Graffiti Distribution in Novi Sad: A Comparative Spatial Analysis Using Field Surveys and Google Street View

GEO 885, Group 8

Jing Tang, Yilin Wei

Department of Geography, University of Zurich

Abstract

In South-Eastern and Central Europe, graffiti is a widespread cultural phenomenon. It is often used to express support for football teams, political affiliation and (national) ideology. Currently, research on graffiti in South-Eastern Europe mainly focuses on the domains of sociology and cultural studies, with an absence of systematic studies from linguistic and geographic perspectives. An integration of these two perspectives can prove especially helpful for the study of graffiti, which in most cases uses language to shape the space in which they are located. The linguistic subfield of linguistic landscape, within which graffiti is often linguistically studied, presents severe issues, mainly due to a lack of systematic methods. Interdisciplinary work combining linguistic and geographic methods is needed in order to improve the scientific quality of linguistic landscape research. Hence, this study aims to adopt a linguistic and geographic perspective, utilizing a combination of field surveys and Google Street View crawling methods to gather graffiti data in the city of Novi Sad, Serbia. Subsequently, an analysis of the spatial distribution patterns and linguistic content of graffiti is conducted to explore whether Novi Sad's graffiti distribution exhibits specific spatial characteristics and whether there are spatial connections among graffiti contents. Through the linguistic and geographic examination of graffiti in Novi Sad, this research not only explores spatial distribution patterns but also delves deeper into the national ideologies of the Serbian population and the identity affiliations across different urban districts.

Keywords: graffiti, street view, machine learning, ideological expression, text recognition, Novi Sad

1. Background

Graffiti could be used to express support for football teams, political affiliation and ideology, which is prevalent in South-Eastern and Central Europe. In the studies of graffiti, the spatial location of graffiti constitutes a crucial research element, as the location and spatial context are important for understanding graffiti (Cresswell, 1992; Ferrell & Weide, 2010). However, current research on graffiti in the South-Eastern mainly focuses on the domains of sociology and cultural studies, with an absence of systematic studies from geographic perspectives (Miklavcic, 2008; Serafis et al., 2018).

Traditionally, graffiti images have been collected through field surveys, where researchers physically visit locations to photograph and document graffiti (Debras, 2019; Campos, 2015; Miklavcic, 2008; Bloch, 2020; Zaimakis, 2015). This method allows for detailed and context-rich data collection but is limited to research area coverage. In contrast, new methods use more data sources to collect graffiti images. One such method involves the use of Google Street View (GSV) images (Alexandros & Eleftheria, 2023; Tessio et al., 2020). This approach allows for the bulk collection of graffiti data across a wider geographic area without the need for physical presence. By developing a reasonable sampling strategy for selecting locations and using deep learning models, researchers can identify street views containing graffiti from GSV images (Javier et al., 2021).

Given the constraints of existing research, this study aims to combine the two data collecting methods:

1) Google Street View Method: Develop a sampling strategy for selecting locations, followed by bulk crawling of Google Street View images from these locations. Utilize a deep learning model to identify street views containing graffiti.

2) Traditional Field Survey: Conduct physical field surveys to collect graffiti images directly.

Then this research will explore the differences and similarities in textual content and spatial distribution of graffiti collected through these two methods.

2. Research goal

2.1 Integration of Perspectives

Our research aims to achieve a comprehensive understanding of graffiti in Novi Sad by integrating both linguistic and geographic perspectives. By examining the spatial locations and contexts of graffiti, we hope to gain insights into the geographical, cultural and social dynamics that drive this form of expression.

2.2 Innovative and Comparative Data Collection

To collect the data, we will employ an innovative and comparative data collection strategy. This involves combining traditional field surveys with Google Street View crawling. The traditional field surveys will allow us to gather detailed and context-rich data by physically visiting locations and documenting graffiti. In contrast, Google Street View crawling will enable us to collect a broader set of graffiti images efficiently.

2.3 Exploration of Graffiti Spatial Distribution Pattern

We will analyze the spatial characteristics and connections among the graffiti's content to determine if there are specific patterns or spatial relationships. This analysis will help us understand how graffiti reflects national ideologies and identity affiliations across different urban districts in Novi Sad.

Through this comprehensive examination, we hope to contribute to graffiti research in South-Eastern and Central Europe by providing new insights from both linguistic and geographic perspectives.

3. Methods and data

3.1 Data Collection Methods

3.1.1 Field Survey

The field survey dataset comprises 95 geolocated images of graffiti, collected during a comprehensive survey conducted last summer in 2023 by our mentor Cristiana Lucchetti in Novi Sad, Serbia. Each image is accompanied by precise geographic coordinates, enabling detailed spatial analysis. This dataset has been pre-processed, with all images intentionally capturing graffiti, thereby eliminating the need for further identification or filtering.

3.1.2 Google Street View Crawling

To explore additional data collection methods, we also gathered image data from Google Street View. The crawling process targeted the same geographic region covered in the field surveys, allowing for a direct comparison. By calculating the heading parameters for each sample point, we sent requests via the Google API to download the images of each point. Due to the lack of updated street view imagery for Novi Sad beyond 2014, the available images reflect historical views from that year. Our script captured images every 5 meters from both sides of the streets within the defined research area, resulting in a comprehensive dataset of 18,124 images.

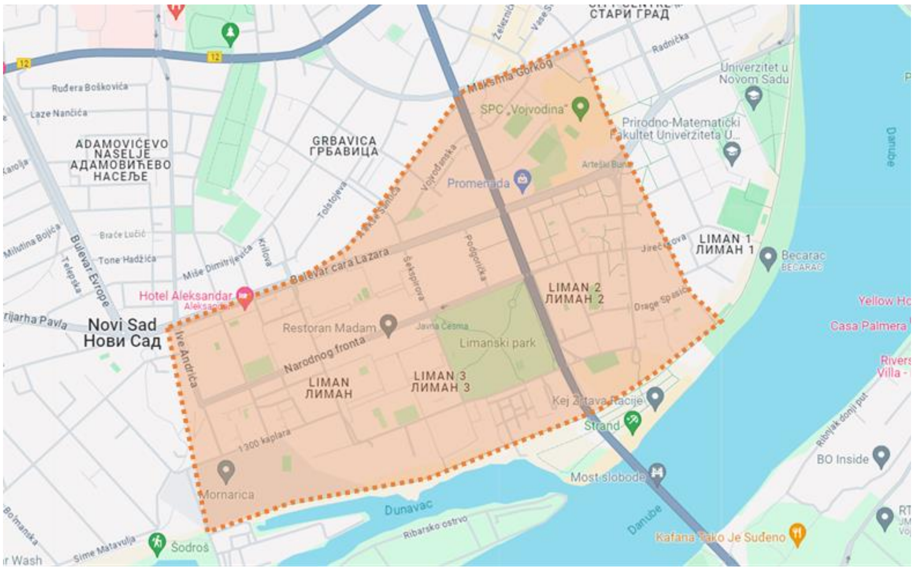

Figure 1. Study Area in Novi Sad

3.2 Data Processing Methods

3.2.1 Image Processing and Graffiti Identification

To process the vast dataset of 18,124 street view images collected via Google Street View, we explored the deep learning method to identify the graffiti in the images. The YOLO (You Only Look Once) model has demonstrated superior performance in object detection (Redmon et al., 2015). This study will utilize transfer learning on the YOLOv8 framework to train a graffiti detection model. The final accuracy of the model in recognizing graffiti is 60%. After identification, there are about 150 graffiti images left.

3.2.2 Manual Verification

Following the automated detection process, manual verification will be implemented. This step involves a detailed review of each image detected by the model to ensure the accuracy of graffiti identification. Manual verification addresses challenges such as varying angles, distances, and image resolutions which will enhance the quality of the data for subsequent analysis. There are 93 readable graffiti images for further analysis.

Figure 2. Automated detection (left) VS Manual verification images (right)

3) Translation and classification

Since the rich cultural and linguistic nuances of the graffiti in Novi Sad, our mentor Cristiana Lucchetti assisted us with this part. The graffiti was categorized into 7 themes: Commerce, Politics, Neighborhood Identity, Sports, Nationalist, Culture & History, and Address to Person, reflecting the range from commercial ads and political views to expressions of local identity and historical moments.

3.3 Data Analysis Methods

To analyze the spatial distribution and thematic content of graffiti, we employed a combination of spatial analysis techniques and comparative evaluation of the two data collection methods.

3.3.1 Map-Based Visualization

To visually display the spatial distribution and content of graffiti, we developed an interactive map. This map allows users to click on each graffiti location to view a popup containing details about the graffiti instance, including its category (e.g., Politics, Nationalist, Neighborhood Identity, etc.) and the specific image. This detailed view helps in understanding the thematic distribution of graffiti across different urban districts in Novi Sad.

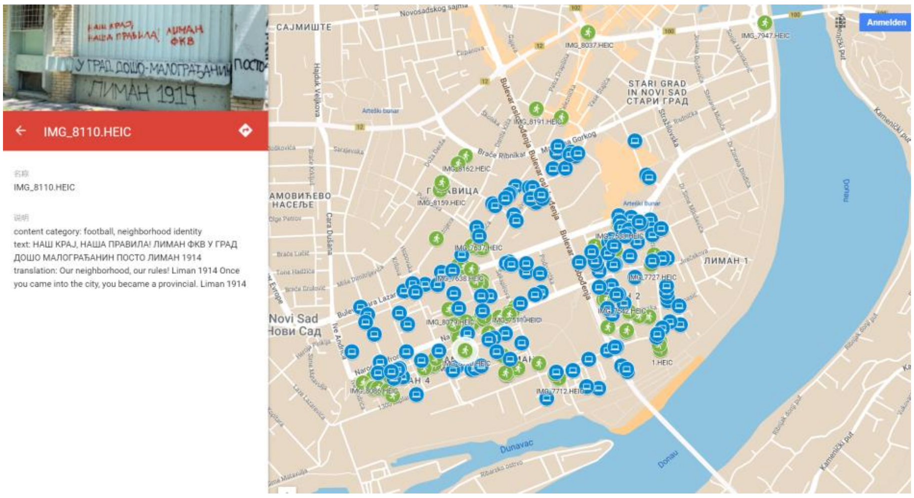

Figure 3. Screenshot of the interactive map

Heat maps are used to provide a more intuitive visualization of the density and distribution of graffiti instances. By overlaying the field survey and Google Street View data on heat maps, we could easily compare the concentration of graffiti in different areas. This visualization highlights the differences in the number of graffiti instances captured through field trips versus those identified via GSV crawling. The heat maps effectively show hotspots and coldspots, indicating areas with high or low graffiti density.

3.3.2 Spatial Autocorrelation Analysis

To understand the spatial distribution patterns of graffiti, we conducted a spatial autocorrelation analysis using the Join Count statistic （Moran, 1948). This analysis aimed to identify whether the spatial distribution of graffiti categories is clustered, dispersed, or random. For each data point, we defined its neighbourhood by selecting the three nearest neighbours (K=3).

The expected joins were calculated, and if the actual BW joins value was less than the expected value, the spatial distribution was considered clustered. Conversely, if the actual BB or WW joins value was greater than the expected value, the distribution was considered dispersed. A z-value of 2 or greater indicated statistical significance.

By integrating these analytical methods, we compare the spatial distribution patterns of graffiti collected through field surveys and Google Street View crawling. This involved analyzing the graffiti content, surrounding spatial elements, and community background to understand the reasons behind graffiti distribution in identified hotspots.

Figure 4. Project Flowchart

Results

Based on the classification and statistical analysis of the graffiti content, we generated Table 1. An image may contain multiple categories of graffiti (Figure 5, translation of graffiti content: "*Liman roots for FK Vojvodina! F89 F89 Firma 1234 Liman 30 years of F89 Suvara*") or a single piece of graffiti can be classified into different graffiti categories (Figure 6, translation of graffiti content: "*since 30 years you are a fortress of Serbianness, we love you, Republika Srpska! 09.01.1992 - Liman people -*"). Therefore, the sum of graffiti in each category does not equal the total number of images. As shown in Table 1, the graffiti in the dataset obtained from the field trip is mainly concentrated in four categories: Politics, Neighborhood identity, Sports, and Nationalist. This indicates that people often use graffiti to express their political stance, national sentiments, or neighborhood identity.

Figure 5. Image containing multiple categories of graffiti

Figure 6. Graffiti can be classified into multiple categories

Table 1. Summary of two datasets

Category	**Human Crawl Data**	**Machine Crawl Data**
Commerce	7	0
Politics	52	20
Neighborhood identity	35	44
Sports	22	67
Nationalist	38	18
Culture & history	5	0
Address to person	11	20

In the category of Politics, graffiti often includes phrases such as "*Death to fascism!*", "*Death to communism!*", and "*Death to capitalism!*", with "*Death to fascism!*" being the most frequent. In the Nationalist category, the phrases "*Kosovo is Serbia*" and "*Serbia to the Serbs*" frequently appear. For Neighborhood identity graffiti, most of them are "*Limma*" and symbols of the Limma district, expressing the strong identification of residents with their neighborhood.

In addition to political statements, support for a particular football team is also an important content of graffiti. In Novi Sad, the most frequently appearing graffiti in support of a football team is for the FKV team. Local teams often become symbols of national, cultural and neighborhood identity, representing local characteristics and cultural traditions. Supporting a team is not just about backing a sports team, it is also about preserving and passing on national culture and history.

In the machine crawl dataset, Sports and Neighborhood Identity graffiti occupy the largest proportion. The most common graffiti in these two categories are "*FKV*" and "*Limma*" because the names of the football team and the district are more common and easier to capture. In the Politics category, "*Antifa*" appears most frequently. Additionally, both the field trip dataset and the Google Street View dataset identified a common piece of graffiti. The graffiti reads "*We will avenge Kosovo.*" The Google Street View images of this area were collected in 2014, which means that even after ten years, while most of the surrounding graffiti has been covered or altered, this particular piece of graffiti has been preserved.

Figure 7. The same graffiti collected by two different data collection methods

Figures 8 and 9 show the heat maps of the graffiti collected by these two different methods. From Figure 8, it can be observed that the spatial distribution of the graffiti data collected manually is more concentrated compared to the graffiti collected through Google Street View. This is mainly because, during field research, researchers have more opportunities to explore the surroundings of their residential areas, being more familiar with these areas, thus collecting more graffiti in these regions. Unlike the images collected by Google Street View, which are mainly concentrated along the sides of the streets, graffiti collected through the field trip can be found in backyards and parks. These are places where people can enter and observe, whereas Google Street View images are generally limited to street views.

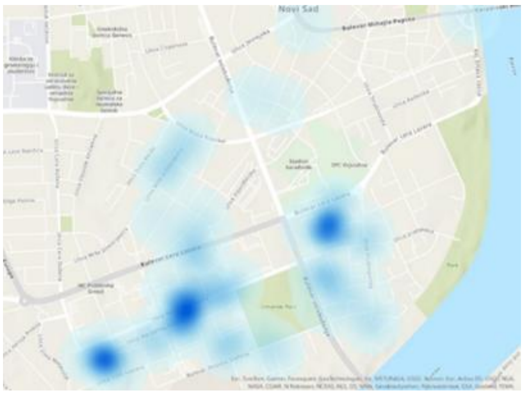

Figure 8. Heatmap of human crawl dataset

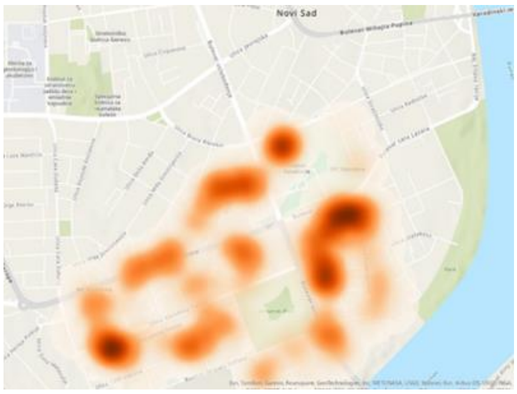

Figure 9. Heatmap of machine crawl dataset

To investigate the spatial distribution patterns of graffiti, distribution maps of different categories of graffiti from the two datasets were created, and a join count statistic was performed for each category in both datasets, resulting in Table 2 and Table 3. From both tables, it can be observed that the join count values for all graffiti categories are greater than the expected join count values, indicating that there is spatial clustering for all graffiti categories. In the field trip dataset, the z-values for Politics, Neighborhood Identity, and Nationalist are greater than 2, indicating significant results. In the Google street view dataset, only the Politics and Sports categories show significant results.

Table 2. Join count statics of human crawl dataset

Category	Joincount	Expected	z-value
Commerce	0.667	0.228	1.807
Politics	16.833	14.413	2.1987
Neighborhood identity	9.167	6.467	2.840
Sports	3.333	2.511	1.190
Nationalist	10.833	7.641	3.215
Culture&history	0.333	0.109	1.317
Address to person	0.833	0.598	0.623

Table 3. Join count statics of machine crawl dataset

Category	Joincount	Expected	z-value
Politics	3.333	2.043	2.029
Neighborhood identity	12.167	10.172	1.885
Sports	26.833	23.774	2.889
Nationalist	1.833	1.645	0.323
Address to person	2.333	2.043	0.457

From the spatial distribution maps, it can be seen that the Politics and Nationalist graffiti in both datasets are mainly concentrated around Limanski Park. This is likely because the park area has higher foot traffic, and as an open public space, creating political and nationalist graffiti in this area can maximize the graffiti's impact and visibility. Additionally, in both datasets, Neighborhood Identity graffiti is concentrated in the Limma neighborhood. This is primarily because the content of Neighborhood Identity graffiti mostly consists of "*Limma*" and the district symbol of Limma, hence it is mainly concentrated within this neighborhood.

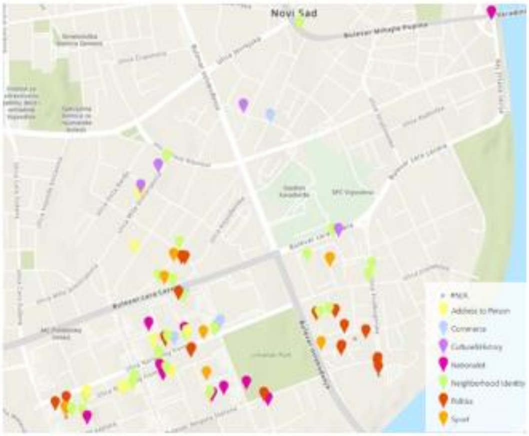

Figure 10. Graffiti distribution by category in human crawl dataset.

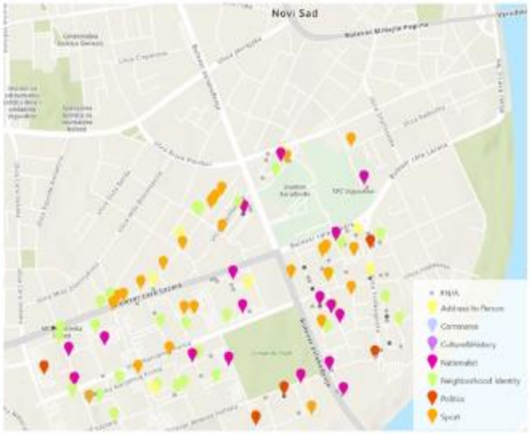

Figure 11. Graffiti distribution by category in machine crawl dataset.

Discussion and conclusion

This research compares the strengths and weaknesses of both field trips and Google Street View in graffiti data collection. Field trips could obtain high-quality data because researchers can filter out meaningless graffiti while taking photos, and can better understand the content of the graffiti by considering the specific context in which it exists. This method also allows for the capture of graffiti in areas that Google Street View does not cover, providing a more comprehensive dataset. However, the main drawbacks include limited area coverage due to time and resource constraints, and the requirement for being present at the site, which can be both labor-intensive and time-consuming.

In contrast, Google Street View can cover large geographical area and the convenience of collecting data online, which can significantly speed up the data collection process. The ability to rapidly capture graffiti from a broad area makes it a powerful tool for initial surveys or large-scale studies. However, the quality of graffiti images varies significantly, as they depend on the conditions when Google Street View images were captured, such as lighting, angle, distance and camera resolution. Additionally, manual verification is needed to ensure the accuracy and relevance of the graffiti data, which can be time-consuming. Furthermore, the quantity of graffiti that can be obtained also depends on the performance of the graffiti detection model; it's possible that images containing graffiti might not be detected by the model. The data collection process is also affected by the sampling distance of Google Street

View. If the sampling distance is small, the collection speed and efficiency will be lower, whereas a larger sampling distance might result in many graffiti being overlooked.

For the spatial pattern of the collected graffiti datasets, the graffiti data collected through field research show a more clustered spatial distribution, whereas the graffiti collected via Google Street View is more uniformly distributed. Additionally, graffiti of the same category in both datasets has a spatial clustering trend. Political and nationalist graffiti is mainly concentrated around parks, while neighborhood identity graffiti is primarily found within neighbourhoods.

Future research could explore hybrid approaches that combine both methods. For instance, Google Street View can be used for broad preliminary surveys, identifying key areas of interest, followed by targeted field trips to these areas for detailed data collection. This approach could optimize resource use while ensuring high-quality data collection. Additionally, the automatic recognition and interpretation of graffiti content by deep learning model is also a potential research direction for the future, aiming to reduce the need for manual verification.

Author contribution

Table 4. Author contribution

Author contribution	Roles and Responsibilities
Both	✓ Conceptualization: Designing the research framework. ✓ Discussion: Participating in discussions on findings and implications, integrating diverse viewpoints and expertise. ✓ Writing: Co-writing the final report.
Jing Tang	✓ Methodology: Developing methodologies for graffiti identification. ✓ Software/Programming: Developing Python scripts for data crawling and analysis. ✓ Formal Analysis: Conducting spatial autocorrelation analysis.
Yilin Wei	✓ Software/Programming: Assisting in image capturing with Python. ✓ Visualization: Creating maps and heat maps. ✓ Validation: Manual verification of data.

References

Alexandros Bartzokas-Tsiompras and Eleftheria Konstantinidou. (2023). Visual analytics of graffiti spatial patterns across 30 European city centres. Environment and Planning B: Urban Analytics and City Science 50, 2 (February 2023), 564–568. https://doi.org/10.1177/23998083221149426

Bloch, S. (2020). Broken Windows Ideology and the (Mis)Reading of Graffiti. Critical Criminology, 28(4), 703–720. https://doi.org/10.1007/s10612-019-09444-w

Campos, R. (2015). Youth, Graffiti, and the Aestheticization of Transgression. Social Analysis, 59(3). https://doi.org/10.3167/sa.2015.590302

Cresswell, T. (1992). The Crucial 'Where' of Graffiti: A Geographical Analysis of Reactions to Graffiti in New York. Environment and Planning D: Society and Space, 10(3), 329–344. https://doi.org/10.1068/d100329

Debras, C. (2019). Political graffiti in May 2018 at Nanterre University: A linguistic ethnographic analysis. Discourse & Society, 30(5), 441–464. https://doi.org/10.1177/0957926519855788

Ferrell, J., & Weide, R. D. (2010). Spot theory. City, 14(1–2), 48–62. https://doi.org/10.1080/13604810903525157

Javier Rozo Alzate, Marta S. Tabares, and Paola Vallejo. (2021). Graffiti and government in smart cities: a Deep Learning approach applied to Medellín City, Colombia. In International Conference on Data Science, E-learning and Information Systems 2021 (DATA'21), June 04, 2021. Association for Computing Machinery, New York, NY, USA, 160–165. https://doi.org/10.1145/3460620.3460749

Miklavcic, A. (2008). Slogans and graffiti: Postmemory among youth in the Italo–Slovenian borderland. American Ethnologist, 35(3), 440–453. https://doi.org/10.1111/j.1548-1425.2008.00045.x

Moran, P. A. (1948). The interpretation of statistical maps. Journal of the Royal Statistical Society. Series B (Methodological), 10(2), 243-251.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2015). You Only Look Once: Unified, Real-Time Object Detection. https://doi.org/10.48550/ARXIV.1506.02640

Serafis, D., Kitis, E. D., & Archakis, A. (2018). Graffiti slogans and the construction of collective identity: evidence from the anti-austerity protests in Greece. Text & Talk, 38(6), 775–797. https://doi.org/10.1515/text-2018-0023

Tessio Novack, Leonard Vorbeck, Heinrich Lorei, and Alexander Zipf. (2020). Towards Detecting Building Facades with Graffiti Artwork Based on Street View Images. ISPRS International Journal of Geo-Information 9, 2 (February 2020), 98. https://doi.org/10.3390/ijgi9020098

Zaimakis, Y. (2015). 'Welcome to the civilization of fear': on political graffiti heterotopias in Greece in times of crisis. Visual Communication, 14(4), 373–396. https://doi.org/10.1177/1470357215593845