter-grid contextual continuity. This pattern is consistent with Figure 13b, where the extracted network exhibits broader spatial coverage along minor streets and at complex intersections, with few false positives in open or mixed-use spaces.

Table 3. Comparison of end-to-end road-extraction performance among the proposed method and two raster-based reference methods [7,38] across four study areas in Wuhan and Changsha.

Experimental Area	Road-Extraction Method	Correctly Extract Length/km	Extract Length/km	Precision
Wuhan experimental zone 1	Ours	130.4843	157.210	83%
	Reference [38]	51.235	62.454	82%
	Reference [7]	211.916	380.958	56%
Wuhan experimental zone 2	Ours	93.512	136.467	68%
	Reference [38]	77.960	95.128	81%
	Reference [7]	116.752	217.926	54%
Changsha experimental zone 1	Ours	85.013	115.122	73%
	Reference [38]	53.488	85.373	62%
	Reference [7]	61.884	78.516	79%
Changsha experimental zone 2	Ours	60.176	73.366	82%
	Reference [38]	35.942	50.261	70%
	Reference [7]	28.283	31.326	90%

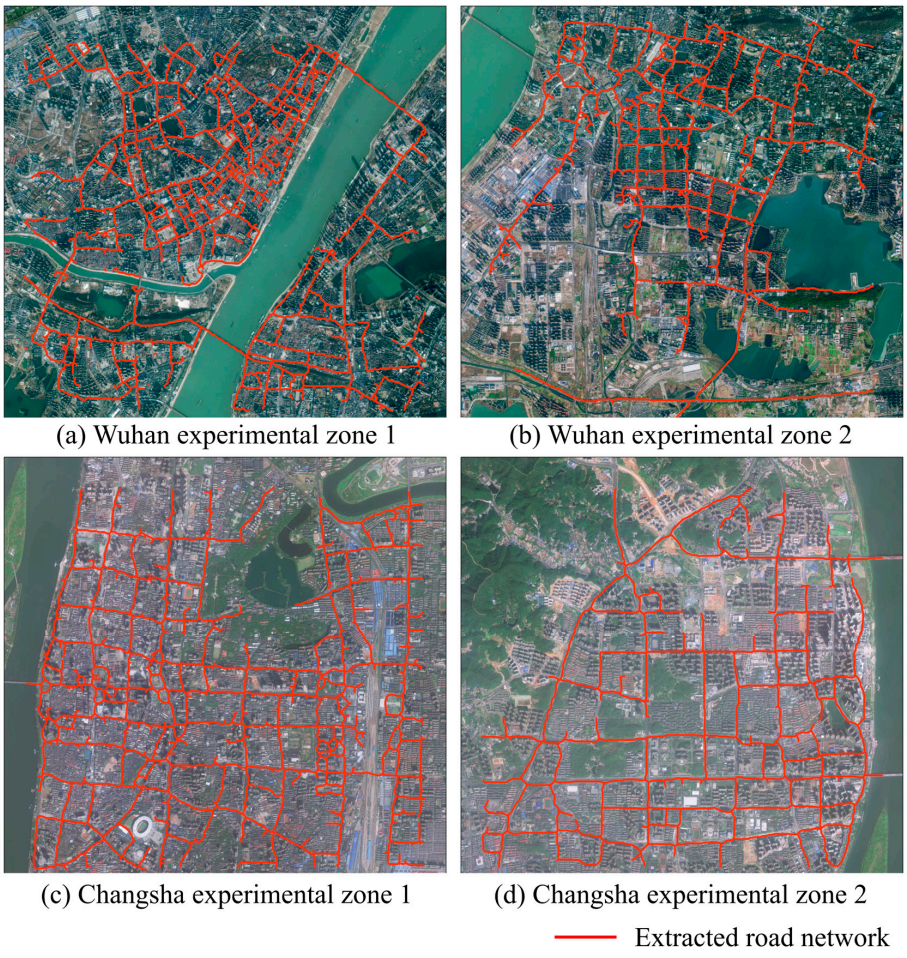


Figure 13. Road network extraction results. (**a**) Wuhan experimental zone 1; (**b**) Wuhan experimental zone 2; (**c**) Changsha experimental zone 1; (**d**) Changsha experimental zone 2. Red lines denote the extracted road network overlaid on satellite imagery; each panel shows the extracted road network overlaid on corresponding satellite imagery, revealing differences in road density and network patterns across urban districts.

Figure 13 provides qualitative overlays of reconstructed road networks on high-resolution imagery. Across the four zones, the proposed method better preserves arterial continuity and junction topology, producing more complete and connected corridor structures (Figure 13a–d). The improvement is particularly visible in dense grid-like neighborhoods (e.g., Changsha Zone 1) and in areas with varying road densities (e.g., Wuhan Zone 2), where the reference methods tend to miss fine structures or yield fragmented segments. These observations are consistent with the quantitative results in Table 3.

3.7. Ablation Study of Feature Components

To examine the contribution of neighborhood context, we conducted a component-level ablation study using the same 200 × 200 grid setting and the same evaluation protocol. Four feature configurations were compared: internal features only, neighborhood features only, internal features plus neighborhood density-difference features, and the full feature set.

As shown in Table 4, the internal-only configuration achieves F1-scores of 0.78 and 0.75 for key and non-key grids, respectively. Adding neighborhood density-difference features improves the F1-scores to 0.80 and 0.77, while the full feature set achieves the best overall performance, with F1-scores of 0.81 and 0.78. These results indicate that neighborhood features provide complementary contextual information and improve classification balance when combined with internal geometric and movement-structure features. In contrast, using neighborhood features alone leads to substantially lower non-key-grid performance, suggesting that neighborhood context is not sufficient as a standalone discriminator but is effective as a contextual supplement to intra-grid features.

Table 4. Component-wide ablation study demonstrating the classification performance under different feature combinations for key-grid detection. (a) Internal features, (b) Neighborhood features (8 + 4), which comprise the 8-dimensional Neighborhood density difference and 4-dimensional Directional density difference, (c) Internal features + eight neighborhood density-difference features, and (d) All features.

Feature Combination	Category	Precision	Recall	F1 Score
Internal features	Key grid	79%	78%	0.78
	Non-key grid	75%	76%	0.75
Neighborhood features (8 + 4)	Key grid	56%	85%	0.68
	Non-key grid	55%	22%	0.31
Internal features + Neighborhood features (8)	Key grid	81%	79%	0.80
	Non-key grid	76%	79%	0.77
All features	Key grid	83%	79%	0.81
	Non-key grid	77%	80%	0.78

4. Discussion

4.1. Interpretation of Model Choice and Scale Effect

The model comparison suggests that Random Forest (RF) yields the most stable performance for key-grid detection under the proposed feature set (Table 2). This outcome is consistent with the nature of trajectory-derived grid features, which are often noisy, heterogeneous, and characterized by nonlinear interactions. As an ensemble of decision trees, RF reduces variance through bagging and can capture nonlinear combinations among directional organization, density variation, and neighborhood continuity—factors that are critical for separating corridor-like road grids from open-space grids. In addition, RF provides an interpretable mechanism for examining feature contributions, which is valuable for understanding how spatial semantics drive classification outcomes in GIS-oriented workflows.

The grid-size sensitivity results reveal a trade-off between within-cell semantic sufficiency and spatial detail. Smaller grids provide more samples but often contain too few trajectory points to support stable internal descriptors, whereas larger grids improve within-cell statistics but blur local structures. The 200 × 200 setting, therefore, provides a practical compromise for road-level, rather than lane-level, extraction. Specifically, the resulting grid size (approximately 6 m × 6 m) is chosen to align with road-level network extraction rather than lane-level modeling. This resolution effectively aggregates scattered trajectory points from multiple lanes into a unified semantic unit, filtering out GPS positioning noise and ensuring a topologically consistent centerline during vectorization. While this may conflate closely parallel lanes, it remains consistent with the standard representation of road graphs in urban GIS applications.

4.2. Transferability and Typical Failure Cases

Applying the Shenzhen-trained model to vehicle trajectories in Wuhan and Changsha indicates encouraging cross-city and cross-modal transferability (Figure 10). The predicted key grids and the final extracted networks visually align with major corridors and intersection structures in dense urban cores and sparser peripheral areas, suggesting that the proposed feature set captures modality-invariant cues of linear mobility corridors, such as directional coherence and neighborhood continuity.

The robustness of this cross-modal and cross-city transfer is fundamentally rooted in the topological invariance of the feature design. While pedestrians and vehicles operate under significantly different physical and dynamic constraints (e.g., movement speed and lane adherence), they share the same underlying physical space when entering road corridors. Through the abstraction and aggregation of the grid, the model effectively filters out micro-level behavioral noise, thereby accurately identifying the structural backbone of the road network based on consistent directional signatures.

Nevertheless, several failure modes are expected under realistic urban conditions. First, elevated roads and underpasses can produce ambiguous trajectory patterns when vertical separation is not accounted for in a 2D formulation, leading to occasional false connections or missing segments. Second, parking-lot aisles and internal roads in large open spaces may generate scattered or multidirectional trajectories, increasing within-grid directional dispersion and DTW-based heterogeneity, thereby reducing the discriminability between road and non-road grids. Third, when sampling becomes very sparse or uneven, internal grid descriptors become unstable, and the extracted network may be fragmented. These limitations motivate the incorporation of additional cues (e.g., elevation/IMU signals) and the enforcement of graph-level constraints to improve topological consistency in future extensions.

4.3. Practical Implications, Limitations, and Future Work

From a practical perspective, the proposed key-grid layer functions as an interpretable candidate-generation step for map updating: it narrows the search space to high-probability corridor cells, thereby reducing manual inspection cost and enabling targeted human validation. Meanwhile, the morphology-plus-smoothing pipeline yields a vectorization-ready skeleton with improved geometric regularity while preserving junction structures (Figures 11–13), which facilitates downstream GIS operations such as network analysis and routing.

Three limitations should be noted. First, buffer-based matching against OSM centerlines may bias evaluation toward well-mapped arterials and penalize newly built or spatially misaligned roads, which can affect the validity of quantitative comparisons. Second, class imbalance is inherent to grid-based road detection because non-road cells

dominate; this can depress recall in fine-resolution settings and makes threshold selection consequential. Third, the current 2D representation cannot fully resolve stacked road structures (e.g., viaducts and underpasses), which may lead to false connections or omissions in complex interchanges.

Future work will therefore target these limitations in a coordinated manner: (a) incorporating multi-source references or curated subsets to reduce evaluation bias and support fair benchmarking, (b) adopting calibrated thresholds or cost-sensitive learning to better balance precision–recall under class imbalance, (c) integrating elevation/IMU cues and multi-layer representations to disambiguate vertical overlaps, and (d) applying graph-based repair and uncertainty estimation to enforce connectivity constraints and support human-in-the-loop validation in operational map updating workflows.

5. Conclusions

This study presented a supervised framework for trajectory-based road network extraction that couples intra-grid semantics with inter-grid neighborhood context. By transforming trajectories into a structured grid space and learning key grids with a Random Forest classifier, the pipeline produces a topology-preserving, vectorization-ready road skeleton through morphology-aware thinning and Kalman smoothing.

Experiments in Shenzhen, Wuhan, and Changsha demonstrated the framework's accuracy and transferability. The 200×200 grid division achieved the most balanced key-grid classification, and the Shenzhen-trained model was directly transferred to vehicle trajectories in other cities without retraining or threshold adjustment. End-to-end evaluation against OSM roads using a 10 m buffer further showed that the proposed method achieved longer correctly extracted road lengths and competitive length-based precision compared with two raster-based reference methods, while also revealing a precision–coverage trade-off in heterogeneous urban areas.

Feature-importance and ablation analyses further confirmed the complementary role of neighborhood context. The full feature set achieved the best overall classification performance, whereas neighborhood features alone were insufficient as standalone discriminators. Future work will extend the framework with stronger topological constraints, calibrated thresholding, and additional sensing cues, such as elevation or IMU data, to improve robustness in complex road environments.

Author Contributions: Conceptualization, Yunfei Zhang and Chaoyang Shi; methodology, Hongjie Zhu and Baifa Wu; software, Naisi Sun and Tianyu Zhong; validation, Naisi Sun and Cuifeng Zhang; formal analysis, Cuifeng Zhang; investigation, Tianyu Zhong; data curation, Baifa Wu and Cuifeng Zhang; writing—original draft preparation, Hongjie Zhu; writing—review and editing, Yunfei Zhang and Chaoyang Shi; visualization, Naisi Sun; supervision, Yunfei Zhang; funding acquisition, Yunfei Zhang and Chaoyang Shi. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant numbers 42371474, 41971421, 41903192; the Hubei Provincial Natural Science Foundation—Innovation and Development Joint Fund Project, grant number 2025AFD740; the Open Fund of the Engineering Laboratory of Spatial Information Technology for Highway Geological Disaster Early Warning in Hunan Province (Changsha University of Science & Technology), grant number kfj230701; the Wuhan Pilot Construction of a Strong Transportation Country Science and Technology Joint Research Projects, grant number 2024-2-6; the Open Topic of the Hunan Engineering Research Center of 3D Real Scene Construction and Application Technology, grant number ReS3D2025Y3; the Postgraduate Scientific Research Innovation Project of Hunan Province, grant number CX20230861.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments: The authors would like to thank the anonymous reviewers and editors for their valuable comments and suggestions, which greatly improved the manuscripts.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. van Winden, K.; Biljecki, F.; van der Spek, S. Automatic Update of Road Attributes by Mining GPS Tracks. *Trans. GIS* **2016**, *20*, 664–683. [CrossRef]
2. Xu, Y.; Xie, Z.; Feng, Y.; Chen, Z. Road Extraction from High-Resolution Remote Sensing Imagery Using Deep Learning. *Remote Sens.* **2018**, *10*, 1461. [CrossRef]
3. Yu, W. Assessing the Implications of the Recent Community Opening Policy on the Street Centrality in China: A GIS-Based Method and Case Study. *Appl. Geogr.* **2017**, *89*, 61–76. [CrossRef]
4. Li, Y.; Xiang, L.; Zhang, C.; Wu, H. Fusing Taxi Trajectories and RS Images to Build Road Map via DCNN. *IEEE Access* **2019**, *7*, 161487–161498. [CrossRef]
5. Zhang, C.; Li, Y.; Xiang, L.; Jiao, F.; Wu, C.; Li, S. Generating Road Networks for Old Downtown Areas Based on Crowd-Sourced Vehicle Trajectories. *Sens.* **2021**, *21*, 235. [CrossRef] [PubMed]
6. Zheng, Y. Trajectory Data Mining: An Overview. *ACM Trans. Intell. Syst. Technol.* **2015**, *6*, 29:1–29:41. [CrossRef]
7. Wen, W.; Zhang, W. Research on Urban Road Network Extraction Based on Web Map API Hierarchical Rasterization and Improved Thinning Algorithm. *Sustainability* **2022**, *14*, 14363. [CrossRef]
8. Huang, J.; Zhang, Y.; Deng, M.; He, Z. Mining Crowdsourced Trajectory and Geo-Tagged Data for Spatial-Semantic Road Map Construction. *Trans. GIS* **2022**, *26*, 735–754. [CrossRef]
9. Wu, T.; Zhu, Y.; Xiang, L.; Qin, J.; Wan, Y. A Topological Model of Trajectories with Road Network Space. *Trans. GIS* **2022**, *26*, 1847–1878. [CrossRef]
10. Cao, C.; Sun, Y. Automatic Road Centerline Extraction from Imagery Using Road GPS Data. *Remote Sens.* **2014**, *6*, 9014–9033. [CrossRef]
11. Yang, W.; Ai, T.; Lu, W. A Method for Extracting Road Boundary Information from Crowdsourcing Vehicle GPS Trajectories. *Sensors* **2018**, *18*, 1261. [CrossRef]
12. Tang, L.; Ren, C.; Liu, Z.; Li, Q. A Road Map Refinement Method Using Delaunay Triangulation for Big Trace Data. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 45. [CrossRef]
13. Zhou, B.; Zheng, T.; Huang, J.; Zhang, Y.; Tu, W.; Li, Q.; Deng, M. A Pedestrian Network Construction System Based on Crowdsourced Walking Trajectories. *IEEE Internet Things J.* **2021**, *8*, 7203–7213. [CrossRef]
14. Dey, T.K.; Wang, J.; Wang, Y. Improved Road Network Reconstruction Using Discrete Morse Theory. In *Proceedings of the 25th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, Redondo Beach, CA, USA, 7 November 2017*; ACM: New York, NY, USA, 2017; pp. 1–4.
15. Guo, Y.; Bardera, A.; Fort, M.; Silveira, R.I. A Scalable Method to Construct Compact Road Networks from GPS Trajectories. *Int. J. Geogr. Inf. Sci.* **2021**, *35*, 1309–1345. [CrossRef]
16. Yang, X.; Tang, L.; Ren, C.; Chen, Y.; Xie, Z.; Li, Q. Pedestrian Network Generation Based on Crowdsourced Tracking Data. *Int. J. Geogr. Inf. Sci.* **2020**, *34*, 1051–1074. [CrossRef]
17. Yang, L.; Ai, M.; Kwan, M.-P.; Zuo, Z.; Zhang, Y.; Zhou, S.; Luo, S.; Chen, Y. Generation of Intra-Community Roads Based on Human-Flow Modeling (HFM). *Int. J. Geogr. Inf. Sci.* **2024**, *38*, 1256–1290. [CrossRef]
18. Zhang, Y.; Zhang, Z.; Huang, J.; She, T.; Deng, M.; Fan, H.; Xu, P.; Deng, X. A Hybrid Method to Incrementally Extract Road Networks Using Spatio-Temporal Trajectory Data. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 186. [CrossRef]
19. Tang, J.; Xia, H.; Peng, J.; Hu, Z.; Ding, J.; Zhang, Y. Outdoor Hiking Navigation Road Network Map Construction Using Crowd-Source Trajectory Data. *J. Geo-Inf. Sci.* **2025**, *27*, 151–166. [CrossRef]
20. Tang, J.; Zhang, T.; Ding, J.; Tao, K.; Yang, C.; Xiang, J.; Ning, X. Three-Dimensional Outdoor Pedestrian Road Network Map Construction Based on Crowdsourced Trajectory Data. *ISPRS Int. J. Geo-Inf.* **2025**, *14*, 175. [CrossRef]
21. Dal Poz, A.P.; Morceli, B.M. Urban and Rural Road Extraction from Smartphone-Based GPS Trajectories. In *Proceedings of the IGARSS 2024-2024 IEEE International Geoscience and Remote Sensing Symposium*; IEEE: New York, NY, USA, 2024; pp. 8105–8108.
22. Zhang, L.; Thiemann, F.; Sester, M. Integration of GPS Traces with Road Map. In *Proceedings of the Third International Workshop on Computational Transportation Science, San Jose, CA, USA, 2 November 2010*; Association for Computing Machinery: New York, NY, USA, 2010; pp. 17–22.
23. Liu, X.; Zhu, Y.; Wang, Y.; Forman, G.; Ni, L.M.; Fang, Y.; Li, M. *Road Recognition Using Coarse-Grained Vehicular Traces. Technical Report HPL-2012-26*; HP Laboratories: Palo Alto, CA, USA, 2012. Available online: https://hdl.handle.net/1783.1/52908 (accessed on 24 March 2026).

24. Kasemsuppakorn, P.; Karimi, H.A. A Pedestrian Network Construction Algorithm Based on Multiple GPS Traces. *Transp. Res. Part C Emerg. Technol.* **2013**, *26*, 285–300. [CrossRef]
25. Wu, T.; Xiang, L.; Gong, J. Updating Road Networks by Local Renewal from GPS Trajectories. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 163. [CrossRef]
26. Zhang, Y.; Liu, J.; Qian, X.; Qiu, A.; Zhang, F. An Automatic Road Network Construction Method Using Massive GPS Trajectory Data. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 400. [CrossRef]
27. Stanojevic, R.; Abbar, S.; Thirumuruganathan, S.; Chawla, S.; Filali, F.; Aleimat, A. Robust Road Map Inference through Network Alignment of Trajectories. In *Proceedings of the 2018 SIAM International Conference on Data Mining, SDM 2018*; SIAM: Philadelphia, PA, USA, 2018; pp. 135–143.
28. Xie, X.; Ou, G. Pedestrian Network Information Extraction Based on VGI. *Geomatica* **2018**, *72*, 85–99. [CrossRef]
29. Buchin, K.; Buchin, M.; Gudmundsson, J.; Hendriks, J.; Hosseini Sereshgi, E.; Silveira, R.I.; Sleijster, J.; Staals, F.; Wenk, C. Roadster: Improved Algorithms for Subtrajectory Clustering and Map Construction. *Comput. Geosci.* **2025**, *196*, 105845. [CrossRef]
30. Sultan, J.; Ben-Haim, G.; Haunert, J.-H.; Dalyot, S. Extracting Spatial Patterns in Bicycle Routes from Crowdsourced Data. *Trans. GIS* **2017**, *21*, 1321–1340. [CrossRef]
31. Karagiorgou, S.; Pfoser, D. On Vehicle Tracking Data-Based Road Network Generation. In *Proceedings of the Proceedings of the 20th International Conference on Advances in Geographic Information Systems, Redondo Beach, CA, USA, 6 November 2012*; ACM: New York, NY, USA, 2012; pp. 89–98.
32. Pu, M.; Mao, J.; Du, Y.; Shen, Y.; Jin, C. Road Intersection Detection Based on Direction Ratio Statistics Analysis. In *Proceedings of the 2019 20th IEEE International Conference on Mobile Data Management (MDM)*; IEEE: New York, NY, USA, 2019; pp. 288–297.
33. Yuan, M.; Yue, P.; Yang, C.; Li, J.; Yan, K.; Cai, C.; Wan, C. Generating Lane-Level Road Networks from High-Precision Trajectory Data with Lane-Changing Behavior Analysis. *Int. J. Geogr. Inf. Sci.* **2024**, *38*, 243–273. [CrossRef]
34. Shen, W.; Wu, W.; Mao, J.; Chen, J.; Cao, S.; Zhao, L.; Zhou, A.; Zhou, L. SAMI: A Shape-Aware Cycling Map Inference Framework for Designated Driving Service. In *Proceedings of the 2023 IEEE 39th International Conference on Data Engineering (ICDE)*; IEEE: New York, NY, USA, 2023; pp. 3269–3281.
35. Lyu, H.; Pfoser, D.; Sheng, Y. Movement-Aware Map Construction. *Int. J. Geogr. Inf. Sci.* **2021**, *35*, 1065–1093. [CrossRef]
36. Wang, J.; Rui, X.; Song, X.; Tan, X.; Wang, C.; Raghavan, V. A Novel Approach for Generating Routable Road Maps from Vehicle GPS Traces. *Int. J. Geogr. Inf. Sci.* **2015**, *29*, 69–91. [CrossRef]
37. Jiao, F.; Xiang, L.; Deng, Y. Automatic Extraction of Road Interchange Networks from Crowdsourced Trajectory Data: A Forward and Reverse Tracking Approach. *ISPRS Int. J. Geo-Inf.* **2025**, *14*, 234. [CrossRef]
38. Jiang, Y.; Li, X.; Li, X.; Sun, J. Geometrical Characteristics Extraction and Accuracy Analysis of Road Network Based on Vehicle Trajectory Data. *J. Geo-inf. Sci.* **2012**, *14*, 165–170. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.