International Journal of Geo-Information

MDPI

Article

Network-Based Hierarchical Feature Augmentation for Predicting Road Classes in OpenStreetMap

Müslüm Hacar [1,2,*], Diego Altafini [3] and Valerio Cutini [2]

[1] Department of Geomatic Engineering, Yildiz Technical University, 34220 Istanbul, Türkiye
[2] Department of Energy, Systems, Territory and Construction Engineering, University of Pisa, 56122 Pisa, Italy; valerio.cutini@unipi.it
[3] Welsh School of Architecture, Cardiff University, Cardiff CF10 3NB, UK; altafinid@cardiff.ac.uk
* Correspondence: mhacar@yildiz.edu.tr; Tel.: +90-212-383-5340

Abstract: The need to enrich the semantic completeness of OpenStreetMap (OSM) data is crucial for its effective use in geographic information systems and urban studies. Addressing this challenge, our research introduces a novel hierarchical feature augmentation approach to developing machine learning classifiers by the features retrieved from various levels of road network connectivity. This method systematically augments the feature space by incorporating measure values of connected road features, thereby integrating extensive contextual information from the network hierarchy. In our evaluation, conducted across diverse urban landscapes in six cities in Italy and Türkiye, we tested two geometry-, six centrality-, and eight semantic-based features to predict road functional classes stored as a *highway* = * key in OSM. The findings indicate a marginal impact of geometric features and city identifiers on classification performance. Utilizing centrality attributes alongside semantic features in a direct, non-hierarchical manner results in an F1 score of 80%. However, integrating these features in our network-based hierarchical feature augmentation process remarkably increases the F1 score up to 85%. The success of our approach underlines the importance of network-based feature engineering in capturing the complex dependencies of geographic data, considering a more accurate and contextually aware OSM classification framework.

Keywords: VGI; machine learning; feature engineering; centrality measure; road network; OpenStreetMap

Citation: Hacar, M.; Altafini, D.; Cutini, V. Network-Based Hierarchical Feature Augmentation for Predicting Road Classes in OpenStreetMap. *ISPRS Int. J. Geo-Inf.* **2024**, *13*, 456. https://doi.org/10.3390/ijgi13120456

Academic Editors: Wolfgang Kainz and Wei Huang

Received: 16 October 2024
Revised: 29 November 2024
Accepted: 16 December 2024
Published: 17 December 2024

1. Introduction

Volunteer Geographic Information (VGI) refers to community-driven efforts that generate freely accessible geographic data. Among the various VGI platforms, OpenStreetMap (OSM) stands out as a leading initiative, enabling users worldwide to contribute and edit map data. However, because OSM relies heavily on volunteer contributions, it can sometimes suffer from inconsistencies in data quality, coverage, and semantic annotations [1,2]. These issues are particularly noticeable in regions with fewer active contributors. To ensure a consistent level of data quality in OSM, particularly in such regions, it is important to inform volunteers about the minimum requirements for their geometric or semantic contributions. A tool based on an artificial intelligence (AI) framework could provide help-assistant-like suggestions and guidance to users as they contribute raw or current data, helping them to assign appropriate classifications as semantic key values and improve the completeness and also the accuracy of their contributions [3,4]. Developing such AI tools requires extensive research on predictive models validated with geographic data from various regions representing distant topographic characteristics (i.e., both natural and constructed) of urban and rural areas. The most critical parameter for these prediction models is the quality of and variation in features derived from the data. In this context, the only available data are the existing OSM elements, tags, and the newly added raw road geometry from the user. Therefore, the feature extraction process must be deeply

investigated in an intrinsic analysis of OSM data structures, enabling the development of robust AI tools that can effectively support users.

Road lines are the main contributions of OSM volunteers, which are parallel to the OSM semantic classes given at OSM Map Features [5]. Also, they are fundamental components of urban systems once they are structured as in network topology, playing a critical role in shaping urban form, accessibility, and mobility patterns [6–8]. Recent studies have indicated the potential of OSM data in analyzing urban studies based on road networks, such as traffic flow, routing, accessibility, and vulnerability [9–11]. Despite their flexibility, working with OSM data is not without its challenges. Being a collaborative dataset, contributions may have different levels of detail, completeness, and accuracy across different regions. Moreover, the quality of the metadata itself can be a problem, with several overlapping, contradictory, or missing semantic tags plus inconsistencies and errors that often require extensive preprocessing [12,13]. In particular, the variation in the completeness of road tags reduces the applicability of a trusted network-based method to places having fewer active contributors [14,15]. Therefore, research focused on developing a prediction model for classifying roads based on their functionalities and integrating the classes with OSM data as semantic tags defined as in OSM Map Features [5] is innovative.

The evaluation of OSM data quality has been a central theme in VGI research, utilizing both intrinsic and extrinsic approaches. Intrinsic evaluations rely on the dataset's own characteristics, emphasizing geometric and semantic coherence, whereas extrinsic evaluations compare OSM data with authoritative datasets to assess accuracy and completeness [16,17]. Reference-based studies have demonstrated OSM's robustness in urban contexts but also identified regional disparities in data coverage and quality [18–20]. However, intrinsic analyses have focused on network analysis, tagging trends, and tag generation [21–28].

Research utilizing machine learning (ML) techniques to derive intrinsic information from OSM data has been growing significantly. Jilani et al. [22] proposed an alternative representation to primal or dual forms of road networks, then suggested an ML model [14] to automate highway tag assessments that refer to OSM road classes. After using the relatively reliable OSM road data in London as a reference, they found that more than 50% of the *residential*, *pedestrian*, *primary*, *motorway*, *primary_link*, and *motorway_link* data were predicted correctly, while less than 40% of the *cycleway*, *bridleway*, *path*, *secondary*, and *secondary_link* data were correct. Corcoran et al. [23] proposed a method for inferring semantic type information from a geometric representation of a road network that involves modeling the network as a probabilistic graphical model and utilizing maximum-margin learning alongside a fusion move approach for inference. This method captures features of individual streets, like linearity, and relationships between streets, such as semantic type co-occurrences. They obtained a test result of 68% in precision. Basiri et al. [3] proposed an approach to creating new map elements or editing current data using mining techniques, which include cartographic generalization and matching steps. Keller et al. [24] developed an ML-based framework that predicts the average speed on rural roads based on road information in OSM. Hacar [25] explored using geometric (rectangularity, density, area, and distance to bus stops and shops) and semantic (amenity) attributes for *leisure* tag identification in OSM data, improving the completeness and quality of the information available. Vargas Munoz et al. [4] developed an ML-based interactive method that assists OSM volunteers in annotating and validating rural building information, which can improve the efficiency and accuracy of humanitarian mapping projects. Alghanim et al. [26] developed a new trustworthiness measure based on the edit history and contributor behavior, which significantly improved the ML model's accuracy for road classification to 87.75% when using trusted data, compared with 57.98% with untrusted data. This demonstrates the importance of data quality in VGI for developing prediction models. Their model was tested in London, which is one of the most actively contributed-to datasets in the OSM repository [27,28]. However, the semantic type of information they used (i.e., *tunnel*, *bridge*, *maxspeed*, and *oneway*) for feature generation could be problematic if the area of interest is devoid of those tags in the OSM history due to having fewer active contributors. Pazoky

and Pahlavani [29] evaluated various ML classifiers for enriching OSM road class data using centrality measures as predictive features (independent variables). Among the classifiers, random forest (RF) demonstrated prominent performance. However, the exclusion of road functional tags such as *motorway* and various *_link* tags from the testing dataset, along with expert-driven filtering of the test data, limits the generalizability of their approach to other OSM data. Juhász et al. [30] proposed an approach to generating semantic road classes by using ChatGPT. They prompted a detailed chat to classify the objects in the input street-view image. Then, they instructed the generative model to suggest the most appropriate tagging for each road in OSM. The experimental result in a small test area in Miami shows that their prompting approach performed at 62% for semantic road categories. Yang et al. [31] developed a neural-based Geographic Knowledge Base Question Answering (GeoKBQA) system that connects natural language questions to OSM data. They created a large GeoKBQA dataset and used an end-to-end entity-linking method to identify and relate spatial entities. Their system translates place-related questions into GeoSPARQL queries, offering a practical and scalable approach to improving the usability of OSM data, overcoming the limitations of earlier rule-based methods. Hochmair et al. [32] evaluated the performance of generative AI. Their study compared the correctness of four prominent chatbots (ChatGPT-4, Gemini, Claude-3, and Copilot) across 76 spatial tasks, highlighting strengths in spatial literacy and GIS theory and underscoring the growing role of LLMs in geoscience, particularly their application in geo-data analysis, GISs, and mapping.

Previous studies, briefly, have led us to consider that there is a need to test the proposed prediction model in different areas of interest for the validation of the harmony between intrinsic features and the network samples. Moreover, the studies show that there are some possibilities:

- to apply ML-based approaches to assess the quality of current OSM road data,
- to apply these approaches to enrich the semantic tags of OSM road data,
- to use intrinsic measures of OSM roads in a prediction model, and
- to use geometry-, semantic-, and centrality-based measures as the predictive features.

OSM data are widely used in navigation systems, urban planning, and geospatial analyses; however, inconsistencies and gaps in road classification present challenges for their effective use. For instance, improved road classification can significantly benefit traffic navigation systems by ensuring that routes are planned based on accurate road functionalities, while urban planners can use these insights to prioritize key road segments for infrastructure improvements. This study introduces a novel method for predicting road classes, addressing these challenges by improving the accuracy and completeness of OSM data, and aims to enrich the semantic information of OSM roads, stored as *highway* = * tags, by developing a prediction model utilizing intrinsic features retrieved from OSM road networks. We propose a novel network-based hierarchical feature augmentation approach that systematically expands the feature space by incorporating aggregated measures from connected road lines across multiple levels of the network. This allows the model to represent not only the local characteristics of each road segment but also the broader contextual information from the surrounding network hierarchy. The evaluation was conducted using data from the three most populous cities in both Türkiye and Italy, ensuring the model's robustness across diverse urban environments with varying local morphology patterns and contributor activities. The subsequent sections are organized as follows. Section 2 first introduces the geometry-, centrality-, and semantic-based measures. It then demonstrates the proposed network-based feature augmentation approach. Finally, it presents the case study areas and the data structure of the road networks. Section 3 gives the results of the ML prediction model both with and without using the feature augmentation and evaluates the model performance across pairwise feature sets. The final section includes the study outcomes and discusses the further optimization of the model.

2. Materials and Methods

2.1. Geometry-, Centrality-, and Semantic-Based Measures

In this study, we employed a diverse set of geometry-, centrality-, and semantic-based measures to effectively represent the characteristics of road lines within the OSM dataset (Table 1). The selection of these measures was driven by their ability to represent various dimensions of road characteristics. Geometric measures provide basic physical properties, centrality measures capture the segment's role and importance within the network, and semantic measures add contextual information. Together, they form a comprehensive feature set that effectively differentiates between various road classes.

Table 1. Geometric-, topologic-, and semantic-based measures.

Measure Type	Measure Name	Formula
Geometry-based	Length (L)	$L = \sum_{i=1}^{n-1} \sqrt{(X_{i+1} - X_i)^2 + (Y_{i+1} - Y_i)^2}$
	Sinuosity (SN)	$SN = \frac{L_{Euclidean}}{L}$
Centrality-based	Straightness (ST)	$ST(i) = \frac{1}{n-1} \sum_{j \in V, j \neq i} \frac{d_{ij}^{Euclidean}}{d_{ij}}$ where $d_{ij}^{Euclidean}$ is the Euclidean distance between nodes i and j along a straight line.
	Betweenness (B)	$B(v) = \sum_{s,t \in V} \frac{\sigma(s,t/v)}{\sigma(s,t)}$ where v represents an intermediary node in the graph, which lies on the shortest path between the source node s and target node t.
	Connectivity (C)	$C = e/v$ Number of edge connections (e) at a junction node (v)
	Angular Connectivity (AC)	$AC = \sum_{u \in N} w(u,v)$ Segment angular connectivity is defined as the cumulative turn angle (w) between a node (v) and its neighbor (u)
	Angular Total Depth (ATD)	$ATD = \sum_{k} d\theta_{Ck}$
	Angular Integration (AIn)	$AIn = 1/n \sum_{k} d\theta_{\pi X,i}$
	Angular Choice (ACh)	$ACh = \sum_{j} \sum_{k} g_{jk}^{(i) \lg(j<k)}_{jk}$
	Normalized Angular Choice ($NACH$)	$NACH = \frac{log(ACh_{\alpha}^{\ell}(x)+1)}{log(ATD_{\alpha}^{\ell}(x)+3)}$
	Normalized Angular Integration ($NAIN$)	$NAIN = \frac{n^{1.2}}{ATD_{\alpha}^{\ell}(x)}$
Semantic-based	Distance to gas station (D_fuel)	$D_i = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2}$
	Distance to bank (D_bank)	
	Distance to pharmacy ($D\_pharmacy$)	
	Distance to cafe (D_cafe)	
	Distance to school (D_school)	
	Distance to restaurant ($D\_restaurant$)	
	Distance to hospital ($D\_hospital$)	
	Distance to bus_station (D_bus_stop)	

Geometric measures provide fundamental information about the physical properties of road segments, which seems significant for distinguishing between different functional road classes. The length (*L*) of a road segment is a basic yet significant geometric property, representing the spatial extent of the road. It helps differentiate between long, continuous roads like highways and shorter, more segmented roads such as local streets. Sinuosity (*SN*) quantifies how much a road deviates from a straight line. This measure is particularly useful in urban areas where winding roads indicate residential zones, as opposed to straighter roads found in commercial or industrial areas. These geometric measures were chosen for their ability to capture basic yet distinct physical characteristics that are often correlated with different road functionalities.

Centrality measures are crucial for understanding the role of each road segment within the broader network. They highlight the importance and connectivity of roads, which are essential for distinguishing major thoroughfares from minor streets. Also, they help in understanding the evolution and organization of road networks across different urban settlements, revealing patterns that can be applied to various urban studies [33–42]. Straightness centrality (*ST*) is theoretically similar to Sinuosity but is measured on nodes. It is calculated as a ratio between the real and Euclidean distance from each node to every other node. Therefore, to make it represent a road line, the mean value of the start and end nodes would be a good approach. Roads with a high degree of straightness are typically key routes for through traffic, suggesting their higher functional importance. Betweenness centrality (*B*) evaluates the significance of each node or edge within the network in terms of travel. It quantifies how frequently a node/edge is utilized when taking the shortest route between each pair of nodes [7,43,44]. It is similar to angular choice in the Space Syntax [45–47], but, in our study, to reflect different characteristics of road networks, *B* is computed on a primal graph instead of a dual one.

Connectivity (*C*) and Angular Connectivity (*AC*) assess the local and angular connectivity of a road segment, respectively, providing insights into how well a segment is integrated into the local road network. High connectivity values often indicate crucial junctions or segments acting as junctions in the network. Angular Total Depth (*ATD*) is the sum of the angular weighted (θ) topological distance (d) among all nodes (streets in the dual approach). Angular Integration (*AIn*) measures the relative accessibility from a node toward all others in the system (n). Angular Choice (*ACh*) is the number of times a certain node is traversed in order to reach all others in the system. It denotes the preferential routes. *ACh*, *ATD*, and *AIn* are used to calculate the Normalized Angular Choice (*NACH*) and Normalized Angular Integration (*NAIN*) [48], detailed in Table 1, capturing the navigational complexity and accessibility of road segments. They are particularly useful for identifying roads that serve as key connectors or destinations within the network. Also, we employed the preprocessing approach presented by Hacar et al. [49], where roads are split at the connections to maintain non-topologically connected road intersections after the following simplification process. To reduce the computational time and speed up the calculation of *C*, *AC*, *NACH*, and *NAIN* in our road networks, the split roads were simplified using the Douglas–Peucker [50] algorithm. Then, these measures were calculated based on the segments of the simplified roads. The inclusion of all these centrality measures allows us to model the influence and accessibility of road segments more accurately, reflecting their functional significance within the urban road network.

Semantic measures were selected to capture the contextual and functional aspects of road segments relative to their surrounding environment. They provide information on the proximity of roads to various amenities, which is crucial for understanding the intended use and importance of these roads. Distances to amenities such as gas stations, banks, pharmacies, and other amenities (Table 1) provide a contextual understanding of the road segment's environment. Each semantic measure was chosen to reflect specific aspects of the road's context, which is essential for distinguishing roads serving different functional purposes, such as residential streets versus highways that typically lack close proximity to amenities.

Geometric and semantic measures were calculated using the software described in [51] and the Python package described in [52], respectively. The centrality metrics *B* and *ST* were computed with the package described in [53], while *C*, *AC*, *NACH*, and *NAIN* were derived utilizing the background of [45,48,54] and the software described in [47].

2.2. The Proposed Network-Based Hierarchical Feature Augmentation

In this section, we present our approach for expanding the feature set of each road line through network-based hierarchical feature augmentation. This method systematically enriches the feature space by aggregating the measures of neighboring road lines over multiple hierarchical levels (Figure 1).

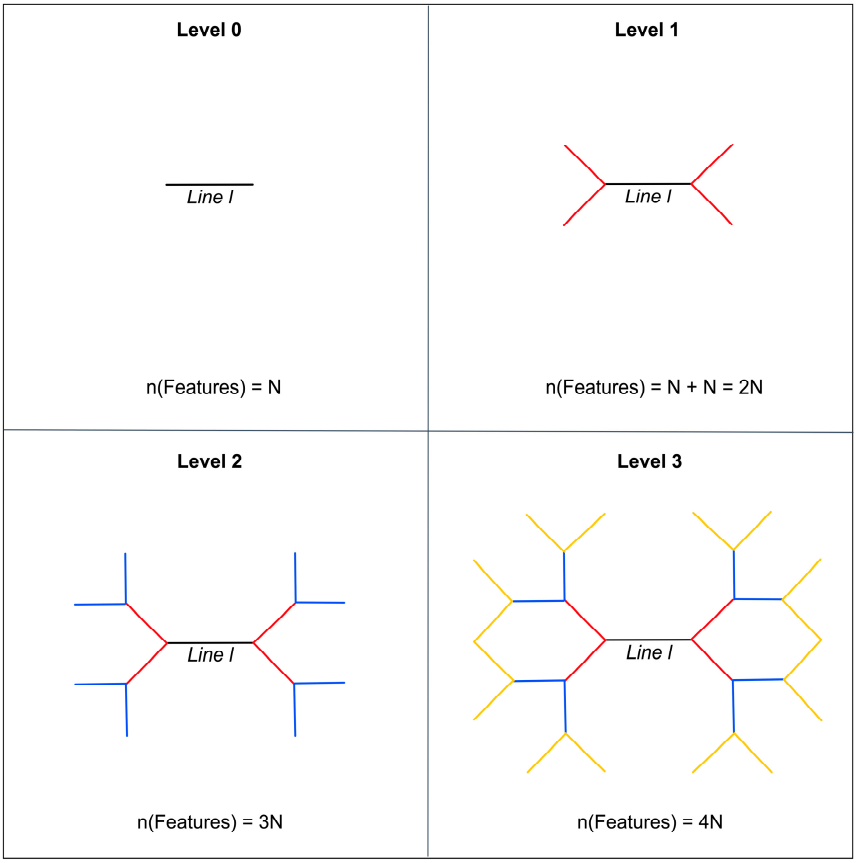

Figure 1. Levels of the network hierarchy in the feature augmentation of *Line l*.

- Level 0: Base Features
  Each road line, referred to as *Line l*, initially possesses its intrinsic geometric, centrality, and semantic measures, as detailed in the previous section (Table 1). These measures form the Level 0 feature set and solely describe the road itself without considering its relationship with the surrounding road network.
- Level 1: Direct Neighbor Aggregation
  To incorporate contextual information, we augment the feature set of *Line l* by including aggregated measures of its directly connected neighboring road lines (red lines in Figure 1). They are the immediate connections to *Line l* and provide additional context about the local network configuration. The resulting doubled features, along with the original Level 0 features, constitute the Level 1 feature set. In this study, we adopted the mean aggregation method for all levels of networks, as it provides a balanced representation of the neighboring characteristics.
- Level 2: Indirect Neighbor Aggregation
  Expanding further, the Level 2 feature set is derived by aggregating the measures of all roads connected to the direct neighbors (Level 1) of *Line l*. This process considers

the indirect neighbors (blue lines in Figure 1), effectively capturing information from a broader section of the road network. The aggregated features are combined with the Level 1 feature set. This approach enables the inclusion of second-degree neighborhood information, ensuring the contextual understanding of each road within the network hierarchy.

- Level 3 and Beyond
  The hierarchical aggregation can be extended to further levels, such as Level 3, by incorporating measures from roads connected to the Level 2 neighbors of *Line l*. At each level, the feature set of the target road line is incrementally augmented by incorporating the aggregated measures of increasingly distant, but converging, lines (yellow lines in Figure 1) within the network. Finally, the number of features is augmented as N+1 times for Level N (Figure 1).

This augmented feature sets reflect a progressively broader range of contextual information, which can be particularly beneficial for complex road networks. It allows for a more comprehensive representation of road lines, capturing both their local and global network characteristics.

2.3. The Network Data and the Study Areas

OSM data are structured with nodes, ways, and relations, each attributed by tags that provide semantic information [55,56]. This study focused on road data extracted from OSM ways, with the datasets sourced from Geofabrik.de [57]. Provincial administrative boundaries from GADM [58] were used to clip the datasets by using the osmosis tool [59]. For this study, only functional road types were filtered as *motorway*, *motorway_link*, *primary*, *primary_link*, *secondary*, *secondary_link*, *tertiary*, *tertiary_link*, *trunk*, *trunk_link*, and *residential* [5]. The study covered the three most populous metropolitan cities in Türkiye (Istanbul, Ankara, and Izmir) and Italy (Rome, Milan, and Naples) (Figure 2). It is important to clarify that the term 'city', as used in this study, refers to the administrative boundaries of provinces in Türkiye and Italy. These provinces include both urban cores and their surrounding rural areas, ensuring a diverse representation of geographic and demographic contexts. The RF classifier employed in the study was randomly fed with data from both urban and rural areas within these provinces. This ensured that the training and testing phases of the classifier accounted for the diverse characteristics of road networks in different geographic contexts, enabling a robust and inclusive analysis.

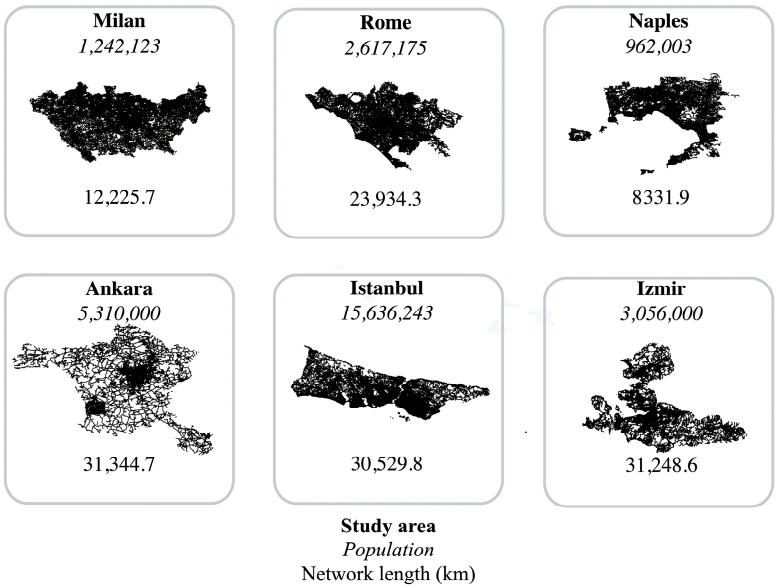

Figure 2. The study areas and data.

3. Results

3.1. Determining the Aggregation Techniques During the Network's Regeneration

Due to the transformation of OSM roads into smaller segments for the calculation of centrality measures, it was necessary to aggregate these segment-level features back to the corresponding road line for the training and testing phases. Therefore, adopting suitable techniques ensures a more accurate and context-aware classification of road networks and increases the performance of the road classification model. In this section, we explain the process of selecting the optimal techniques for transferring centrality-based measures from network segments back to the original OSM road lines. Various aggregation techniques were considered, including the mean, maximum, minimum, and sum, to summarize the segment-level measures into road-level features. The RF classifier was used to evaluate the performance of each technique in predicting road classes, and the results are depicted in the feature importance chart below (Figure 3). The minimum technique yielded the highest feature importance for the *B* and *NAIN* measures, suggesting it is the most effective method for these centrality metrics. On the other hand, the mean technique showed superior performance for all other centrality measures (*ST*, *C*, *AC*, and *NACH*). Moreover, the centrality-based measures calculated for each city are shown in Figure 4, where each measure highlights distinct characteristics of the roads in the respective city. The dataset used in the ML model integrates both urban and rural samples, ensuring a diverse geographic representation. Features like *ST* and *C* effectively present the linearity and junction density of rural roads, while measures like *NACH* and *NAIN* highlight the navigational complexity and accessibility inherent to urban networks. By incorporating this broad range of features, the model is able to distinguish urban roads, which typically serve as hubs within a dense network, from rural roads, which are often more isolated and linear. Figure 4 highlights these differences visually, illustrating how the centrality measures vary spatially within each province. These patterns allow the model to inherently account for urban–rural distinctions without the need for explicit categorization.

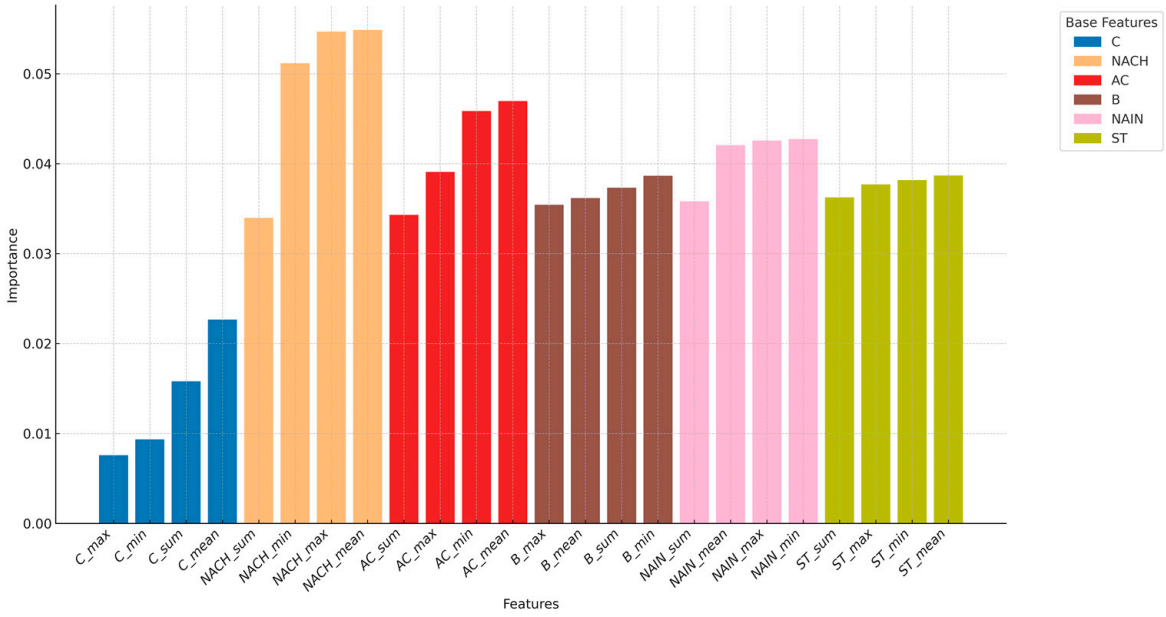

Figure 3. Feature importance of alternative aggregation techniques.

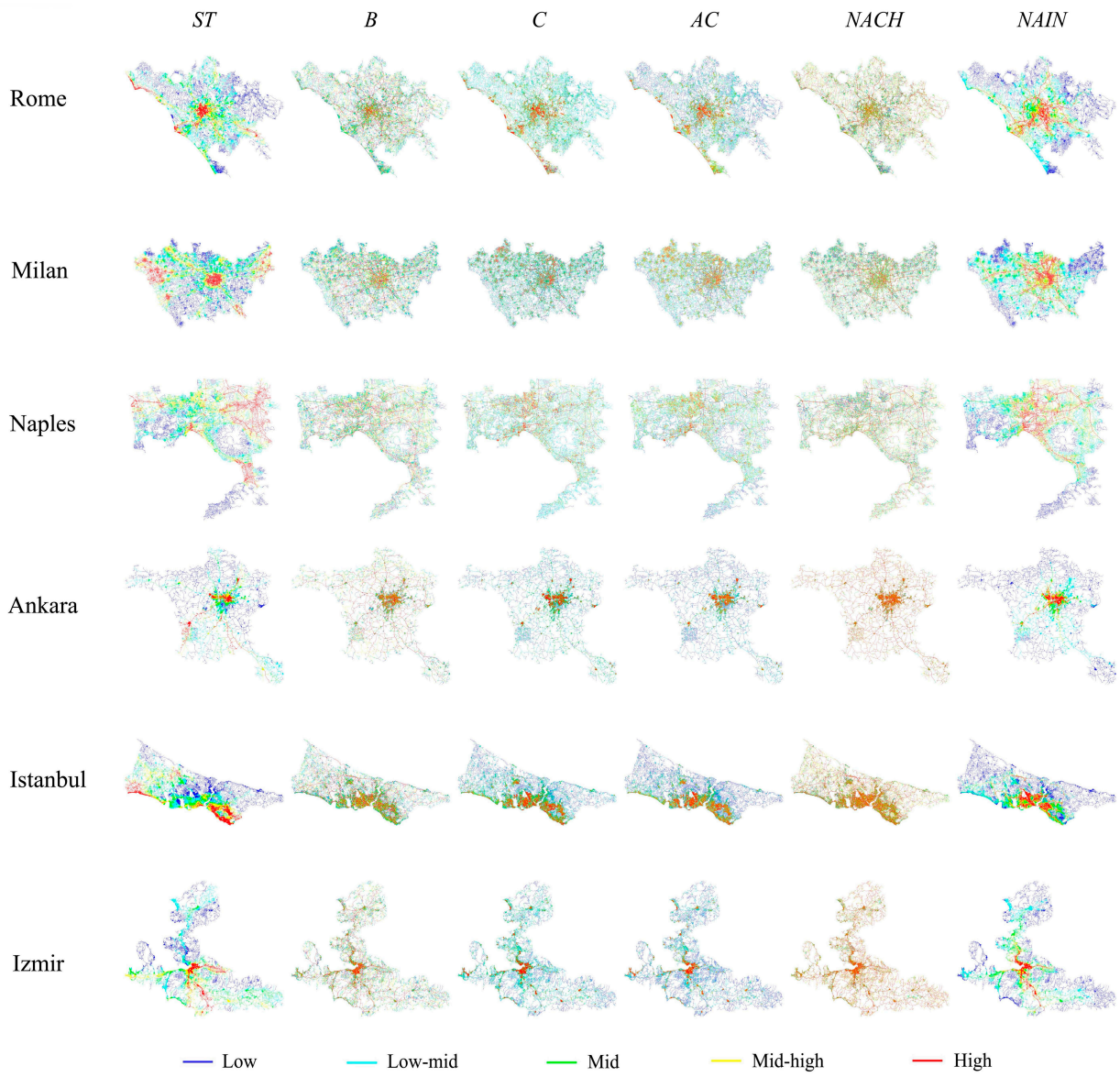

Figure 4. Distribution of centrality-based measures across each city's network.

3.2. Evaluation of the Prediction Results Before Feature Augmentation

We used the RF, XGBoost, and LightGBM classifiers, with the geometry-, centrality-, and semantic-based features mentioned in Table 1, to predict road functional classes stored in the *highway* = * tag. As a preprocess for all alternative prediction models, all measures were normalized. While they were used as independent prediction features, the *highway* feature was used as the dependent variable in our model. *highway* values were selected as *motorway*, *motorway_link*, *primary*, *primary_link*, *residential*, *secondary*, *secondary_link*, *tertiary*, *tertiary_link*, *trunk*, and *trunk_link*.

The RF model performed well in identifying various road types within the OSM data, as reflected by the precision, recall, and F1-scores in Table 2. It was the most accurate with *residential* roads, showing high scores across precision and recall, indicating a strong ability to correctly identify and capture the majority of residential roads in the dataset. *motorways* and *trunk* roads also saw high precision, suggesting that the model reliably recognized these road types when they were present. However, the model was less effective with link roads, particularly *primary_link*, *secondary_link*, and *tertiary_link* roads, as evidenced by lower recall scores. This suggests that, although the model usually correctly classifies link roads when it considers them suitable for classification, it misses a significant proportion

of classes related to the use of existing features. The overall weighted scores across all road types were solid, indicating robust model performance in general, but with room for improvement in distinguishing certain road types, particularly those represented as link roads in the dataset. Also, the overall accuracies of XGBoost and LightGBM were 80.1% and 80.2%, respectively, without using city identifiers.

Table 2. The results in RF prediction by geometry-, semantic-, and centrality-based features.

Key Value	Precision	Recall	F1-Score	Object Count	Area of Interest
motorway	0.85	0.85	0.85	876	All cases without a city identifier
motorway_link	0.73	0.39	0.51	971	
primary	0.75	0.56	0.64	3366	
primary_link	0.68	0.18	0.28	895	
residential	0.85	0.98	0.91	55,912	
secondary	0.69	0.44	0.54	4911	
secondary_link	0.77	0.11	0.19	562	
tertiary	0.65	0.48	0.55	10,444	
tertiary_link	0.84	0.11	0.19	427	
trunk	0.86	0.64	0.73	1210	
trunk_link	0.78	0.23	0.35	1116	
Weighted	0.80	0.82	0.80	80,690	

The overall weighted scores for precision, recall, and F1-score are all slightly higher in Table 3, which includes city identifiers, pointing to a marginal enhancement in model performance. This indicates that the geographic context plays a role in the effectiveness of the prediction model. For the *motorway* class, the precision and recall increased slightly, indicating that the model became more accurate and consistent when city data were factored in. Similarly, for *primary* and *tertiary* roads, we saw a marginal increase in both precision and recall, suggesting that including city identifiers helps the model to better classify these road types. However, the performance on link roads still shows low recall values. The notable increase in precision for *tertiary_link* roads, up to 0.86, is tempered by the recall remaining low (0.13), indicating that these types of roads are still challenging for the model to consistently identify. Moreover, *trunk* roads also saw an improvement in precision and recall, but *trunk_link* roads, much like the link roads, although better in precision, still lagged in recall. This pattern across the link road categories suggests that, while the model is refining its predictions with the additional city context, there is still a tendency to miss identifying some instances of less common road types. The overall accuracies of XGBoost and LightGBM are 80.8% and 81.0%, respectively, indicating similar weighted trends with RF. As a result, while the incorporation of city identifiers and the combination of geometry-, semantic-, and centrality-based features have improved the model's performance in predicting road types, there remains the potential for further improvement. Specifically, the model's ability to recognize and accurately classify link roads could benefit from additional refinement on features.

Analyzing the performance of the RF classifier using different feature sets can provide information on the strengths and weaknesses of each approach and highlight opportunities for improvement. When we look at Table 4, which presents the results using only geometry-based features, we see that the model had a particularly challenging time identifying most road types, except for *residential* roads. For instance, the precision for *motorway* was just 0.20, and the recall was even lower at 0.06, leading to an F1-score of 0.10. This suggests that geometric features alone are not sufficient for the model to accurately predict *motorway*. The recall figures across all road types are low, which indicates that the model missed a significant number of true cases. However, *residential* roads stand out as an exception, with relatively high precision and recall values, which could be due to the higher prevalence of these road types in the dataset or the specific distinction in geometric properties (i.e., L and SN), making it easier for the model to learn and predict.

Table 3. The results in RF prediction by a city identifier and geometry-, semantic-, and centrality-based features.

Key Value	Precision	Recall	F1-Score	Object Count	Area of Interest
motorway	0.88	0.87	0.87	876	All cases with city identifiers
motorway_link	0.73	0.41	0.53	971	
primary	0.75	0.58	0.66	3366	
primary_link	0.62	0.24	0.34	895	
residential	0.85	0.98	0.91	55,912	
secondary	0.70	0.47	0.56	4911	
secondary_link	0.78	0.14	0.24	562	
tertiary	0.68	0.49	0.57	10,444	
tertiary_link	0.86	0.13	0.22	427	
trunk	0.86	0.67	0.75	1210	
trunk_link	0.78	0.26	0.39	1116	
Weighted	0.81	0.83	0.81	80,690	

Table 4. The results in RF prediction by only geometry-based features.

Key Value	Precision	Recall	F1-Score	Object Count	Area of Interest
motorway	0.20	0.06	0.10	876	All cases without a city identifier
motorway_link	0.02	0.01	0.01	971	
primary	0.08	0.03	0.04	3366	
primary_link	0.03	0.01	0.01	895	
residential	0.73	0.92	0.81	55,912	
secondary	0.12	0.05	0.07	4911	
secondary_link	0.03	0.01	0.01	562	
tertiary	0.32	0.16	0.21	10,444	
tertiary_link	0.02	0.00	0.01	427	
trunk	0.12	0.04	0.06	1210	
trunk_link	0.01	0.00	0.00	1116	
Weighted	0.56	0.66	0.60	80,690	

In Table 5, we see a marked improvement in precision and recall across almost all road types. For example, the precision for *motorway* improved to 0.75 and the recall to 0.56, resulting in a higher F1-score of 0.64 compared with the geometry-based model. This improvement indicates that semantic features have a significant impact on the model's ability to classify road types accurately. The results for *secondary* roads and links are notably better than those in Table 4, demonstrating that semantic features provide valuable information that aids the classification process.

Table 5. The results in RF prediction by only semantic-based features.

Key Value	Precision	Recall	F1-Score	Object Count	Area of Interest
motorway	0.75	0.56	0.64	876	All cases without a city identifier
motorway_link	0.65	0.53	0.58	971	
primary	0.66	0.44	0.53	3366	
primary_link	0.51	0.32	0.39	895	
residential	0.84	0.97	0.90	55,912	
secondary	0.70	0.47	0.56	4911	
secondary_link	0.54	0.30	0.38	562	
tertiary	0.75	0.47	0.58	10,444	
tertiary_link	0.55	0.33	0.42	427	
trunk	0.62	0.45	0.52	1210	
trunk_link	0.59	0.45	0.51	1116	
Weighted	0.80	0.81	0.79	80,690	

Looking at Table 6, which outlines the performance using only centrality-based features, we observe that the RF model performed moderately well, especially for *residential* roads, with a precision of 0.82 and recall of 0.97. Also, it predicted *motorway* roads better than the usage of only semantic-based features. However, the model's effectiveness varied for other road types. For instance, the precision for *motorway_links* was 0.49 with a recall of 0.25, and other link roads had a maximum recall of 0.08. This variability suggests that, while centrality features contribute to the model's predictive capability, they may not be as reliable when used in isolation, particularly for less common road types.

Table 6. The results RF in prediction by only centrality-based features.

Key Value	Precision	Recall	F1-Score	Object Count	Area of Interest
motorway	0.78	0.73	0.75	876	All cases without a city identifier
motorway_link	0.49	0.25	0.33	971	
primary	0.57	0.39	0.46	3366	
primary_link	0.37	0.04	0.08	895	
residential	0.82	0.97	0.88	55,912	
secondary	0.47	0.26	0.33	4911	
secondary_link	0.55	0.02	0.04	562	
tertiary	0.49	0.33	0.39	10,444	
tertiary_link	0.67	0.01	0.02	427	
trunk	0.69	0.48	0.56	1210	
trunk_link	0.48	0.08	0.14	1116	
Weighted	0.72	0.76	0.73	80,690	

Comparing these results to those in Tables 2 and 3, we can see that the model achieves the best performance when it utilizes a combination of city, geometry-, semantic-, and centrality-based features. The weighted scores in Tables 2 and 3 are generally higher than those in Tables 4–6, which demonstrates the benefits of a multi-feature approach over single-feature models.

The overall findings indicate that, while there is potential in using semantic- and centrality-based features for road type classification, there is still a need for (1) fine-tuned existing features or (2) new features to better capture the complexities of road type classification in OSM data. Our first approach to increasing the RF classifier's performance was using K-means clusters of the lowest features, which were *SN* and *L* in the geometric-based features. The number of clusters was determined to be 3, which is supported by the Elbow graph (Figure 5). Table 7 is a modified version of Table 3, using geometric-based clusters instead of original features. Although the *motorway* and *primary* road prediction performance experienced little increases, there were areas where the introduction of geometric-based clusters did not lead to an increase. For instance, the recall values for *primary_link*, *secondary_link*, and *tertiary_link* roads remained low, which implies that the model still struggled to identify all instances of these road types. Secondly, we used all features as K-means clusters in their respective measure types in Table 1. We used three clusters for each feature group (geometric, semantic, and centrality). However, none of the alternatives yielded improved results. The F1-scores of the RF classifications were 81%, 76%, and 76% when each feature group of geometry-, semantic-, and centrality-based measures was clustered, respectively. Some of the measures, like *SN* and distance to amenities, were grouped into a quantile interval. Nonetheless, the performance remained on par with the model utilizing clustering techniques, suggesting no tangible benefit from this level of measure categorization. The underperformance of measure categorization could be attributed to the intrinsic mechanics of the RF classifier. RF is inherently adept at handling continuous variables and exploiting the subtle variances within them to make decisions. By categorizing the measures, we may inadvertently obscure these nuances, reducing the algorithm's ability to leverage the full spectrum of information provided by the raw data. The categorization may also introduce artificial boundaries that do not align

with the natural distributions of the data, leading to a loss of valuable context that could aid in classification.

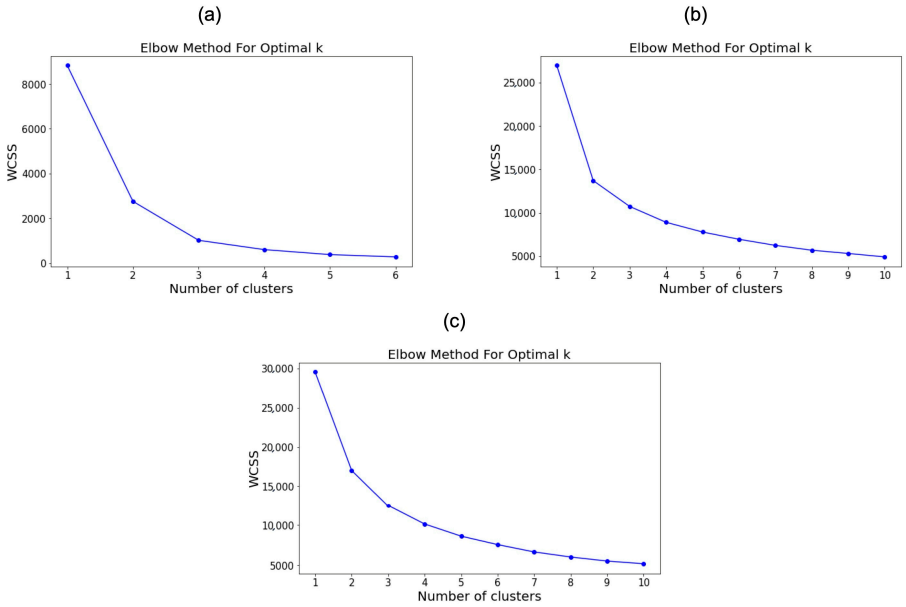

Figure 5. Elbow graphs of geometry- (**a**), semantic- (**b**), and centrality-based (**c**) features.

Table 7. The results in RF prediction by a city identifier, semantic- and centrality-based features, and geometric-based clusters.

Key Value	Precision	Recall	F1-Score	Object Count	Area of Interest
motorway	0.89	0.88	0.89	876	All cases with a city identifier
motorway_link	0.76	0.47	0.58	971	
primary	0.77	0.62	0.68	3366	
primary_link	0.65	0.17	0.27	895	
residential	0.85	0.98	0.91	55,912	
secondary	0.73	0.48	0.58	4911	
secondary_link	0.81	0.09	0.16	562	
tertiary	0.72	0.47	0.57	10,444	
tertiary_link	0.72	0.07	0.12	427	
trunk	0.88	0.70	0.78	1210	
trunk_link	0.75	0.31	0.44	1116	
Weighted	0.82	0.83	0.81	80,690	

3.3. Evaluation of the Prediction Results After Feature Augmentation

We increased the number of features in each city by using the network-based hierarchical feature augmentation approach detailed in Section 2.2. The RF classifier's performance in the feature enrichment process is reflected in the F1-score matrices of Tables 8–10 with pairwise feature sets. Firstly, as we incorporated higher levels of surroundings for geometry (G) and centrality (CN) features, the F1-scores improved without using semantic-based features and a city identifier (Table 8). We observed the highest scores when the level of the surroundings reached 50 for both types of features (G-50, CN-50), with an F1-score of 83%.

Table 8. Prediction results using geometry- and centrality-based features.

F1-Score Matrix (%)	Without CN	CN-0	CN-5	CN-10	CN-25	CN-50
Without G	-	73	77	78	80	81
G-0	60	74	78	79	81	82
G-5	65	75	79	80	81	82
G-10	65	75	80	80	81	82
G-25	66	76	80	81	82	83
G-50	67	76	81	82	83	83

G, level of the surroundings in geometry-based features; CN, level of the surroundings in centrality-based features.

Table 9. Prediction results using geometry- and semantic-based features.

F1-Score Matrix (%)	Without S	S-0	S-5	S-10	S-25	S-50
Without G	-	79	79	79	79	80
G-0	60	74	77	78	79	80
G-5	65	71	76	77	79	80
G-10	65	71	76	77	78	80
G-25	66	71	76	77	79	80
G-50	67	72	76	78	79	80

G, level of the surroundings in geometry-based features; S, level of the surroundings in semantic-based features.

Table 10. Prediction using semantic- and centrality-based features.

F1-Score Matrix (%)	Without CN	CN-0	CN-5	CN-10	CN-25	CN-50
Without S	-	73	77	78	80	81
S-0	79	80	80	81	81	82
S-5	79	82	83	83	83	83
S-10	79	83	83	83	83	83
S-25	79	84	84	84	84	84
S-50	80	84	85	84	84	84

CN, level of the surroundings in centrality-based features; S, level of the surroundings in semantic-based features.

Table 9, focusing on only geometry- and semantic-based features, also shows an upward trajectory in F1-scores with increasing levels of surroundings. Notably, when the level of the surroundings in semantic features (S) was maximized to 50 (S-50), alongside a high level of geometry-based surroundings (G-50), the model achieved an F1-score of 80%. Table 10 offers the most compelling evidence for the effectiveness of feature-level augmentation in the predictive model. The model achieves its highest performance, with an F1-score of 85%, when semantic-based features are expanded to the maximum level of 50 (S-50) and centrality-based features are enriched to just a level of 5 (CN-5). This targeted strategy leads to a more efficient and effective predictive model, emphasizing that the road to optimization may require a balance rather than a uniform maximization of feature levels. As a result, this particular combination suggests that a comprehensive semantic context, paired with a moderate level of centrality context, substantially increases the model's performance in predicting road types within the OSM data.

If we were to follow the trend observed in previous tables, where increased levels of feature aggregation generally led to improved predictive performance, we might hypothesize that higher levels of geometric features (e.g., G-25, G-50) would further maximize the model's performance. However, the analysis of Table 11's results shows that adding the geometric context to the already robust S-50 and CN-5 feature set, with or without a city identifier, does not provide any change in the model's performance.

Table 11. Prediction using all levels of geometric-based features with Level 50 semantic-based (S-50) and Level 5 centrality-based (CN-5) features.

	Weighted Precision	Weighted Recall	Weighted F1-Score	Object Count	Area of Interest
Without G	85	86	85	80,690	All cases with or without a city identifier
G-0					
G-5					
G-10					
G-25					
G-50					

G, level of the surroundings in geometry-based features.

Table 12 shows that, for cases without a city identifier, the model with the S-50 and CN-5 combination performs commendably across various road types. Notably, the precision and recall for *motorway_link* roads are higher than any previous testing, indicating that the model can reliably identify these road types. *residential* roads also had very high precision and recall values, likely due to their prevalence in the dataset, which allows the model to learn from a large number of examples. For *tertiary_link* and *secondary_link*, the model exhibits better but not ideal performance, which could be an area of focus for future improvements.

Table 12. The results in prediction by S-50 and CN-5 features.

Key Value	Precision	Recall	F1-Score	Object Count	Area of Interest
motorway	0.85	0.90	0.87	876	All cases without a city identifier
motorway_link	0.78	0.77	0.78	971	
primary	0.76	0.68	0.72	3366	
primary_link	0.70	0.45	0.55	895	
residential	0.88	0.98	0.93	55,912	
secondary	0.75	0.57	0.65	4911	
secondary_link	0.67	0.31	0.43	562	
tertiary	0.80	0.53	0.64	10,444	
tertiary_link	0.73	0.28	0.40	427	
trunk	0.81	0.70	0.75	1210	
trunk_link	0.74	0.65	0.69	1116	
Weighted	85	86	85	80,690	

In Table 13, where city identifiers are included alongside the S-50 and CN-5 features, the precision and recall for most road types are slightly improved or remain stable compared with Table 12. For example, *motorway* roads show a slight increase in recall, and motorway links maintain high precision and recall. *primary_link* and *secondary_link* roads see a slight improvement in precision, which might suggest that smaller city identifiers (e.g., district or town boundary files) contribute to the model's ability to discern these road types better. The overall weighted performance remains consistent with that in Table 12. As a result, the model has reached a peak where the inclusion of additional geometric context or a city identifier does not contribute additional discriminative power for the prediction task.

Table 13. The results in prediction by a city identifier and S-50 and CN-5 features.

Key Value	Precision	Recall	F1-Score	Object Count	Area of Interest
motorway	0.84	0.91	0.88	876	All cases with a city identifier
motorway_link	0.79	0.78	0.78	971	
primary	0.77	0.68	0.72	3366	
primary_link	0.72	0.46	0.56	895	
residential	0.89	0.98	0.93	55,912	
secondary	0.75	0.57	0.65	4911	
secondary_link	0.71	0.31	0.43	562	
tertiary	0.79	0.54	0.64	10,444	
tertiary_link	0.77	0.29	0.42	427	
trunk	0.83	0.72	0.77	1210	
trunk_link	0.74	0.63	0.68	1116	
Weighted	85	86	85	80,690	

The heatmap in Figure 6 provides a comparative analysis of F1-scores across different cities, revealing distinct patterns in model performance using only the respective city's samples for training. Each cell represents the F1-score of the prediction test within each city, with the color intensity indicating the level of performance. This visualization highlights the comparative strengths and weaknesses of the model's performance in different urban contexts. In other words, the RF classification results offer a window into the model's variable performance across different cities, revealing how urban-specific characteristics influence predictive accuracy. Milan and İzmir stand out with the highest accuracy and weighted F1-score at 89% and 88%, respectively, showcasing the model's strong predictive capabilities in these environments. Particularly in Milan, the results across various road classes are consistently satisfying, indicating a well-trained model for the city's road types. A consistent pattern emerges where residential and motorway roads achieve the highest predictive accuracy in each city, a testament to the robustness of the model for these categories. Conversely, link roads exhibit lower predictive performance, suggesting that these might be more challenging to classify due to their varied and less consistent representation in the data. Interestingly, the model more accurately predicts *motorway_link*, *primary_link*, and *trunk_link* roads compared with other link types. This trend suggests that the more significant (major) a link road is within the road network hierarchy, the more accurately the model can predict its functional class. As a result, the study underscores the comparative strengths and weaknesses of the model, highlighting the impact of our model's performance on different urban areas.

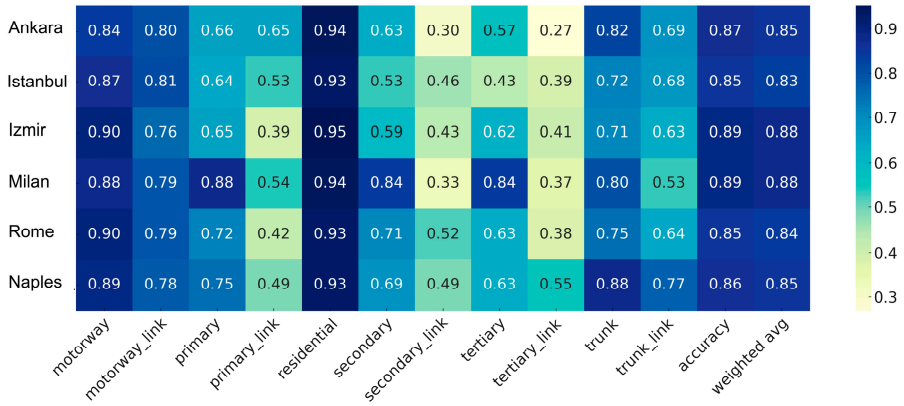

Figure 6. Heatmap displaying the F1-scores across six cities.

4. Discussion

The results of this study highlight the effectiveness and challenges of predicting road functional classes in OSM using a machine learning approach enhanced with network-

based hierarchical feature augmentation. The findings are discussed below in relation to implications for OSM data enrichment, the model's feature engineering, and performance variations across road classes and cities.

The findings of this study present significant implications for improving the completeness and accuracy of OSM data. The proposed hierarchical feature augmentation approach provides a scalable framework for enriching OSM road tags, making the dataset more reliable for applications such as urban planning, navigation systems, and disaster management. The ability to predict road classes with an average F1-score of 85%, without relying on geometric properties, demonstrates the potential of this method to contribute missing semantic information in OSM data.

The evaluation of the prediction results obtained before augmenting the features revealed key strengths and limitations of the model. The RF classifier performed exceptionally well in predicting *residential* roads, with F1-scores exceeding 90%, due to the prevalence of these roads and their distinct features. Similarly, major road types such as *motorway* and *trunk* roads were classified with high precision and recall, reflecting the reliability of the feature set for these categories. However, the model struggled with link roads (e.g., *secondary_link* and *tertiary_link*), as evidenced by lower recall scores. These road types often lack consistent tagging or distinctive features, making them harder to classify. The inclusion of city identifiers marginally improved the performance for some road types, suggesting that geographic context plays a role in improving classification accuracy.

The inclusion of hierarchical levels, as detailed in Section 3.3, further enriched the feature set by representing both local and global network characteristics. The highest F1-scores were achieved when semantic-based features were expanded to Level 50 (S-50) and centrality-based features to Level 5 (CN-5), underscoring the importance of balancing feature complexity and computational efficiency. The model's performance in predicting road types across six cities highlights notable differences in prediction accuracy, with cities such as Milan and Izmir achieving the highest weighted F1-scores (88%), indicating that the feature level combination tested with the whole dataset is also compatible with specific study areas (Izmir and Milan). Also, it seems that there may be a better combination for Istanbul, where the S-50 and CN-5 combination performed at 83%. Residential roads consistently showed strong performance across all cities, reflecting their prevalent tagging and distinct characteristics in the dataset. Conversely, *secondary_link* and *tertiary_link* roads demonstrate lower scores, particularly in Ankara, where the less consistent tagging and variations in the representation pose challenges for classification [42]. These findings underline the model's effectiveness in densely populated urban areas and its limitations in handling underrepresented road types, emphasizing the importance of data completeness and quality across diverse urban contexts for future optimization.

There are some minor differences among the prediction performances using only samples of the city of interest for training. However, the model's hierarchical, network-based approach seems to present geographic and morphological distinctions, potentially making explicit city information redundant once it is trained with the random data in all cities. This could account for the observed insubstantial impact of the city identifier on the model's performance, as the network features might already represent spatial relationships and contextual properties inherent to specific urban settings.

Despite its success in general prediction performance, this study has certain limitations. The model's reliance on existing OSM data may introduce biases in regions with fewer contributors or incomplete tagging since the proposed semantic features are computed from the intrinsic *amenity* tags contributed by volunteers. Additionally, the challenges in predicting link roads highlight the need for further refinement in feature engineering.

5. Conclusions

Our conclusions align with the principle of parsimony in model building, where the goal is to achieve the maximum predictive power with the simplest model possible, avoiding overfitting and ensuring that the model is as generalizable as possible. The

predictive model has reached a performance threshold where the further inclusion of geometric data or city identifiers does not significantly optimize the predictive accuracy. This plateau suggests that the combination of semantic- and centrality-based features at certain levels already captures essential classification information, rendering additional geometric detail superfluous. This discovery is novel, as it utilizes a simplified model structure, minimizing unnecessary complexity and computational demands. The prediction performance also highlights the model's reliability, particularly for well-defined road types such as *residential*, *motorway*, *trunk*, and *primary* roads. Also, the use of intrinsic road network characteristics, including centrality measures and semantic features, significantly improved the prediction accuracy by capturing the functional and structural distinctions within road networks.

Our feature augmentation approach creates a hierarchical network around each road, aggregating features from directly connected, secondarily connected, and sequentially graded connected road lines. These aggregated features likely provide a rich set of data (N+1 times more than Level 0) that reflects not just the properties of the individual road but also the characteristics of the Level N surrounding road network.

Future studies will involve the model's refinement, exploring diverse feature level combinations with new centrality- and semantic-based features, and maintaining an automated, iterative process to determine the best level of the network hierarchy per feature set. Moreover, understanding the influence of volunteer contributions on the model's accuracy could provide a new step for improving data consistency and reliability. In addition, answering the question of how the quality of the data contributed by OSM volunteers impacts the performance of this prediction model will be important to visualize the complex behavior of volunteers. Finally, integrating the proposed approach into an AI-based assistant tool within the OSM platform could assist contributors in accurately assigning semantic values when adding or editing road line objects that lack road class information.

Author Contributions: Conceptualization, Müslüm Hacar and Valerio Cutini; methodology, Müslüm Hacar and Diego Altafini; software, Müslüm Hacar and Diego Altafini; validation, Müslüm Hacar, Diego Altafini and Valerio Cutini; formal analysis, Müslüm Hacar; investigation, Müslüm Hacar, Diego Altafini and Valerio Cutini; resources, Müslüm Hacar; data curation, Müslüm Hacar and Diego Altafini; writing—original draft preparation, Müslüm Hacar; writing—review & editing, Müslüm Hacar, Diego Altafini and Valerio Cutini; visualization, Müslüm Hacar; supervision, Valerio Cutini; project administration, Müslüm Hacar, Diego Altafini and Valerio Cutini; funding acquisition, Müslüm Hacar, Diego Altafini and Valerio Cutini. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by The Scientific and Technological Research Council of Türkiye (TÜBİTAK), grant number 1059B192201248, and the United Kingdom Research and Innovation Post Doctoral Fellowship Guarantee Scheme set over the European Union's Horizon Europe—Marie Skłodowska Curie Actions Post Doctoral Fellowships, UKRI grant number 101107846-DECIDE/EP/Y028716/1. The views and opinions expressed are those of the authors only, and do not necessarily reflect those of the United Kingdom or the European Union. Neither the United Kingdom, the European Union, nor the granting authority can be held responsible for them.

Data Availability Statement: OpenStreetMap data were downloaded from Geofabrik's web page (http://download.geofabrik.de/ (accessed on 4 November 2022)).

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

References

1. Haklay, M.; Weber, P. OpenStreetMap: User-Generated Street Maps. *IEEE Pervasive Comput.* **2008**, *7*, 12–18. [CrossRef]
2. Mooney, P.; Minghini, M. A review of OpenStreetMap data. In *Mapping and the Citizen Sensor*; Ubiquity Press: London, UK, 2017; pp. 37–59. [CrossRef]
3. Basiri, A.; Amirian, P.; Mooney, P. Using crowdsourced trajectories for automated OSM data entry approach. *Sensors* **2016**, *16*, 1510. [CrossRef] [PubMed]
4. Vargas Munoz, J.E.; Tuia, D.; Falcão, A.X. Deploying machine learning to assist digital humanitarians: Making image annotation in OpenStreetMap more efficient. *Int. J. Geogr. Inf. Sci.* **2021**, *35*, 1725–1745. [CrossRef]

5. OpenStreetMap Wiki. (n.d.) Map Features. Available online: https://wiki.openstreetmap.org/wiki/Map_features (accessed on 15 December 2024).
6. Porta, S.; Crucitti, P.; Latora, V. The network analysis of urban streets: A dual approach. *Phys. A* **2006**, *369*, 853–866. [CrossRef]
7. Porta, S.; Crucitti, P.; Latora, V. The network analysis of urban streets: A primal approach. *Environ. Plan. B Plan. Des.* **2006**, *33*, 705–725. [CrossRef]
8. Batty, M. The size, scale, and shape of cities. *Science* **2008**, *319*, 769–771. [CrossRef] [PubMed]
9. Jiang, B. Street hierarchies: A minority of streets account for a majority of traffic flow. *Int. J. Geogr. Inf. Sci.* **2009**, *23*, 1033–1048. [CrossRef]
10. Ugalde, C.; Braga, A.; Altafini, D. Spatial reconfigurations within flood-prone areas: The case of Porto Alegre Metropolitan Region—Brazil. In Proceedings of the 13th Space Syntax Symposium, Bergen, Norway, 20–24 June 2022.
11. Pezzica, C.; Altafini, D.; Cutini, V. Urban-regional dynamics of street network resilience: The spatial outcomes of Genoa's and Bologna's bridge crashes. In Proceedings of the 13th Space Syntax Symposium, Bergen, Norway, 20–24 June 2022.
12. Zielstra, D.; Zipf, A. A comparative study of proprietary geodata and volunteered geographic information for Germany. In Proceedings of the 13th AGILE International Conference on Geographic Information Science, Guimarães, Portugal, 10–14 May 2010.
13. Sehra, S.S.; Singh, J.; Rai, H.S. Assessing OpenStreetMap data using intrinsic quality indicators: An extension to the QGIS processing toolbox. *Future Internet* **2017**, *9*, 15. [CrossRef]
14. Jilani, M.; Corcoran, P.; Bertolotto, M. Automated highway tag assessment of OpenStreetMap road networks. In Proceedings of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, Dallas/Fort Worth, TX, USA, 4–7 November 2014; pp. 449–452.
15. Meng, S.; Zheng, H. A personalized bikeability-based cycling route recommendation method with machine learning. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *121*, 103373. [CrossRef]
16. Haklay, M. How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. *Environ. Plan. B* **2010**, *37*, 682–703. [CrossRef]
17. Girres, J.F.; Touya, G. Quality assessment of the French OpenStreetMap dataset. *Trans. GIS* **2010**, *14*, 435–459. [CrossRef]
18. Mondzech, J.; Sester, M. Quality analysis of OpenStreetMap data based on application needs. *Cartographica* **2011**, *46*, 115–125. [CrossRef]
19. Da Costa, J.N. Novel tool for examination of data completeness based on a comparative study of VGI data and official building datasets. *Geod. Vestn.* **2016**, *60*, 495–508. [CrossRef]
20. Zhang, H.; Malczewski, J. Accuracy Evaluation of the Canadian OpenStreetMap Road Networks. *Int. J. Geospat. Environ. Res.* **2018**, *5*, 1–14. Available online: https://dc.uwm.edu/ijger/vol5/iss2/1/ (accessed on 20 October 2021).
21. Davidovic, N.; Mooney, P.; Stoimenov, L. An analysis of tagging practices and patterns in urban areas in OpenStreetMap. In Proceedings of the AGILE 2016 Conference, Helsinki, Finland, 14–17 June 2016.
22. Jilani, M.; Corcoran, P.; Bertolotto, M. Multi-granular street network representation towards quality assessment of OpenStreetMap data. In Proceedings of the Sixth ACM SIGSPATIAL International Workshop on Computational Transportation Science, Orlando, FL, USA, 5–8 November 2013; pp. 19–24.
23. Corcoran, P.; Jilani, M.; Mooney, P.; Bertolotto, M. Inferring semantics from geometry: The case of street networks. In Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems, Bellevue, WA, USA, 3–6 November 2015; pp. 1–10.
24. Keller, S.; Gabriel, R.; Guth, J. Machine learning framework for the estimation of average speed in rural road networks with OpenStreetMap data. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 638. [CrossRef]
25. Hacar, M. Using geometric and semantic attributes for semi-automated tag identification in OpenStreetMap data. In Proceedings of the GISRUK 2021, Cardiff, UK, 14–16 April 2021. [CrossRef]
26. Alghanim, A.; Jilani, M.; Bertolotto, M.; McArdle, G. Leveraging road characteristics and contributor behaviour for assessing road type quality in OSM. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 436. [CrossRef]
27. OSMstats. (n.d.) Statistics of the Free Wiki World Map. Available online: https://osmstats.neis-one.org/ (accessed on 15 December 2024).
28. Taginfo. (n.d.) OpenStreetMap Taginfo. Available online: https://taginfo.openstreetmap.org/ (accessed on 15 December 2024).
29. Pazoky, S.H.; Pahlavani, P. Developing a multi-classifier system to classify OSM tags based on centrality parameters. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *104*, 102595. [CrossRef]
30. Juhász, L.; Mooney, P.; Hochmair, H.H.; Guan, B. ChatGPTas a mapping assistant: A novel method to enrich maps with generativeAI and content derived from street-level photographs. In Proceedings of the Fourth Spatial Data Science Symposium, distributed/online symposium, 5–6 September 2023. [CrossRef]
31. Yang, J.; Jang, H.; Yu, K. Geographic Knowledge Base Question Answering over OpenStreetMap. *ISPRS Int. J. Geo-Inf.* **2024**, *13*, 10. [CrossRef]
32. Hochmair, H.H.; Juhász, L.; Kemp, T. Correctness Comparison of ChatGPT-4, Gemini, Claude-3, and Copilot for Spatial Tasks. *Trans. GIS* **2024**, *28*, 2219–2231. [CrossRef]
33. Buhl, J.; Gautrais, J.; Reeves, N.; Solé, R.V.; Valverde, S.; Kuntz, P.; Theraulaz, G. Topological patterns in street networks of self-organized urban settlements. *Eur. Phys. J. B* **2006**, *49*, 513–522. [CrossRef]

34. Crucitti, P.; Latora, V.; Porta, S. Centrality measures in spatial networks of urban streets. *Phys. Rev. E* **2006**, *73*, 036125. [CrossRef] [PubMed]
35. Jiang, B. A topological pattern of urban street networks: Universality and peculiarity. *Phys. A* **2007**, *384*, 647–655. [CrossRef]
36. Barthélemy, M.; Flammini, A. Modeling urban street patterns. *Phys. Rev. Lett.* **2008**, *100*, 138702. [CrossRef]
37. Masucci, A.P.; Smith, D.; Crooks, A.; Batty, M. Random planar graphs and the London street network. *Eur. Phys. J. B* **2009**, *71*, 259–271. [CrossRef]
38. Masucci, A.P.; Stanilov, K.; Batty, M. Limited urban growth: London's street network dynamics since the 18th century. *PLoS ONE* **2010**, *8*, e69469. [CrossRef] [PubMed]
39. Strano, E.; Nicosia, V.; Latora, V.; Porta, S.; Barthélemy, M. Elementary processes governing the evolution of road networks. *Sci. Rep.* **2012**, *2*, 296. [CrossRef] [PubMed]
40. Corcoran, P.; Mooney, P.; Bertolotto, M. Analysing the growth of OpenStreetMap networks. *Spatial Stat.* **2013**, *3*, 21–32. [CrossRef]
41. Zhao, P.; Jia, T.; Qin, K.; Shan, J.; Jiao, C. Statistical analysis on the evolution of OpenStreetMap road networks in Beijing. *Phys. A* **2015**, *420*, 59–72. [CrossRef]
42. Hacar, M.; Kılıç, B.; Şahbaz, K. Analyzing OpenStreetMap road data and characterizing the behavior of contributors in Ankara, Turkey. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 400. [CrossRef]
43. Freeman, L.C. A set of measures of centrality based upon betweenness. *Sociometry* **1977**, *40*, 35–41. [CrossRef]
44. Brandes, U. On variants of shortest-path betweenness centrality and their generic computation. *Soc. Netw.* **2008**, *30*, 136–145. [CrossRef]
45. Turner, A. Angular analysis. In Proceedings of the 3rd International Space Syntax Symposium, Georgia Institute of Technology, Atlanta, GA, USA, 1 January 2001.
46. Turner, A. *Depthmap 4: A Researcher's Handbook*; Bartlett School of Graduate Studies, University College: London, UK, 2004.
47. depthmapX Development Team. (n.d.) depthmapX (Version 0.8.0) [Computer Software]. Available online: https://github.com/SpaceGroupUCL/depthmapX/ (accessed on 15 December 2024).
48. Hillier, W.R.G.; Yang, T.; Turner, A. Normalising least angle choice in Depthmap-and how it opens up new perspectives on the global and local analysis of city space. *J. Space Syntax* **2012**, *3*, 155–193.
49. Hacar, M.; Mara, F.; Altafini, D.; Cutini, V. Preprocessing Open Data for Optimizing Estimation Times in Urban Network Analysis: Extracting, Filtering, Geoprocessing, and Simplifying the Road-Center Lines. In Proceedings of the 12th International Conference on Innovation in Urban and Regional Planning: INPUT 2023, L'Aquila, Italy, 6–8 September 2023.
50. Douglas, D.H.; Peucker, T.K. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Cartographica* **1973**, *10*, 112–122. [CrossRef]
51. QGIS. (n.d.) Firenze Version 3.28 LTR. 2022. Available online: http://www.qgis.org/en/site/index.html (accessed on 15 December 2024).
52. Jordahl, K.; Van den Bossche, J.; Wasserman, J.; McBride, J.; Gerard, J.; Fleischmann, M.; Tratner, J.; Perry, M.; Farmer, C.; Hjelle, G.A.; et al. Geopandas/Geopandas: v0.6.0. *Zenodo*. 2019. Available online: https://zenodo.org/records/3463125 (accessed on 15 December 2024).
53. Fleischmann, M. Momepy: Urban morphology measuring toolkit. *J. Open Source Softw.* **2019**, *4*, 1807. [CrossRef]
54. van Nes, A.; Yamu, C. *Introduction to Space Syntax in Urban Studies*; Springer: Cham, Switzerland, 2021; ISBN 978-3-030-59140-3.
55. Neis, P.; Zielstra, D. Recent developments and future trends in volunteered geographic information research: The case of OpenStreetMap. *Future Internet* **2014**, *6*, 76–106. [CrossRef]
56. OpenStreetMap Wiki. (n.d.) Welcome to OpenStreetMap. Available online: https://wiki.openstreetmap.org/wiki/Main_Page (accessed on 1 December 2024).
57. Geofabrik. (n.d.) Available online: https://www.geofabrik.de/data/download.html (accessed on 1 December 2024).
58. GADM. (n.d.) GADM: Global Administrative Areas. Available online: https://gadm.org/ (accessed on 1 December 2024).
59. Osmosis. (n.d.) Osmosis. Available online: https://wiki.openstreetmap.org/wiki/Osmosis (accessed on 1 December 2024).