Analysis of Irregular Spatial Data with Machine Learning: Classification of Building Patterns with a Graph Convolutional Neural Network

Xiongfeng Yan
School of Resource and Environmental Sciences, Wuhan University, Wuhan, China
xiongfeng.yan@whu.edu.cn
https://orcid.org/0000-0003-4748-464X

Tinghua Ai
School of Resource and Environmental Sciences, Wuhan University, Wuhan, China
tinghuaai@whu.edu.cn
https://orcid.org/0000-0002-6581-9872

Abstract

Machine learning methods such as Convolutional Neural Network (CNN) are becoming an integral part of scientific research in many disciplines, the analysis of spatial data often failed to these powerful methods because of its irregularity. By using the graph Fourier transform and convolution theorem, we try to convert the convolution operation into a point-wise product in Fourier domain and build a learning architecture of graph CNN for the classification of building patterns. Experiments showed that this method has achieved outstanding results in identifying regular and irregular patterns, and has significantly improved in comparing with other methods.

2012 ACM Subject Classification Computing methodologies → Machine learning

Keywords and phrases Building pattern, Graph CNN, Spatial analysis, Machine learning

Digital Object Identifier 10.4230/LIPIcs.GIScience.2018.69

Category Short Paper

Acknowledgements The first author would like to express thankfulness to Shuaixiong Rao for providing the computing resources, which is very helpful to this paper.

1 Introduction

With the improvement of computing power and the advent of data era, machine learning methods are becoming an integral part of scientific research in many disciplines. As a supervised learning method, CNN has excellent performance in many fields, such as computer vision and natural language processing. These successes are mainly attributed to its two important properties: first, inspired by neuronal processing, CNN focuses on local structures (Local Receptive Fields), and combines them into a whole, which can be applied to classification or identification. Second, local structure of different regions can be detected by the same convolution kernel, that is, weights sharing. The former accords with the decomposability of object and hierarchy of cognition, the latter reduces complexity and improves learnability.

However, it should be noted that both the local connection and weight sharing properties require that the local structure of data is fixed, normative, and can be clearly defined. For example, the images in visual analysis are organized in a grid of pixels as a processing unit, and sentences in natural language processing are organized in a linear arrangement of words as a processing unit. But, for most of spatial graphics data in GIS fields, the arrangement, combination, or connection between objects may be more diversified, and it is often difficult to satisfy this requirements of specification, for example, the group relationships between plane buildings. Therefore, this kinds of data cannot directly use these powerful learning methods for pattern recognition and knowledge discovery.

Although spatial data is irregular and cannot be organized according to grid or linear structure, it is still possible to represent by graph structure. The graph cannot define a convolution operation in the vertex domain directly, but in virtue of graph Fourier transform and convolution theorem, the operation can be transformed into a point-wise product in the Fourier domain, which is similar to the transformation of spatial domain convolution into frequency domain convolution in image processing. Based on this idea, we propose a graph CNN for identifying patterns and mining knowledge of irregular spatial data.

In this study, we focus on using machine learning to solve the problem of building group pattern recognition, which can be used in many fields, such as urban morphology and environmental analysis. Although the related researches have been carried out for decades, there are still some problems such as incomplete taxonomy and inconsistent recognition rules. The introduction of machine learning method is an effective attempt and supplement to solve such classical problems in spatial analysis. In the following sections, we will describe detailed methods, then conduct experiments and compare with other similar methods, and finally discuss and conclude this study.

2 Methodology

2.1 Definition of Building Pattern Classification

Building patterns refer to visually salient structures exhibited collectively by a group of buildings[4]. Traditional patterns detection methods are to predefine some specific perceptual rules according to the characteristics of azimuth angle, direction difference and proximity, and then to inquire whether there is a local group that satisfy such rules[3][8][10]. But these rules are difficult to formalize and too rigid, which inevitably lead to an unsatisfactory result[6].

Similar to image processing, determining which pattern a building group visually belongs to is essentially an issue of classification. A building group is an analogy to a picture, and each building is analogy to the pixel and its shape features are analogy to color channels, in spite of the relationship between them is an irregular graph structure, not a fixed grid.

2.1.1 Features of single building

Single building has spatial features that describe its graphical structures and semantic features that describe its attributes, which in combination can effectively reflect its basic form. For the description of these features, dozens of indices have been proposed[8]. In this study, we mainly consider area A_b, main direction α, and three shape indices including length-width ratio R_{lw}, area ratio R_A, and compactness C. These indices illustrated in figure 1.

2.1.2 Graph representation of building group

Graph is an ideal tool to describe the relationships between multiple objects. Delaunay triangulation (DT) and Minimum Spanning Tree (MST) are the two most commonly used ways due to they can take spatial constraints and other contextual constraints into consideration, such as proximity.

Regardless of whether DT or MST, they can be defined as $\mathcal{G}= (\mathcal{V}, \mathcal{E}, \mathcal{W})$, where $\mathcal{V}$ and $\mathcal{E}$ is a finite set of $|\mathcal{V}| = n$ vertices and edges, respectively, $\mathcal{W} \in \mathbb{R}^{n\times n}$ is a adjacency matrix

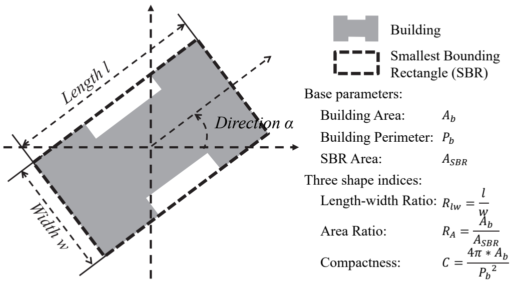

Figure 1 Input feature indices.

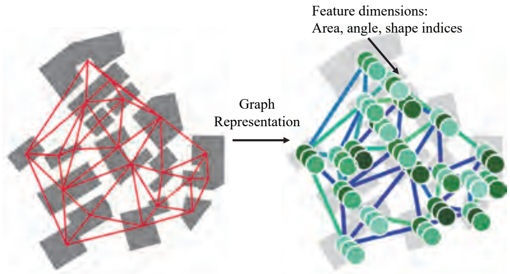

Figure 2 Graph construction.

encoding the weight between two vertices, and each vertex also contains one or several features, as seen in figure 2.

2.2 Graph Convolutional Neural Network

2.2.1 Graph Fourier transform

Fourier transform is an effective tool in signal analysis and image processing, it decomposes an original function (e.g., a signal or image) into the frequencies that make it up. The process is essentially a linear transformation by using given orthogonal basics $\langle f, e^{i\omega t} \rangle$.

For the graph structure, we utilize the eigenvectors χ_ℓ of Laplacian as the decomposition basics instead of complex exponentials, then define the graph Fourier transform as:

$$\hat{f}\left(\lambda_\ell\right) = \langle f, \chi_\ell \rangle = \int_{n=1}^{N} \chi_\ell^T(n) f(n)\,(1) \tag{1}$$

Where, λ_ℓ is the eigenvalue and the inverse Fourier transform as:

$$f\left(n\right) = \int_{n=1}^{N} \hat{f}\left(\lambda_\ell\right) \chi_\ell(n)\,(2) \tag{2}$$

This definition is precise analogy to the classical one, and it can be interpreted as an expansion of f in terms of the eigenvectors of the Laplacian[5][7].

2.2.2 Convolution on graph

As we cannot conduct the convolution in vertex domain directly, we can try to convert this operation into a point-wise product in Fourier domain by means of graph Fourier transform and convolution theorem, and it can be defined as:

$$\left(f * g\right)\left(n\right) = \int_{n=1}^{N} \hat{f}\left(\lambda_\ell\right) \hat{g}\left(\lambda_\ell\right) \chi_\ell(n)\,(3) \tag{3}$$

Using notation from the matrix theory, the convolution also can be written as:

$$f * g = \mathcal{X}\left(\left(\mathcal{X}^T f\right) \bullet \left(\mathcal{X}^T g\right)\right) = \mathcal{X} diag\left(\hat{g}\left(\lambda_1\right),\ \ldots, \hat{g}\left(\lambda_N\right)\right) \mathcal{X}^T f = \mathcal{X} \hat{g}_\theta\left(\Lambda\right) \mathcal{X}^T f\,(4) \tag{4}$$

Where, $diag\left(\hat{g}\left(\lambda_1\right),\ \ldots, \hat{g}\left(\lambda_N\right)\right)$ can be understood as a set of free parameters θ in the Fourier domain (the Eigenspaces of Laplacian), or a function of the eigenvalues $\hat{g}\left(\Lambda\right)$.

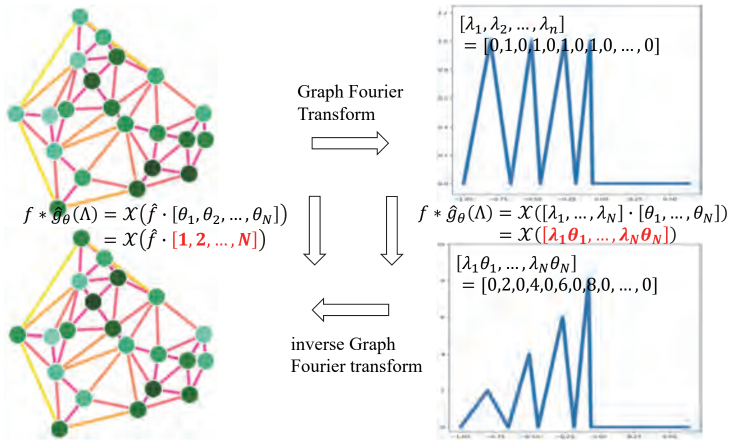

Figure 3 The convolution of a graph f and a kernel of free parameters.

2.2.3 Polynomial approximation for fast localized convolution

The above definition of convolution operation on graph still has two limitations: 1) in each operation, the Eigen decomposition need to be performed, which will bring lots of computational cost; 2) without considering the locality in space, the features of a vertex may be related to global vertices after this operation, it is not consistent with the local connection property of classical CNN[1][2].

In response to these problems, Hammond[5] proposed a fast localized convolution based on low-order polynomial approximation that represent $\hat{g}_\theta(\Lambda)$ as a polynomial function of eigenvalues:

$$\hat{g}_\theta(\Lambda) = \int_{k=0}^{K-1} \theta_k \Lambda^k \, (5) \tag{5}$$

Then, the Equation (4) can be rewritten as:

$$f * g = \mathcal{X}\left(\int_{k=0}^{K-1} \theta_k \Lambda^k\right) \mathcal{X}^T f = \left(\int_{k=0}^{K-1} \theta_k \left(\mathcal{X}\Lambda^k\mathcal{X}^T\right)\right) f = \left(\int_{k=0}^{K-1} \theta_k \mathcal{L}^k\right) f \, (6) \tag{6}$$

As we can see, no need to perform the Eigen decomposition anymore, and the feature values of vertex are related only to its K-order neighboring vertices, which satisfies the locality in space.

2.2.4 Architecture of convolutional neural network on graph

Based on the above-defined graph convolution, we propose a learning architecture of CNN on graph for the classification of building patterns, as seen in figure 4. This architecture includes convolutional, subsampling, and full connected layers, where subsampling layer is optional and the full connected layer is the same as the classical CNN. We input a building group that has already been represented as graph to this architecture, after the steps of feature extraction and classification, we can get the probability that it belongs to each class and choose the class with maximum probability as the final classification result.

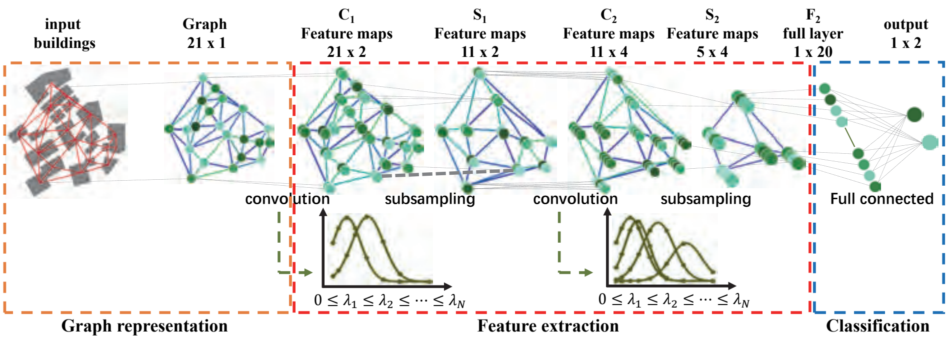

Figure 4 The architecture of convolutional neural network on graph.

Table 1 Accuracies of the proposed method and other methods.

Method	SVM	Random Forest	Graph CNN
Accuracy	90.2%	93.4%	98.04%

3 Experiments

The experimental buildings were extracted from a large-scale 1:2000 topological map of the city of Guangzhou, China. We divided them into separated groups by using road network division and clustering techniques, and each group contains 20-128 buildings. Then, we manually identified the two patterns of regular and irregular from all groups. Each group was estimated by at least three participants to ensure the correctness, and these ambiguous groups were discarded. At last, there are 2647 and 2646 available groups for regular and irregular pattern respectively and contain a total of 318 598 buildings. Each group can serve as a sample for the graph CNN, all samples were split into training, validation and test sets by 6:2:2, and input features of all data were standardized using training set.

We used a shallow graph CNN architecture with four convolutional layers and one full connected layer to test the datasets, each convolutional layer contains 24 third-order polynomial convolution kernels. The more convolutional layers, the more complex the model is and the more samples are required. In addition, regularization and dropout techniques are also used to control the complexity, and their parameters are referenced from empirical values and fine-tuned. The accuracy is 98.04%, which is better than that of SVM[9] and random forest[6] methods, the comparison results are shown in table 1.

The activation of a sample is shown in figure 5 and the input volume stores the graph of building group (left) and the last volume holds the scores for each class (right).

In this model, the order K of the polynomial is one of the important parameters. We tested the values of 1, 2, 3, 4, 5, and 6 respectively, the performances on the validation set are shown in figure 6a. The comparison found that it achieved the best performance when K=3. The larger of K, the more complex of the training and the longer it takes. We further tested the effect of input features of individual building on the classification of group patterns. We tried to train and learn by using 4 features of them or only one, these results are shown in figure 6b and we found that the area was one of the important features and the accuracy could also reach 96.34% when only area feature was used. This may be due to the fact that areas of buildings in a regular pattern are more homogeneous.

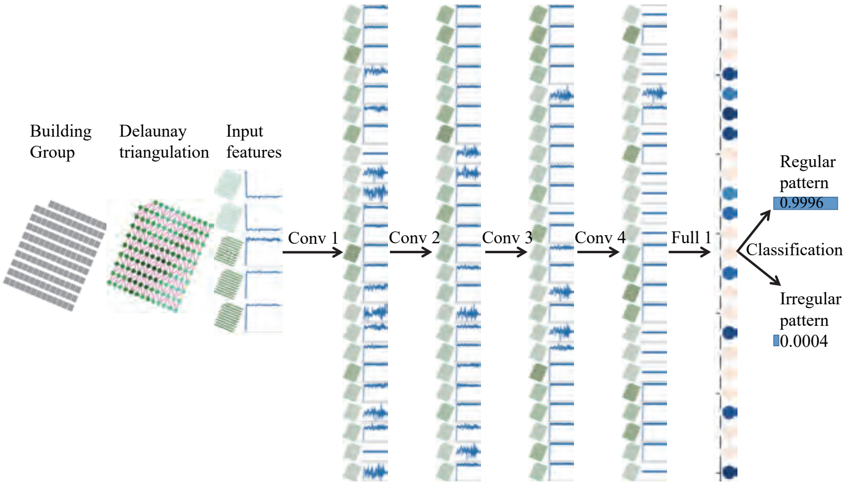

Figure 5 The activations of an example graph CNN architecture.

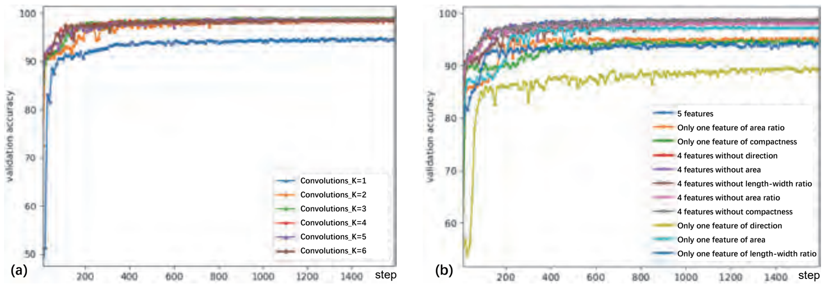

Figure 6 Performances when taking different K values or inputting different features.

4 Discussion and Conclusion

As a classical problem in the analysis of irregular spatial data, the traditional building pattern classification method needs to manually extract features and design rules for specific patterns. In this paper, we propose a graph CNN that represent the building groups by graph and convert the convolution of vertex domain into a point-wise product in Fourier domain. It can directly extract patterns characteristics based on the training and learning of sample data. Experiments showed that proposed method has achieved outstanding results in identifying regular and irregular patterns, and has significantly improved in comparing with other methods. Meanwhile, it has great potential to extend to other analyses of irregular spatial data, such as classification of road patterns and identification of point clouds.

The difficulties of this method lie in the selection of input features and the training process. We have selected five features for experiments, but there are still many other descriptive indices that can be selected. Determining which indices can better describe building patterns and applying them to training and learning still requires more experiments, and the principal component analysis may be a worthwhile approach to try. The training of graph CNN requires more samples, otherwise it will easily lead to overfitting, especially for deep networks with many convolutional layers. In the follow-up work, Volunteer Geographic Information (VGI) is a desirable and feasible data source.

References

1 Joan Bruna, Wojciech Zaremba, Arthur Szlam, and Yann LeCun. Spectral networks and locally connected networks on graphs. *CoRR*, 2013. URL: `http://arxiv.org/abs/1312.6203`.

2 Michaël Defferrard, Xavier Bresson, and Pierre Vandergheynst. Convolutional neural networks on graphs with fast localized spectral filtering. *CoRR*, 2016. URL: `http://arxiv.org/abs/1606.09375`.

3 Shihong Du, Liqun Luo, Kai Cao, and Mi Shu. Extracting building patterns with multilevel graph partition and building grouping. *ISPRS Journal of Photogrammetry and Remote Sensing*, 122:81–96, 2016. `doi:10.1016/j.isprsjprs.2016.10.001`.

4 Shihong Du, Mi Shu, and Chen Chieh Feng. Representation and discovery of building patterns: a three-level relational approach. *International Journal of Geographical Information Science*, 30(6):1161–1186, 2015. `doi:10.1080/13658816.2015.1108421`.

5 David K. Hammond, Pierre Vandergheynst, and Rémi Gribonval. Wavelets on graphs via spectral graph theory. *Applied and Computational Harmonic Analysis*, 30(2):129–150, 2009. `doi:10.1016/j.acha.2010.04.005`.

6 Xianjin He, Xinchang Zhang, and Qinchuan Xin. Recognition of building group patterns in topographic maps based on graph partitioning and random forest. *ISPRS Journal of Photogrammetry and Remote Sensing*, 136:26–40, 2018. `doi:10.1016/j.isprsjprs.2017.12.001`.

7 David I Shuman, Benjamin Ricaud, and Pierre Vandergheynst. Vertexfrequency analysis on graphs. *Applied and Computational Harmonic Analysis*, 40(2):260–291, 2016. `doi:10.1016/j.acha.2015.02.005`.

8 Zhiwei Wei, Qingsheng Guo, Lin Wang, and Fen Yan. On the spatial distribution of buildings for map generalization. *Cartography and Geographic Information Science*, pages 1–17, 2018. `doi:10.1080/15230406.2018.1433068`.

9 Liqiang Zhang, Hao Deng, Dong Chen, and Zhen Wang. A spatial cognition-based urban building clustering approach and its applications. *International Journal of Geographical Information Science*, 27(4):721–740, 2013. `doi:10.1080/13658816.2012.700518`.

10 Xiang Zhang, Tinghua Ai, Stoter Jantien, Kraak Menno Jan, and Molenaar Martien. Building pattern recognition in topographic data: examples on collinear and curvilinear alignments. *GeoInformatica*, 17(1):1–33, 2013. `doi:10.1007/s10707-011-0146-3`.