International Cartographic Association
Association Cartographique Internationale

MLP Feature Extraction from Coordinates for Building Footprint Simplification using Graph Convolutional Networks

Martin Kada [a*], Julien Baerenzung [a], Lilli Kaufhold [a]

[a] *Institute of Geodesy and Geoinformation Science, Technische Universität Berlin, 10553 Berlin, Germany, martin.kada@tu-berlin.de, baerenzung@tu-berlin.de, lilli.kaufhold@tu-berlin.de*

* Corresponding author

Keywords: building simplification, cartographic generalization, graph convolutional network, feature extraction

Abstract:

Building footprint simplification is crucial for various applications in urban planning, mapping, and navigation. With recent advances in deep neural networks, this task has often been addressed in research in the image domain with convolutional neural networks (CNN) or in the vector domain using graph neural networks (GNN). For the latter, hand-crafted geometric features are typically computed for the points of the building footprint polygon (e.g. line segment lengths, triangle areas, angles, etc.) and used as input to the neural network instead of the point coordinates. However, this does not reflect the notion of deep learning, in which feature extraction is performed by the neural network itself. In this study, we employ the graph convolutional networks (GCNs) approach proposed by Zhou et al. (2023), and add a multilayer perceptron (MLP) module at the beginning of the neural network that performs a feature extraction directly on the point coordinates (see Figure 1). We show that this approach is able to learn the relevant features for the task of cartographic simplification of building footprints, and that the network performs on par with using pre-computed features.

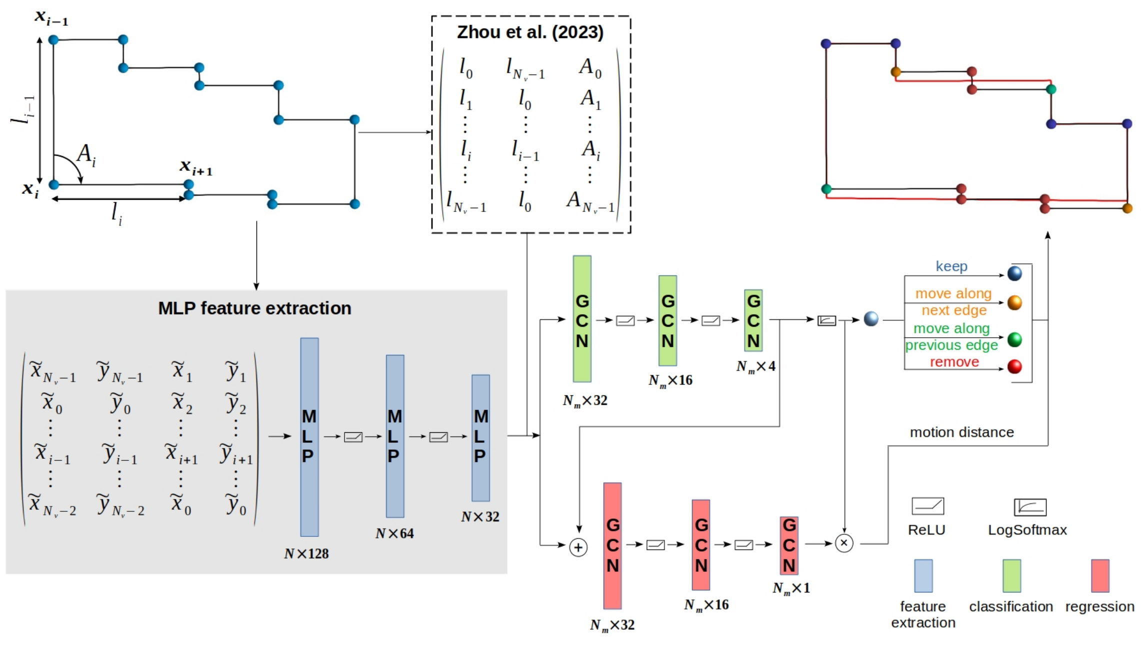

Figure 1. GCN architecture with (adapted) pre-computed features as input (dashed box) following Zhou et al. (2023), and with the proposed MLP based feature extraction module (gray box).

To study the benefits of such an approach we implemented two models: The first one represents a minor adaptation of the original model introduced by Zhou et al. (2023). The second introduces our new parametrization of inputs and the MLP module for feature extraction (see Figure 1). For both models, the primary objective is to determine whether a vertex should be kept, removed or moved, and for the moved vertices by how much they are moved along the two incident edges. Note that contrary to Zhou et al. (2023), we move a vertex either along the preceding or the succeeding edge (and not along both), and therefore have two move classes (instead of one) and one measure of move distance.

Abstracts of the International Cartographic Association, 8, 11, 2024.
2024 ICA Workshop on AI, Geovisualization, and Analytical Reasoning – CartoVis24, 7 September 2024, University of Warsaw, Poland.
https://doi.org/10.5194/ica-abs-8-11-2024 |

In the GCN baseline model adapted from Zhou et al. (2023), the network is supplied with input features derived from the geometric attributes of incident edges. For each vertex, these features encompass three parameters: the lengths of the preceding and succeeding edges, denoted as l_i and l_{i-1}, respectively, and the signed angle between these edges, denoted as A_i. We do not use the angle direction as in the original work.

In contrast, for the GCN model with the MLP feature extraction, input features for each vertex i are directly formulated using its spatial coordinates, $\mathbf{x_i} = (x_i, y_i)$, as well as the spatial coordinates of its preceding and succeeding vertices, denoted as $\mathbf{x_{i\text{-}1}}$ and $\mathbf{x_{i+1}}$, respectively. These coordinates are then centered around $\mathbf{x_i}$, resulting in $\tilde{\mathbf{x}}_{\mathbf{i\text{-}1}} = (x_{i\text{-}1} - x_i, y_{i\text{-}1} - y_i)$, $\tilde{\mathbf{x}}_{\mathbf{i}} = (0,0)$, and $\tilde{\mathbf{x}}_{\mathbf{i+1}} = (x_{i+1} - x_i, y_{i+1} - y_i)$. Notably, as $\tilde{\mathbf{x}}_{\mathbf{i}} = (0,0)$ for each vertex i, we do not regard this quantity as important anymore and it is excluded from the input features, resulting in an input vector $\mathbf{X}$ of dimension $N = N_v \times 4$ for a building footprint containing N_v vertices. The presented input tensor is then not directly fed into the GCN, but first goes through a three layer MLP module.

As in the study of Zhou et al. (2023), we employed the dataset from Feng et al. (2019), specifically focusing on building footprints in the Stuttgart region. We also chose a scale of 1 : 5000 for the source map and 1 : 10000 for the target map. We restricted our analysis to footprints comprising a minimum of 5 vertices, resulting in a dataset containing 25185 footprints. Subsequently, this dataset was partitioned into training, evaluation, and testing sets, in the proportions 8:1:1.

The outcomes of the vertex classification part are presented in Table 1, showcasing the prediction accuracy attained on the test set. Overall, the GCN model with MLP feature extraction demonstrates marginally superior predictive performance, particularly concerning vertices needing displacement.

GCN model with	Keep	Move along preceding edge	Move along succeeding edge	Remove
hand-crafted features	96.5%	76.7%	78.4%	95.3%
MLP features	97%	82.8%	81.7%	96.9%

Table 1. Accuracy of vertex classification predictions.

To evaluate the fidelity of predicted outcomes against actual targets within the test set, two comparison metrics were computed: the Intersection over Union (IoU), expressed as a percentage, and the Hausdorff distance (HD), measured in meters. Statistical summaries of these metrics are provided in Table 2. While the minimum, maximum, and median values of IoU and HD exhibit close proximity between both models, discernible disparities emerge in their mean values. This suggests that, on average, the GCN model with MLP feature extraction generates more precise predictions.

	IoU (%)				HD (m)			
GCN model with	Min	Max	Mean	Median	Min	Max	Mean	Median
hand-crafted features	0.318	1	0.978	0.996	0.003	47.9	0.601	0.058
MLP features	0.367	1	0.984	0.997	0.003	47.9	0.461	0.059

Table 2. Similarity metrics between predicted and real target building footprints.

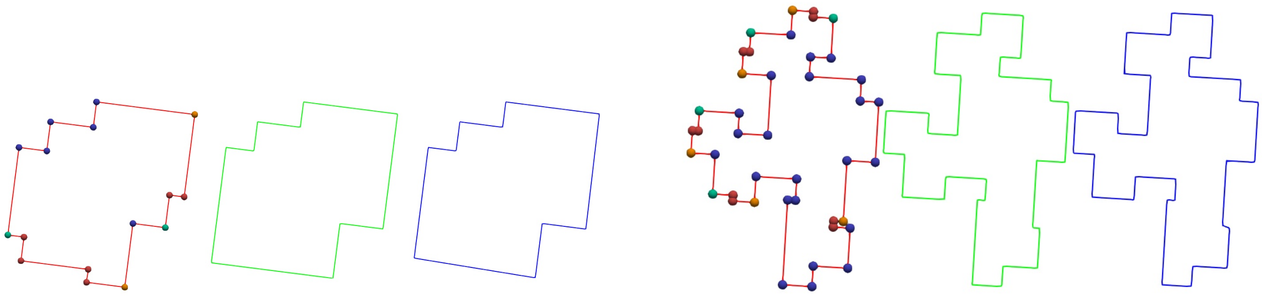

Figure 2. Examples of simplified building footprints using the MLP feature extraction module with input (red), ground truth (green), and predicted target (blue) footprints, respectively.

This study shows that an MLP is capable of extracting features directly from vector coordinates for the simplification of building footprints using a GCN. The findings suggest that the approach can also be used to extract effective features from coordinates for other tasks related to 2D vector geodata using neural networks without relying on hand-crafted features.

References

Feng, Y., Thiemann, F. and Sester, M., 2019. Learning cartographic building generalization with deep convolutional neural networks. *ISPRS International Journal of Geo-Information*.

Zhou, Z., Fu, C. and Weibel, R., 2023. Move and remove: Multi-task learning for building simplification in vector maps with a graph convolutional neural network. *ISPRS Journal of Photogrammetry and Remote Sensing* 202, pp. 205–218.

Abstracts of the International Cartographic Association, 8, 11, 2024.
2024 ICA Workshop on AI, Geovisualization, and Analytical Reasoning – CartoVis24, 7 September 2024, University of Warsaw, Poland.
https://doi.org/10.5194/ica-abs-8-11-2024 |