Scholar Publishing Group
Machine Learning Theory and Practice
https://doi.org/10.38007/ML.2022.030309
ISSN 2790-0983 Vol. 3, Issue 3: 69-77

Pattern Recognition Method of Urban Road Network Based on Support Vector Machine Algorithm

Feng Liu[1, 2*]

[1]*Philippine Christian University, Philippine*

[2]*Jining Normal University, Jining, China*

415592180@qq.com

[*]*corresponding author*

Keywords: Support Vector Machine, Urban Road, Network Mode, Distribution Mode

Abstract: As an important material carrier of urban space, road network is considered as the "fingerprint" of a city. Its model reflects the modernization process of a city. Extracting and analyzing the structural model of urban road network is helpful to understand urban development and build a better urban form. In order to solve the shortcomings of the existing research on pattern recognition methods of urban road network, this paper discusses the pattern classification of urban road network, the pattern recognition of road network distribution and the support vector machine, discusses the selection and establishment of road network samples and the recognition parameters of road network patterns, and discusses the design of the pattern recognition model of urban road network. Finally, the recognition model designed in this paper is compared with CNN, RF and GBDT. The experimental results show that the recall rate, accuracy rate and recall rate of SVM urban road network pattern recognition reach 97.8%, which is superior to the other three recognition methods. Therefore, it is verified that the pattern recognition method of urban road network based on support vector machine algorithm has high practical value.

1. Introduction

Road network is one of the most basic elements in maps and geospatial data, and it is the basis for expressing and analyzing urban spatial activities. Pattern recognition of map elements can automatically mine hidden knowledge, relationships or internal patterns from vector maps, which provides powerful support for cartographic generalization, construction and updating of multi-scale spatial databases, and has a certain impact on applications such as cyberspace analysis.

Nowadays, more and more scholars have done a lot of research on pattern recognition methods of urban road network through various technologies and system tools, and have also made certain research achievements through practical research. Bohua aims to directly extract the local road network topology in the complex urban environment from the bird's eye view (BEV). The road topology is represented by a set of directed lane curves and their interactions, which are captured using their intersections. In order to capture topology better, the concept of minimum cycle and its coverage is introduced. The minimum loop is the minimum loop formed by a directed curve segment (between two intersections). The cover is a group of curves whose line segments participate in forming the minimum cycle. First, it is shown that the cover is sufficient to uniquely represent the road topology. Then, the covering is used to supervise the depth neural network, as well as the lane curve supervision. These studies predict road topology from a single input image. The results of NuScenes and Argverse benchmark tests are significantly better than those of baseline tests [1]. In order to solve the problems in the existing road network spatial pattern recognition methods, a new urban center recognition method based on road line kernel density is proposed. First, the continuous density surface of the road is generated by kernel density estimation. Then, according to the statistics of density values, contour lines are used to depict the city center. These centres fall into three categories, namely global centres, local centres and fake centres. The false center is removed according to the region threshold and spatial relationship. Finally, the global centers and local centers are classified according to the centrality index [2]. A spatial and interactive spatial graph inference (SPIN) module is proposed. When it is inserted into ConvNet, the module infers the graphics constructed from the space projected by the feature map and the interactive area. Spatial reasoning extracts the correlation between different spatial regions and other contextual information. SPIN extracts long-distance dependencies between road sections and effectively describes roads from other semantics. A SPIN pyramid is also introduced, which performs SPIN graph inference across multiple scales to extract multi-scale features. A road segmentation network based on stacked hourglass modules and SPIN pyramid is proposed. Compared with existing methods, this network has better performance [3]. Although the existing research on pattern recognition methods of urban road network is very rich, there are still many problems in its practical application.

This paper combines the theories and methods related to road network pattern and support vector machine algorithm, conducts in-depth research on urban road network pattern on the map, uses support vector machine algorithm to identify grid pattern in road network, and enriches the theories and methods of spatial data mining for cartographic generalization and multi-scale expression. In the part of urban road network pattern recognition, this paper first analyzes some existing road network recognition algorithms, and then makes an in-depth study of the principle of SVM classification. While solving the problem of SVM multi classification, the concept of core similarity merging is proposed, and this classification algorithm is applied to urban road network pattern classification, which greatly reduces the number of SVM classifiers, and thus improves the efficiency of license plate recognition.

2. Design and Research of Urban Road Network Pattern Recognition Method Based on Support Vector Machine Algorithm

2.1. Identification of Urban Road Network

(1) Type of urban road network mode

1) Geometry mode

It refers to that part or all of the road network is geometrically close to a certain shape in human perception, such as linear mode, grid mode, ring mode, etc. [4]. For example, linear mode refers to the mode formed by connecting road segments as long as possible while maintaining smoothness. Grid mode refers to the road network mode formed by the intersection of two groups of approximately parallel roads. The ring pattern refers to a nearly circular road pattern around the city center [5].

2) Functional structure mode

The functional structure mode refers to the parts of the road network that play an important role in transportation and urban form, such as overpasses, multi lane roads, etc. They bear the main traffic flow of the city, connect different functional areas of the city, and generally have a higher road grade [6]. The recognition of road functional structure pattern is of great significance for road mapping synthesis, road network updating, etc. [7].

3) Topology mode

It refers to the topological network analysis based on road composition, from which the statistical characteristics of various topological parameters are extracted and the laws and patterns that the characteristics conform to are summarized, such as degree distribution, small world effect, degree correlation, hierarchical organization, etc. [8].

4) Attachment mode

It refers to the mode reflected by the phenomenon that depends on the road, such as traffic flow, infrastructure distribution, etc; Derived mode refers to a high-order mode that can be extracted from the spatial distribution of roads, such as urban center, sight line, etc. [9].

(2) Pattern recognition of road network distribution

Road network spatial distribution pattern recognition is a process of extracting road network structure patterns by measuring one or more attributes of the road network (such as length, direction, connectivity, density, etc.) in combination with statistical classification, syntax model or machine learning methods [10]. In addition to the roads and road nodes themselves, the following data models are also required for the identification of road network distribution patterns [11].

1) Road mesh

A road mesh is a closed polygon enclosed by road lines. The road mesh is a face object constructed artificially according to the structure of the road network, not a real geographical entity [12]. According to the complexity, the road mesh can be divided into basic mesh and composite mesh. The road mesh can be used as the basic unit of circular pattern and grid pattern recognition. Different types of pattern recognition adopt different description parameters. For example, circular pattern can be measured by roundness index and shape angle parameter of mesh, while grid pattern is usually described by rectangle degree, size and shape parameter of mesh [13].

2) Road network diagram

Road network graph is a network model that describes the topological structure of road network by graph theory. It is composed of many nodes and their relationships. The topology of the network is reflected in the correlation and order between nodes [14]. Road network graph is the basic data model of road network distribution pattern recognition: the original graph is used to measure the geometric characteristics of adjacent roads, such as road length, included angle and turn; The dual graph is used to describe the topological importance of roads, such as network connectivity, proximity centrality and intermediary centrality [15].

2.2. Support Vector Machines

(1) Statistical learning

Statistical learning variables: v and v input a dependency relationship, that is, in a set of sample probability $W(u,v)$, statistical learning problem m identification samples $(u_1,v_1),(u_2,v_2),...(u_m,v_m)$, and in a group of functions $\{w(u,q)\}$, find an optimal function $w(u,q_0)$ to pre identify samples.

$$G(q_0) = \gamma F(v, w(u, q_0)) aW(u, v) \tag{1}$$

Minimum, where $\{w(u,q)\}$ is called evaluation function set and $q \in \Omega$ is generalized sample parameter, so $\{w(u,q)\}$ can represent multiple sample data. Data identification expectation definition:

$$G_{emp}(q_0^{'}) = \frac{1}{m}\sum_{i=1}^{m} F(v, w(u_i, q)) \tag{2}$$

Where, $w_0^{'}$ is the generalized sample vector that minimizes equation (2). Because equation (2) identifies the parameters defined by sample data [16].

(1) Structural risk minimization principle

Formula (3) is expressed by the bound measuring the generalization ability of the classifier.

$$G(\beta) \le G_{emp}(\beta) + \varphi(\frac{k}{i}, \frac{\ln(\mu)}{i}) \tag{3}$$

Where, $G(\beta)$ is the actual recognition result, $G_{emp}(\beta)$ is the empirical recognition result, k is the VC dimension, i is the number of training set samples, and $1-\mu$ is the confidence level.

3. Research on Pattern Recognition Method of Urban Road Network Based on Support Vector Machine Algorithm

3.1. Selection and Establishment of Road Network Samples

In this paper, the 1:10000 scale road network data of a city is selected as the experimental data to ensure that the selected samples include data from different geographical regions such as cities, rural areas and suburbs. Samples are divided manually by participants. Participants will manually cut out samples of grid mode, irregular mode and radiation mode. During cutting, due to insufficient sample data, there is regional overlap between samples. At the same time, because the number of samples of radiation mode is too small, this paper adds a batch of radiation samples by manual means (moving nodes, moving edges, adding roads, deleting roads) [17]. The original data and the data of linear division are shown in Table 1.

Table 1. Data samples

Project	Numerical value
Total number of roads	14328
Grid Mode Road	7654
Irregular mode	5231
Network mode linear unit	7116
Irregular mode linear unit	6541

3.2. Cognitive Parameters of Road Network Model

Considering that this paper uses the linear subdivision model of road network to model the road network, we need to use visual variables to describe the visual features of linear units, mainly considering the location, direction and connection features. However, as the road network is a complex spatial network, additional attention should be paid to its connection characteristics. It should be noted that the semantic attributes of roads are also an important part of Gestalt's theory, which will also have an impact on spatial cognition. The semantic attributes of roads mainly refer to the grades and categories of roads. This paper summarizes the cognitive parameters of the road network model from three aspects: location, perspective and connection characteristics, as shown in Table 2 [18].

Table 2. Cognitive parameters of road network

Variable	Index	Computing method	Meaning
Positional feature	Center distance	$\sqrt{(u_x - u_y)^2 + (v_x - v_y)^2}$	Midpoint distance of linear unit
	Center direction	α	The direction of the connecting line between the midpoint and the central point of the linear element
Angular feature	Line element direction	$\tan^{-1}\dfrac{v_{x1} - v_{y2}}{u_{x1} - u_{y2}}$	Direction of linear unit
Connection characteristics	Degree of node	-	Number of sides connected by linear cells

4. Research on the Application of Support Vector Machine Algorithm to Urban Road Network Pattern Recognition

4.1. Pattern Recognition Process of Urban Road Network Based on Support Vector Machine Algorithm

Figure 1 shows the process of road network pattern recognition system based on support vector machine classification. It mainly consists of five parts: road network modeling, feature extraction, road network segmentation, road network pattern recognition and classified output.

(1) Road network diagram modeling. In this part, after cleaning the vector data of the road network, the road segment is divided into several linear elements by linear subdivision, and then the linear elements of the road network are taken as graph nodes, the connection relationship between the elements is taken as the edge of the graph, and the road network is modeled as an undirected plan.

(2) Feature extraction. This part needs to extract the graph node features of the road network, including location features, angle features, and connection features.

(3) In the road network segmentation part, the vertical and horizontal projection method is mainly used to segment the road network in combination with the prior knowledge of the road network. In the road network recognition part, it is actually the road network classification stage. Correct classification of the divided road network can produce correct recognition results.

(4) In the part of road network recognition, the SVM algorithm of kernel similarity merging proposed in this paper is mainly used to classify and recognize the segmented road network.

(4) Classified output. The validation samples are input into the trained support vector machine model, and the output samples are classified into three categories: radial mode, grid mode and irregular mode.

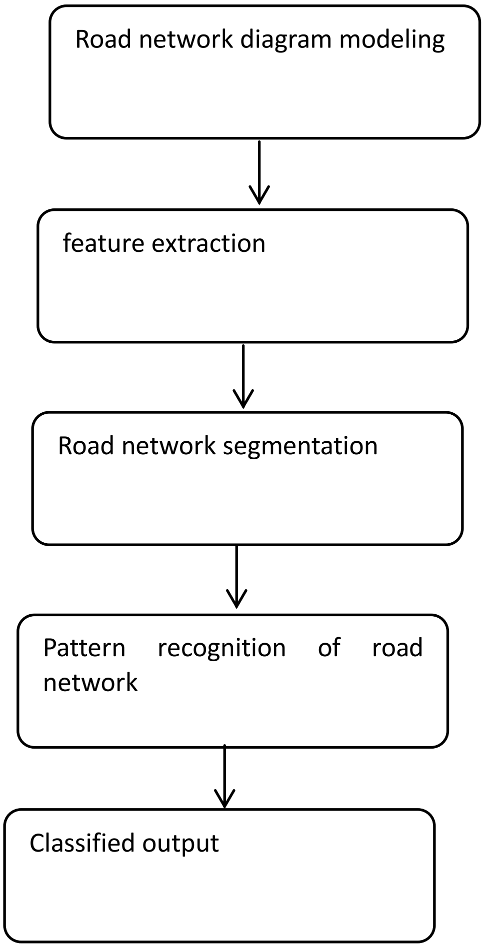

Figure 1. Process of urban road network pattern recognition system

4.2. Application of Support Vector Machine Algorithm to Urban Road Network Pattern Recognition

In order to compare the recognition effect of the support vector machine algorithm with other methods, this paper selects three classic classification methods, CNN, RF and GBDT, as reference, selects the 1:1000 scale road network data of a city as experimental data, and similarly selects 20% samples as training training. The experimental results are shown in Table 3.

Table 3. Algorithm recognition performance

Algorithm	SVM	RF	GBDT	CNN
Recall ratio	93.1%	78.5%	76.5%	85.4%
Accuracy	96.3%	79.4%	78.2%	87.8%
Call back rate	97.8%	81.2%	77.3%	86.7%

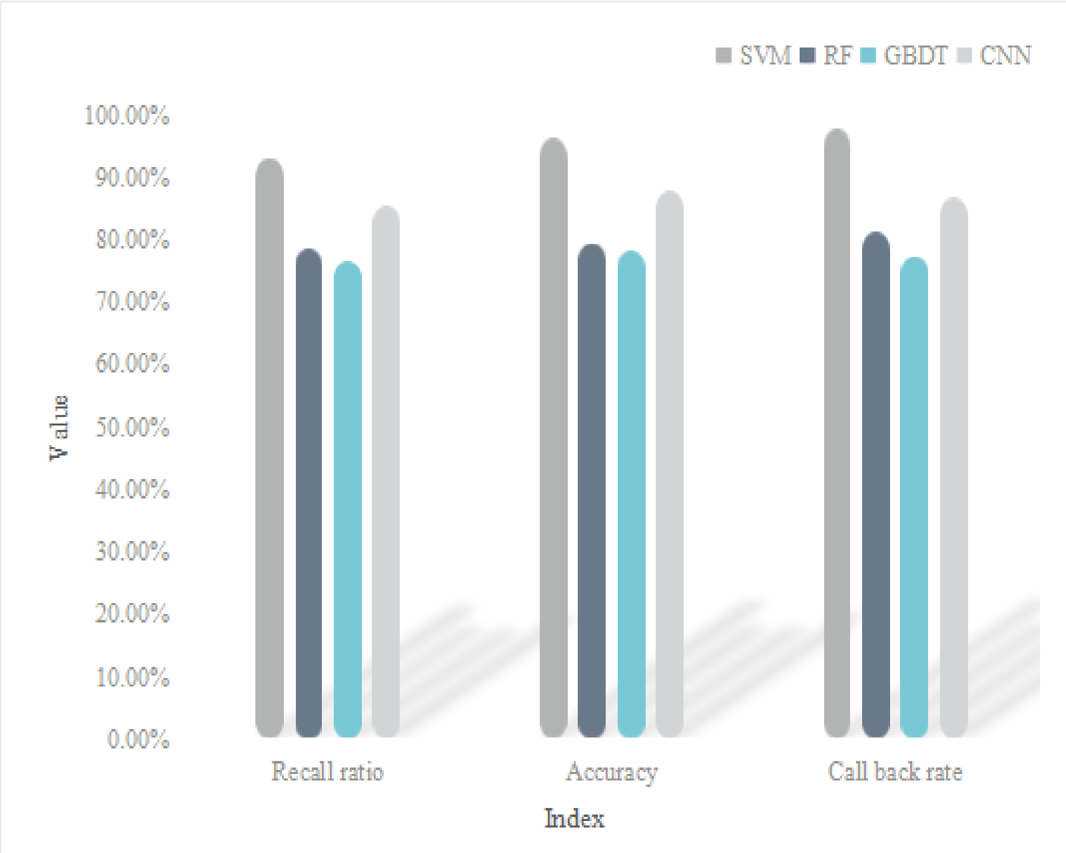

Figure 2. Comparison of algorithm recognition performance

As can be seen from the data in Figure 2, except for GBDT, the other three methods have good phenotypes. Among them, the SVM algorithm and CCN model proposed in this paper have better classification effects than other methods. GBDT performs very poorly in the case of uneven sample distribution, which is significantly worse than other methods. Compared with other machine learning methods, the recall rate of SVM algorithm in urban road network pattern recognition is 93.1%, while the recall rates of RF, GBDT and CNN are 78.5%, 76.5% and 85.4% respectively. The accuracy of SVM algorithm in pattern recognition of urban road network is up to 96.3%. The recall rates of RF, GBDT and CNN reached 79.4%, 78.2% and 87.8% respectively. The recall rate of SVM algorithm in pattern recognition of urban road network is as high as 97.8%. The recall rates of RF, GBDT and CNN reached 81.2%, 77.3% and 86.7% respectively.

5. Conclusion

This paper summarizes the cognitive parameters of the road network model from three aspects: location, angle and connection characteristics, and uses the corresponding parameters to describe these characteristics. The pattern recognition model of urban road network based on support vector machine algorithm is designed. The support vector machine algorithm is used to recognize the grid pattern, and experiments are carried out on the grid pattern as the main pattern data of urban roads. Support vector machine algorithm is more comprehensive in grid pattern recognition, and the recognition results have higher accuracy. Through comparative experiments, it is found that the effect of support vector machine in grid pattern recognition is better than that of RF, GBDT and CNN recognition models, which proves the effectiveness of the algorithm in this paper.

Funding

This article is not supported by any foundation.

Data Availability

Data sharing is not applicable to this article as no new data were created or analysed in this study.

Conflict of Interest

The author states that this article has no conflict of interest.

References

[1] Bohua, HUANG, Wei, et al. Hierarchical Area Partitioning Method of Urban Road Networks Matching. Journal of Geodesy and Geoinformation Science, 2019, v.2(03):57-69.

[2] Williams, Julie, C, et al. Soluble CD14, CD163, and CD27 biomarkers distinguish ART‐suppressed youth living with HIV from healthy controls. Journal of Leukocyte Biology: An Official Publication of the Reticuloendothelial Society, 2018, 103(4):671-680. https://doi.org/10.1002/JLB.3A0717-294RR

[3] Roy J, Han J, Weeks A, et al. P.138 A small RNA signature from extracellular vesicles in patient plasma correlates with recurrence or progression of high-grade gliomas. Canadian Journal of Neurological Sciences / Journal Canadien des Sciences Neurologiques, 2022, 49(s1):S43-S43. https://doi.org/10.1017/cjn.2022.222

[4] Nino, Kopaleishvili. Perception of Myth in Print and Online Media. Language and Semiotic Studies, 2020, v.6(01):35-45.

[5] Chiu A P. When Rifles Need a Makeover: SARCO Inc.'s Reproduction G43 Magazine. Small arms review, 2019, 23(8):44-47.

[6] Ewelina, Wawryk-Gawda, Beata, et al. Dietary supplements – consumer assessment based on questionnaire survey. Przeglad epidemiologiczny, 2018, 72(1):111-120.

[7] Tsertsvadze G. The Probabilistic Model of Canonically Conjugate Fuzzy Subsets. Bulletin of the Georgian Academy of Sciences, 2018, 12(4):53-58.

[8] Gitonga, Njeru. Roads of Success Ethiopia's new extensive road network to boost industrialization development. ChinAfrica, 2019, v.11(03):34-35.

[9] Memon I. Distance and clustering‐based energy‐efficient pseudonyms changing strategy over road network. International Journal of Communication Systems, 2018, 31(11):e3704.1-e3704.22. https://doi.org/10.1002/dac.3704

[10] Dhanpat J, Higginson A, Brooks K. Estimation of the Effect of Bio-Admixtures on Concrete Workability Using Linear Regression and Support Vector Machines. IFAC-PapersOnLine, 2021, 54(21):133-138. https://doi.org/10.1016/j.ifacol.2021.12.023

[11] Kahl M, Kroll A. Extending Regularized Least Squares Support Vector Machines for Order Selection of Dynamical Takagi-Sugeno Models. IFAC-PapersOnLine, 2020, 53(2):1182-1187. https://doi.org/10.1016/j.ifacol.2020.12.1331

[12] Kantavat P, Kijsirikul B, Songsiri P, et al. Efficient Decision Trees for Multi-Class Support Vector Machines Using Entropy and Generalization Error Estimation. International Journal of Applied Mathematics & Computer Science, 2018, 28(4):705-717. https://doi.org/10.2478/amcs-2018-0054

[13] Mathew J, Pang C K, Luo M, et al. Classification of Imbalanced Data by Oversampling in Kernel Space of Support Vector Machines. Neural Networks and Learning Systems, IEEE Transactions on, 2018, 29(9):4065-4076. https://doi.org/10.1109/TNNLS.2017.2751612

[14] Kim, Sungdo, Mun, et al. Data depth based support vector machines for predicting corporate bankruptcy. Applied Intelligence: The International Journal of Artificial Intelligence, Neural Networks, and Complex Problem-Solving Technologies, 2018, 48(3):791-804. https://doi.org/10.1007/s10489-017-1011-3

[15] Mohsen, Farhang, Ali, et al. Automatic modulation recognition for DVB-S2 using pairwise support vector machines. International journal of autonomous and adaptive communications systems: IJAACS, 2018, 11(2):144-153. https://doi.org/10.1504/IJAACS.2018.092020

[16] Silvina, Niell, Florencia, et al. Beehives biomonitor pesticides in agroecosystems: Simple chemical and biological indicators evaluation using Support Vector Machines (SVM). Ecological indicators: Integrating, monitoring, assessment and management, 2018, 91(Aug.):149-154. https://doi.org/10.1016/j.ecolind.2018.03.028

[17] Hossain M Z, Gedeon T, Sankaranarayana R. Using Temporal Features of Observers' Physiological Measures to Distinguish Between Genuine and Fake Smiles. IEEE Transactions on Affective Computing, 2020, 11(1):163-173. https://doi.org/10.1109/TAFFC.2018.2878029

[18] Williams, Julie, C, et al. Soluble CD14, CD163, and CD27 biomarkers distinguish ART‐suppressed youth living with HIV from healthy controls. Journal of Leukocyte Biology: An Official Publication of the Reticuloendothelial Society, 2018, 103(4):671-680. https://doi.org/10.1002/JLB.3A0717-294RR