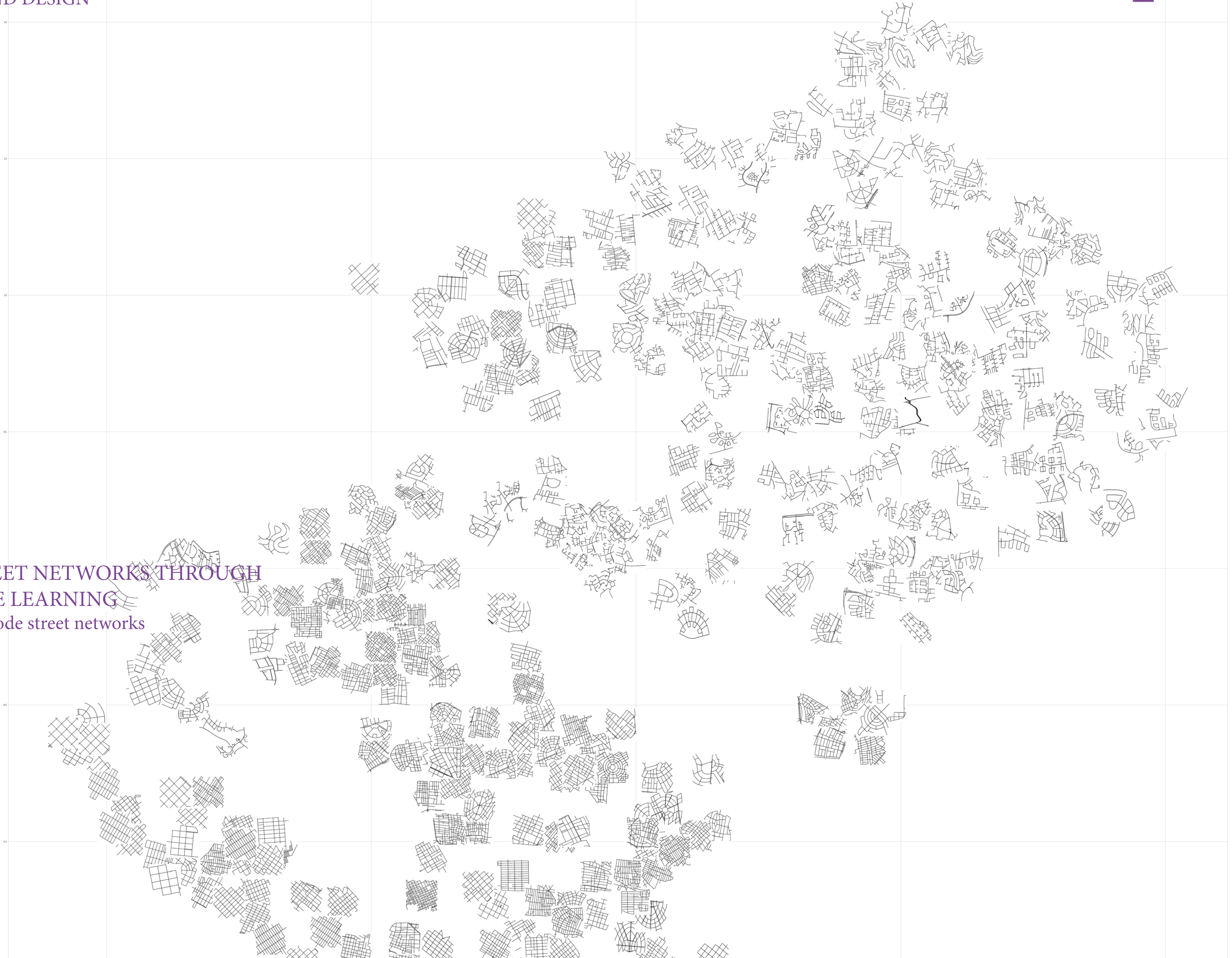
MASTER IN ADVANCED COMPUTATION FOR
ARCHITECTURE AND DESIGN



thesis studio

ANALYZING STREET NETWORKS THROUGH GRAPH MACHINE LEARNING

a study on how to encode street networks

2021-2022

MASTER IN ADVANCED COMPUTATION FOR
ARCHITECTURE AND DESIGN

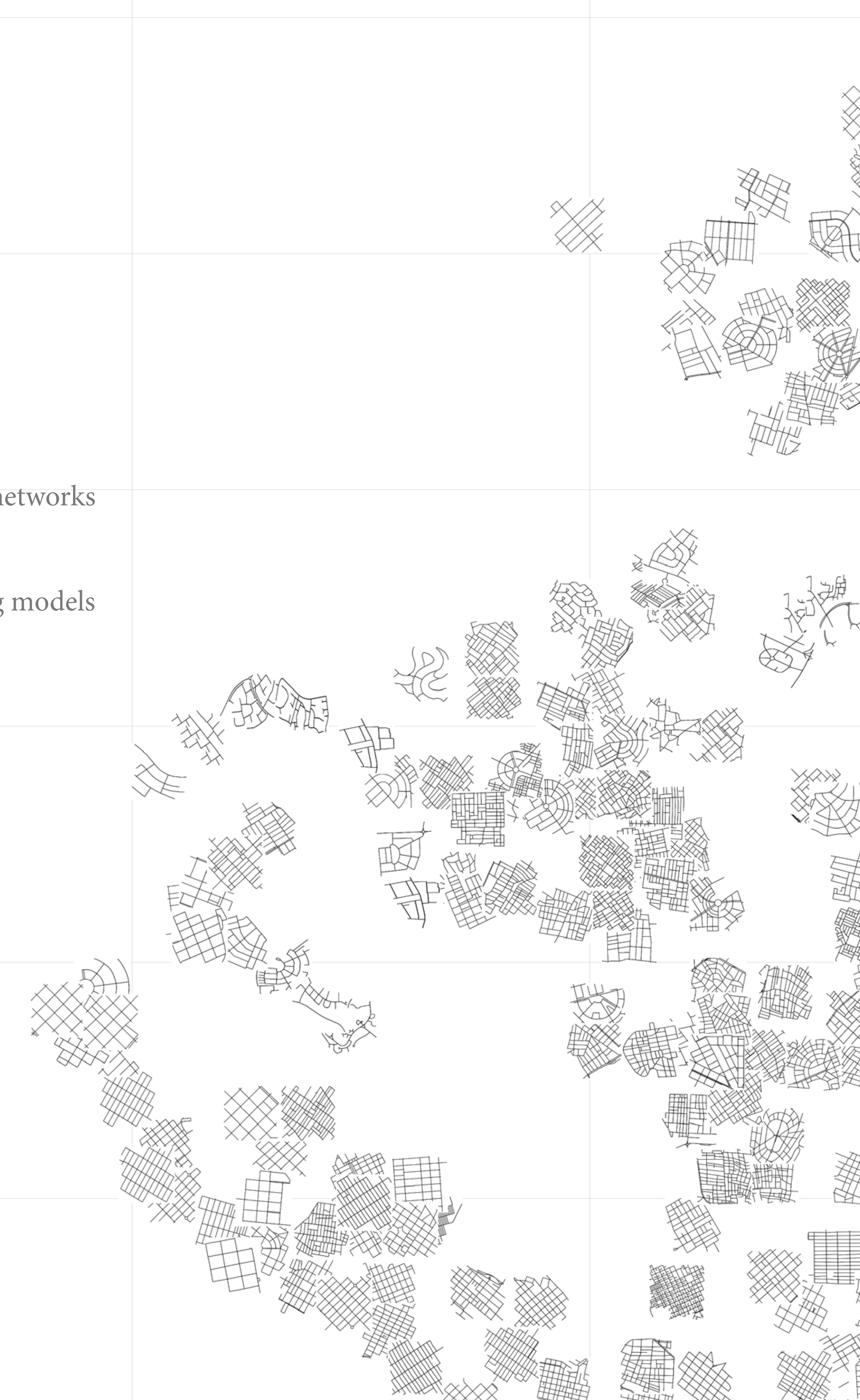
ANALYZING STREET NETWORKS THROUGH GRAPH MACHINE LEARNING

a study on how to encode street networks

Faculty:
DAVID ANDRES LEON

Student:
ERIDA BENDO

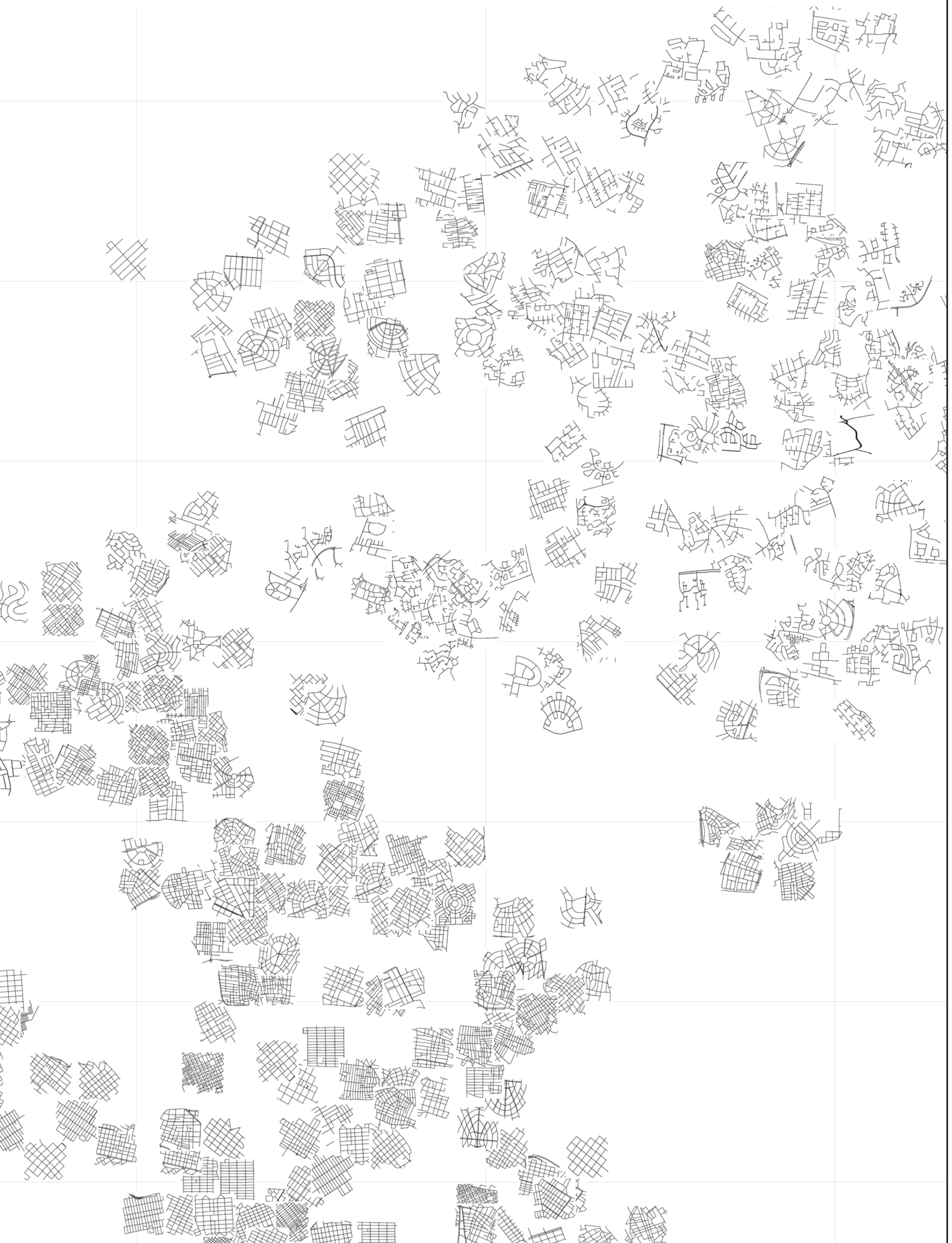
INDEX



ABSTRACT

Urban road networks are reflections of several environmental, social, and economic factors evolving with different speeds at different times. Sometimes these factors are linearly related and traceable, in other cases street patterns remain formal manifestations. This study represents an attempt of analyzing the geometric properties of street networks, through graph machine learning. Through unsupervised and supervised models, two ways of encoding them are tested: representing streets as graphs with the primal and dual approach. It concludes with some advantages and disadvantages of each encoding method and further opens a discussion on the prospects of using graph machine learning methods when analyzing or generating street patterns.

street networks, graph machine learning, encoding, primal/dual graphs



01. STATE OF THE ART

Research into street networks has gained prominence over the last fifty years[1], modeling them as graph data.

Graph neural networks have recently gained attention and have been applied to street networks only in a few studies.

1 (Marshall, et al. 2018)

01. STATE OF THE ART

ENCODING STREET NETWORKS AS GRAPHS

Street networks have commonly been treated as sets of linear elements, connecting locations and intersecting at junctions.[1]

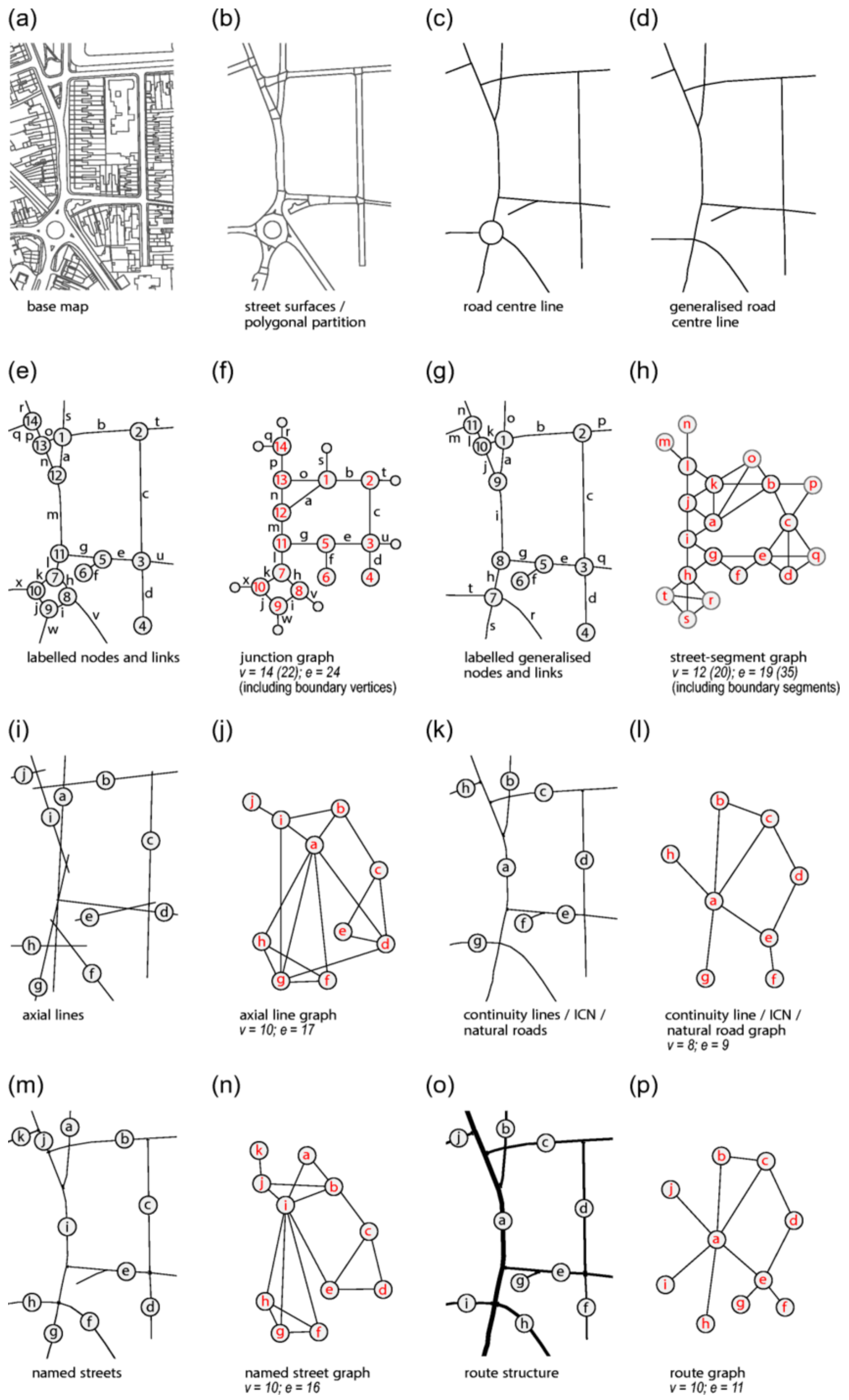
In mathematics, more specifically graph theory, a graph is a structure of relations between elements, the so called vertices. In a diagrammatic form, a graph can be depicted as a set of circles (vertices) connected by lines (edges). A large number of social, biological and man-made systems can be represented in the form of networks, such as: communities, the internet, transsportation systems etc. [2] Street networks, on the other hand have been discussed widely through graph theory, ever since Euler wrote about the Seven Bridges of Konigsberg in 1736.

It becomes clear that it is possible to model street networks as graphs, and building on the wealth of computational methods applicable to graph data, it can open new fronts of analysing them.

Graph theory can be used in the generalization of networks in two different ways: by deriving quantitative measures of topological or metrical properties of edges and nodes, or by identifing and representing important topological information which is needed for the effective application of generalization procedures.

Street networks can be modeled as graphs in two major categories: the primal and the dual approach. In the first one, quite intuitively, vertices represent junctions and edges the street, while in the second one the opposite happens.

1 (Marshall, et al. 2018)
2 (Porta, Crucitti and Latora 2004)



Different graph representations
(Marshall, et al. 2018)

The primal approach is the most studied case, being the most simple way to capture one of the most crucial components of the geographic dimension: distance.[2]

The dual approach on the other hand offers the possibility of calculating a the connectivity of the graph topologically in a non Euclidean step-distance.

01. STATE OF THE ART

GRAPH MACHINE LEARNING

The field of street network analysis has not yet benefited from much of the recent work on graph machine learning.[1]

Graph machine learning is a branch of machine learning that deals with graph data. In simple terms, these methods seek to generate embeddings on the node/edge or graph level, so that it can be able to classify nodes, make link predictions or predict characteristics of entire graphs.

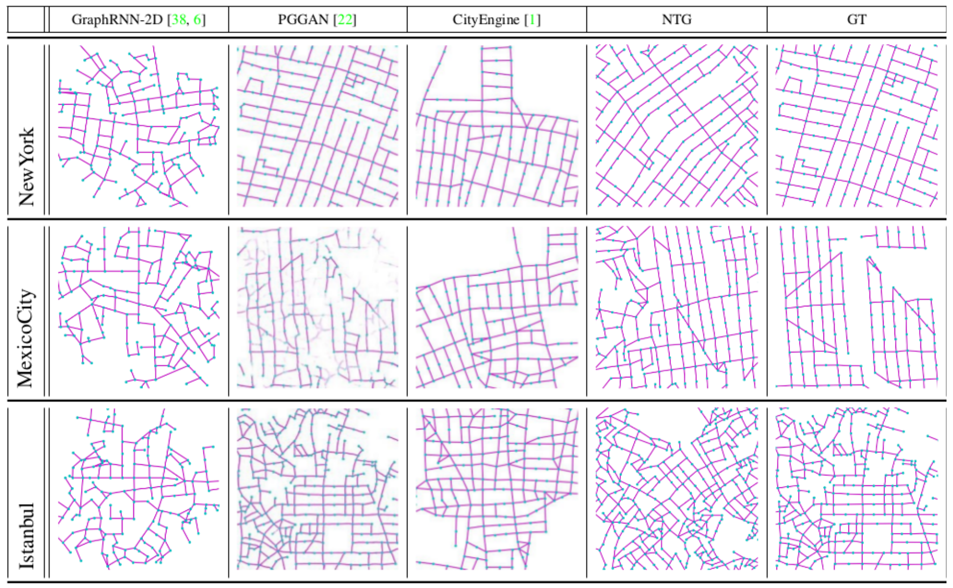
Studies dealing with graph machine learning and street networks mostly focus on the analysis of specific features.

In 'graph embeddings for street network analysis'[1], node embeddings are used for predicting traffic congestion. DeepMind is also making traffic predictions through graph neural networks.

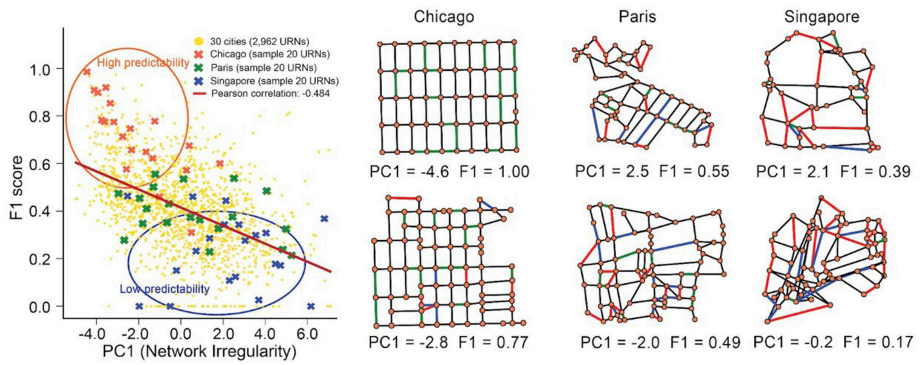
(Xue, et al. 2022)[2] were able to calculate an index, called the spatial homogeneity of road networks, in order to quantify the similarities of subnetworks in cities.

In the field of graph generation (Chu, et al. 2019) [3] were succesful into generating street networks that bear similarities to certain cities.

1 (DeMichele, Santos und Scheinfeld 2019)
2 Quantifying the spatial homogeneity of urban road networks via graph neural networks
3 Neural Turtle Graphics for Modeling City Road Layouts



City road layout generation[3]



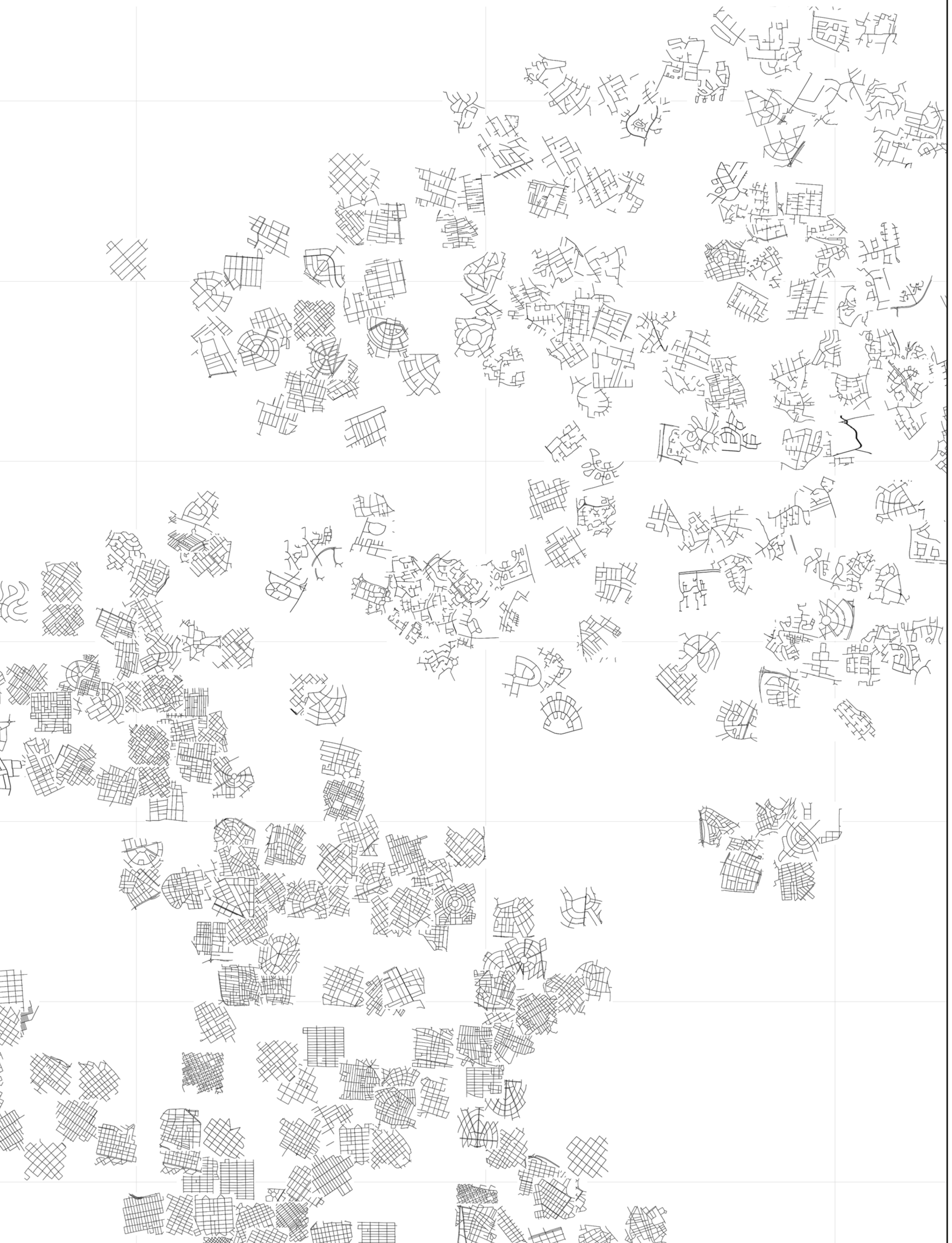
Spatial homogeneity of road networks[2]

(AlHalawani, et al. 2014)[1] provide another interesting study about understanding which factors play the most dominant roles for determining the character of a city. It doesn't work with graph data, but with images attaching to them topological and geometric features such as: street density, connectivity index, intersection density, 4 way crossing proportion etc.

(Kempinska und Murcio 2019)[2] and (Fang, Yang und Jin 2020)[3] also provide examples of analyzing city patterns through images.

Analyzing street networks through images however does not give in depth insights about these patterns.

1 What Makes London Work Like London?
2 Modelling urban networks using Variational Autoencoders
3 DeepStreet: A deep learning powered urban street network generation module



02. ENCODING STREET NETWORKS

While graphs are able to capture the topological qualities of street networks, its geographical features still need to be translated as features, either of nodes or edges.

02. ENCODING STREET NETWORKS

The purpose of this study is to analyze the geometrical qualities of street networks, thus the created dataset is composed of handpicked patterns which can be visually categorized into different typologies.

Based on the street network classification done by (Southworth und Ben-Joseph 1997)[1] around 400 graphs where chosen, labelled as either gridiron, fragmented parallel, warped parallel, loops and lollipops or lollipops on a stick.

It should be noted that choosing samples only visually is subjective and better ways of differentiating street networks in typologies might yield to better reasoning.

1. encoded as primal graphs.

(4 features per node/1 feature per edge)
Node (junction) features:

1.Average intersection angle
2.Maximum intersection angle
3.Minimum intersection angle
(Max/min angle proportion)
4.Length proportion -
(Max / min length of edges starting at the node)

Edge (street) feature:

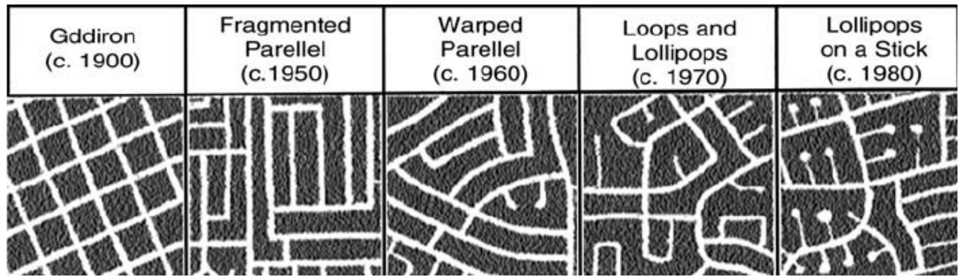
Length

2. encoded as dual graphs.

(2 features per node)
Node (street) features:

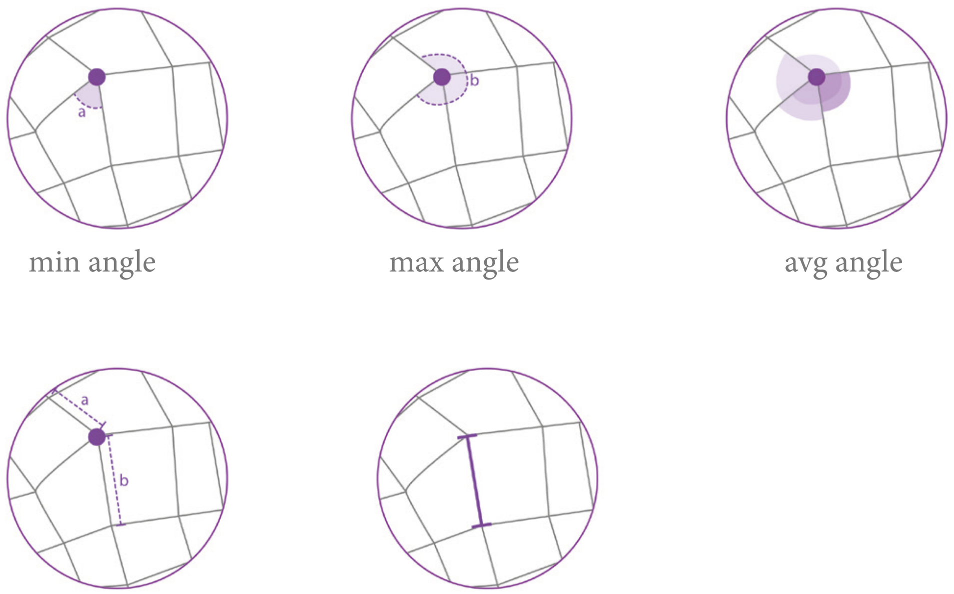
1.Length
2.Bearing

1 Streets and the Shaping of Towns and Cities



Street pattern typologies[1]

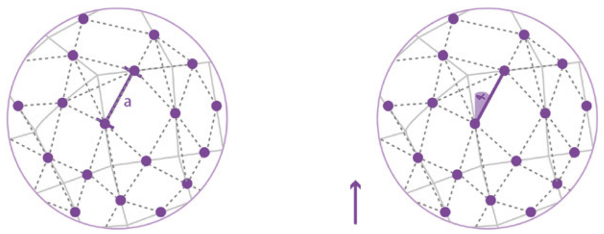
PRIMAL



min angle · max angle · avg angle

prop length · length

DUAL



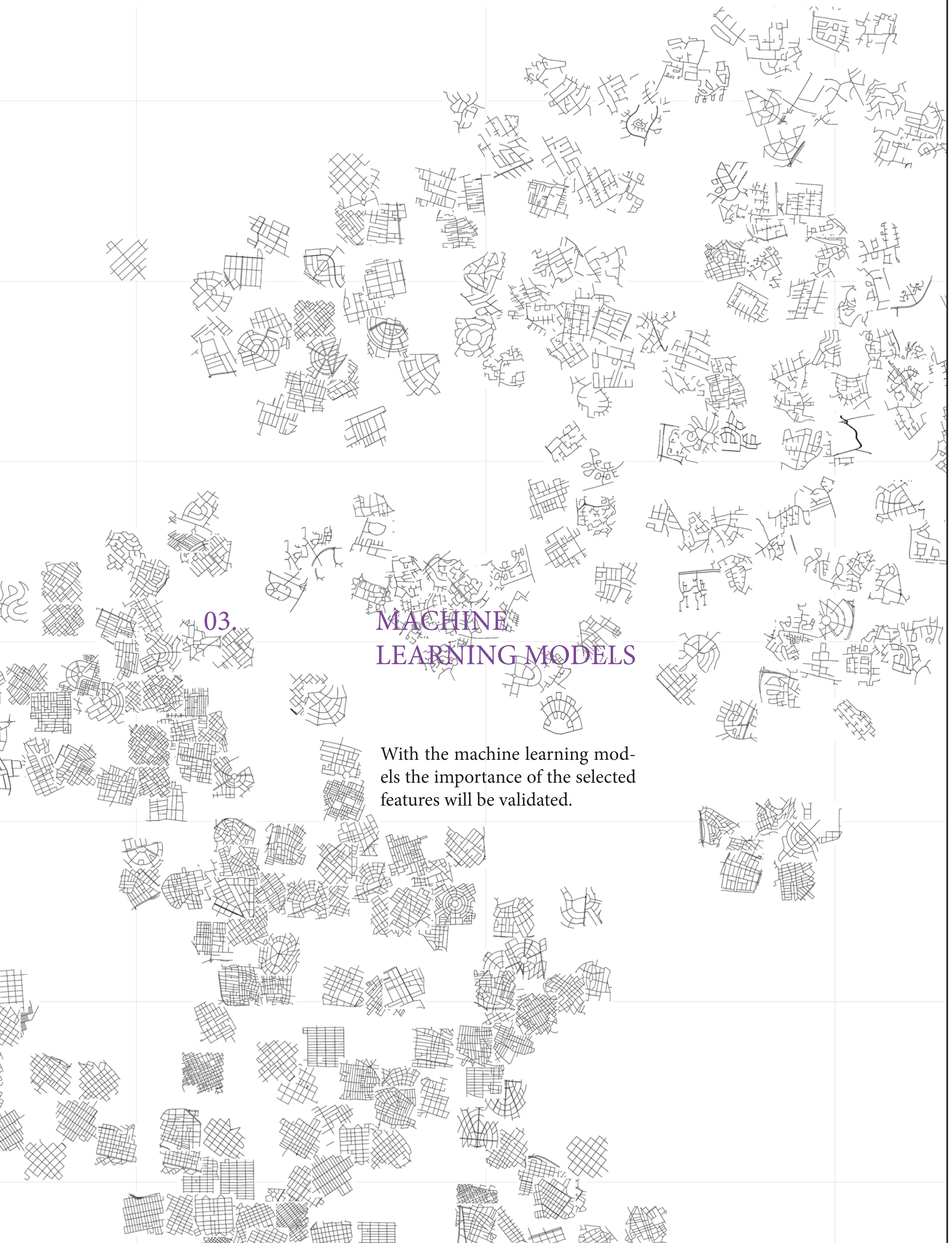
length · bearing

Notes:

The used machine learning models were tested on a smaller dataset with three scales, retrieving tiles of 600 m , 1200, 1800 m width and height dimension. The smaller one gave better results and thus tiles of 600 m were used for testing the final dataset.

a missing feature that might have improved the performance of recognizing the warped parallel typology would be straightness (ratio of length / aerial distance).

The publicly available road network data from OpenStreetMap (https://www.openstreetmap.org/) via the OSMnx Python package (https://github.com/gboeing/osmnx) was used for retrieving the street network samples.



03. MACHINE LEARNING MODELS

With the machine learning models the importance of the selected features will be validated.

03.1 UNSUPERVISED

Unsupervised models will be the first approach for testing our encoding strategy.

03.1 UNSUPERVISED

The chosen unsupervised model was based on the Unsupervised Inductive Graph-level Representation Learning via Graph-Graph Proximity by (Bai, et al. 2019). This model is able to generate graph level embeddings by training on a set of graphs, best preserving their graph-graph proximity. The researchers also state that the generated embeddings would be able to provide meaningful visualization in a two-dimensional space, besides using them for traditional tasks of graph classification and graph similarity/ distance computation[1].

In our case we have used an implementation of this model with the stellargraph library for tensorflow, where the used graph-graph proximity metric is the laplacian matrix.

Just like in the study we use the generated embeddings for visualizing the street patterns in a two dimensional space. As a second step we perform clustering to see if the model can guess the previously labelled typologies.

In the tsne visualization of the primal encoding scenario we can observe that the generated embeddings from this model are able to produce a meanigful representation. Besides the differences in their topology we can see that through the assigned features it is also able to capture subtle geometric nuances.

The model is not succesful in differentiating between fragmented parallel and warped parallel, so we can conclude that the selected features are not enough when dealing with curved patterns.

tsne visualization of the graphs without features

1 (Bai, et al. 2019)

tsne visualization primal

03.1 UNSUPERVISED

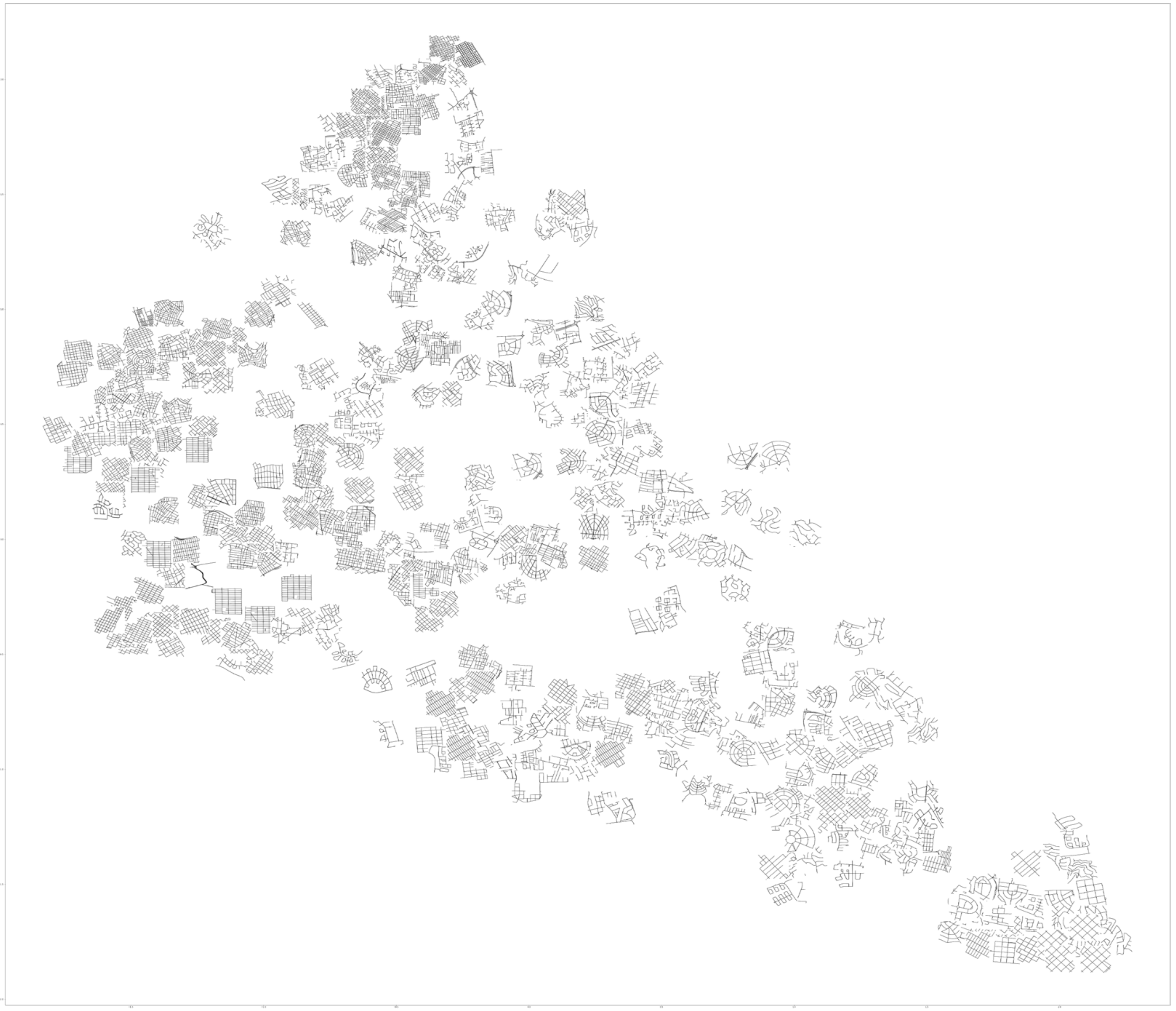
In the tsne visualization of the dual encoding scenario we can observe that the groups of the 5 topologies are not clearly distinguishable.

However we observe that street networks that are visually very similar are grouped together, concluding that this might be a succesful way of analayzing patterns of a similar topology.

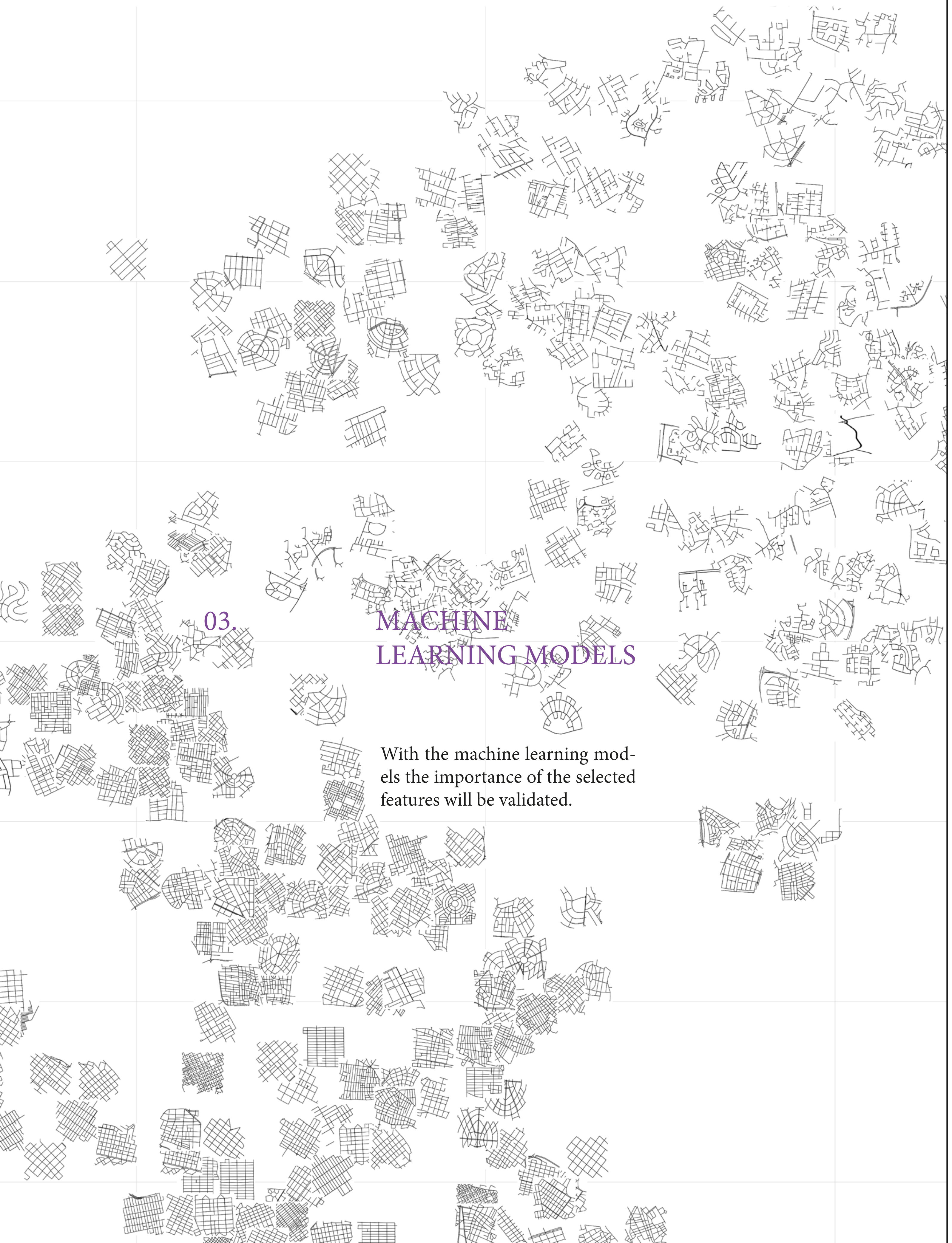
The second step consisted in clustering the graph embeddings into the 5 group typologies.

While the tsne dimensionality reduction performs well, k-means clustering does not produce clear groups of patterns that can be distinguished visually.

Two of the clusters contain typologies of fragmented/ warped paralell and gridiron, one mostly loops and lollipops and lollipops on stick and the other ones loops and lollipops, lollipops on stick and parallel fragmented.



tsne visualization
dual



03. MACHINE LEARNING MODELS

With the machine learning models the importance of the selected features will be validated.

03.2 SUPERVISED

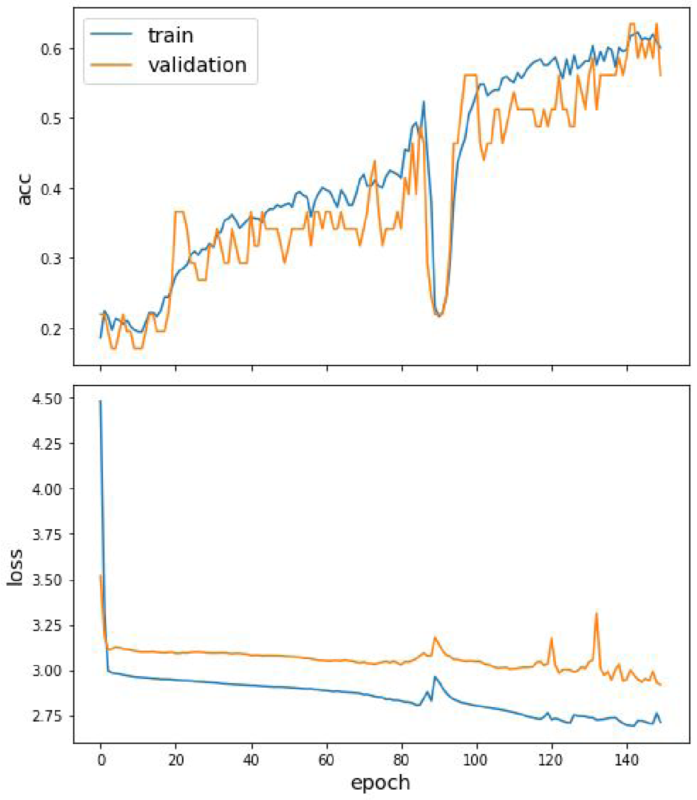
an attempt to build a street network typology classifier.

03.2 SUPERVISED

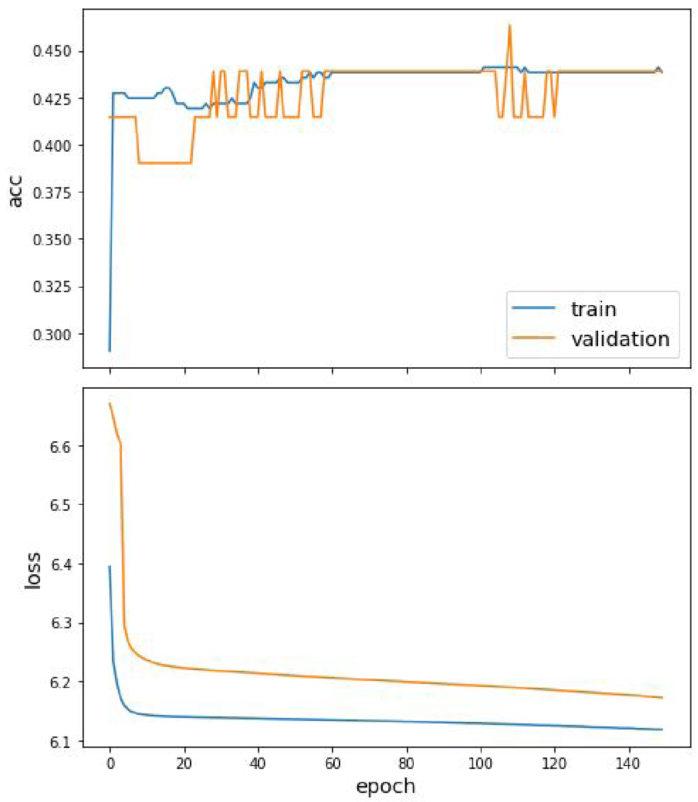
For this classifying task two models were tested: GCN (graph convolutional layers) and DGCNN[1] (deep graph convolutional neural network). GCN failed to train satisfactory models while the results from DGCNN were acceptable.

We can observe that the chosen features increase accuracy by one third. Also by removal of the features that encode length the accuracy is not hindered. In this case the features that encode angle intersection information seem to be more important than the ones encoding length, raising the hypothesis that angles are more important for classifying typologies than length.

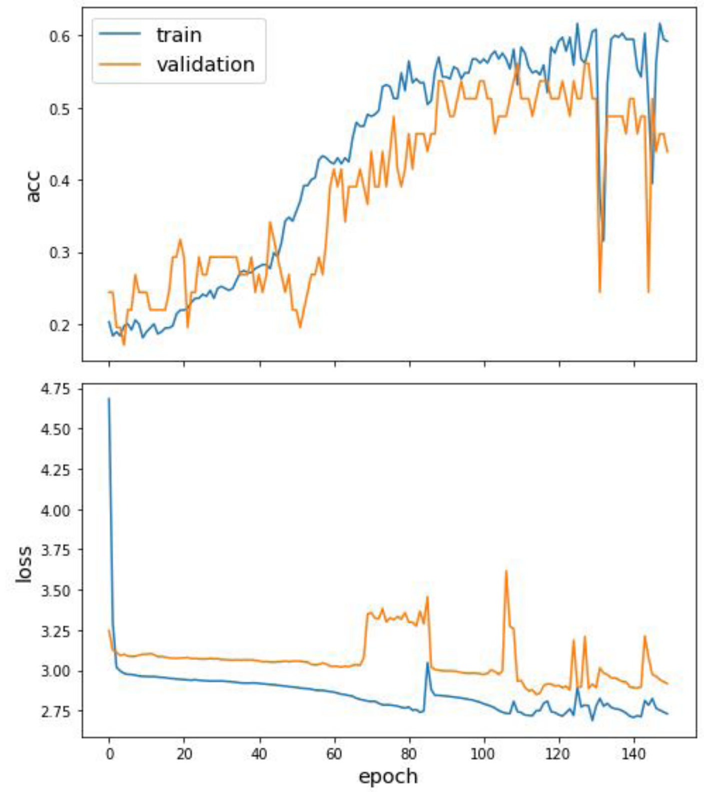
1 (Zhang, et al. 2018)



accuracy and loss for primal encoding



accuracy and loss for primal encoding
no features

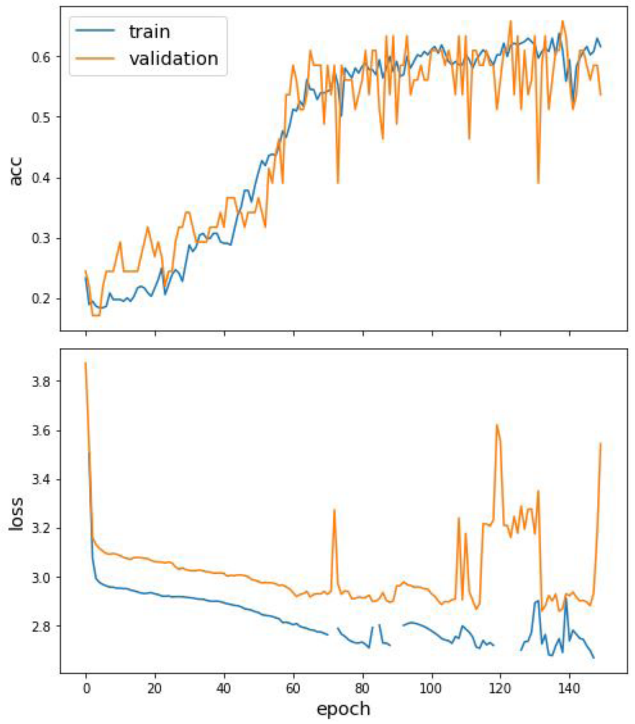


accuracy and loss for primal encoding
no features representing length

03.2 SUPERVISED

The same DGCNN model trained on the dual representation yields similar accuracy results to the primal one.

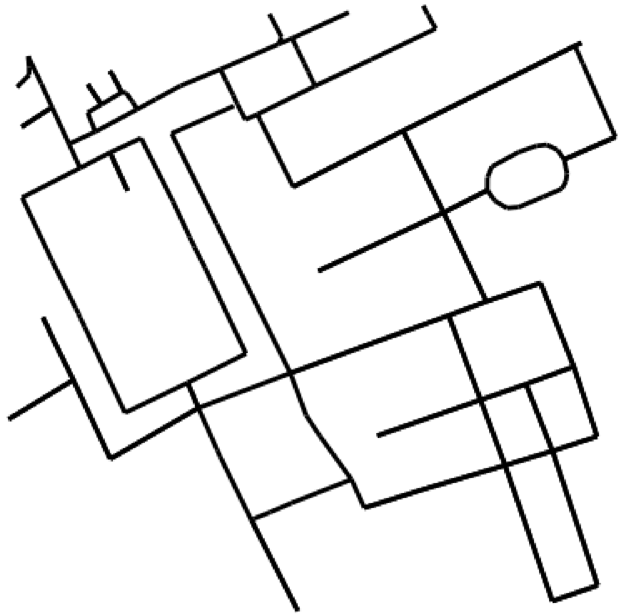
In both cases the need for a bigger dataset is felt, both for preventing overfitting in the case of training over 150 epochs and also for the amount of validation and test sets.



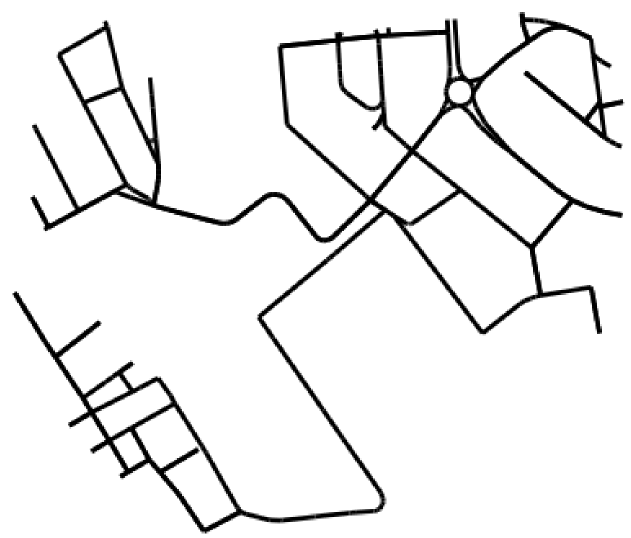
accuracy and loss for dual encoding

Testing the trained model on unseen and unlabeled graphs also provides interesting insights on street networks where clear typologies can not be clearly defined. We can observe that gridion typologies usually get succesfully identified. Warped parallel on the other side is the least succesful one, also owing to the fact that it had the least samples in the dataset.

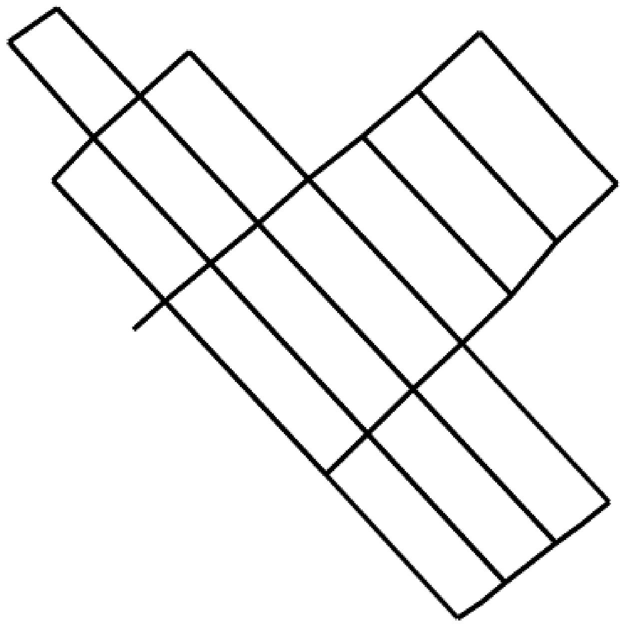
0	Gridiron
1	Fragmented Parallel
2	Warped Parallel
3	Loops and Lollipops
4	Lollipops on a stick



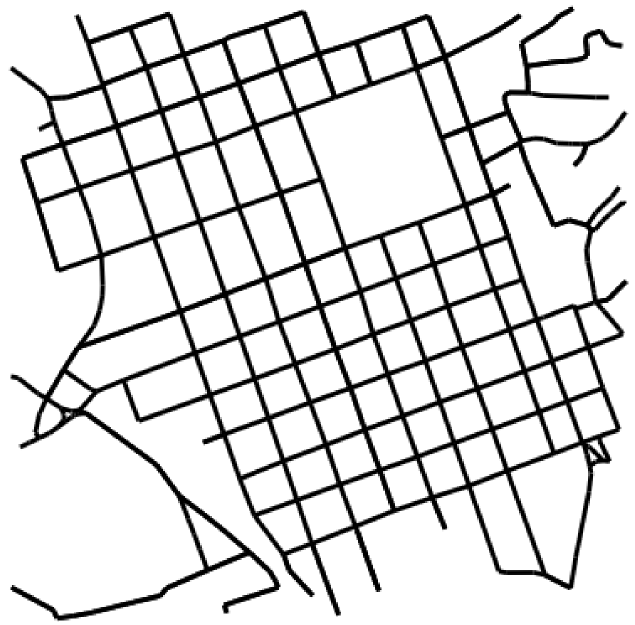
0	1	2	3	4
0.39	0.27	0	0.13	0.21



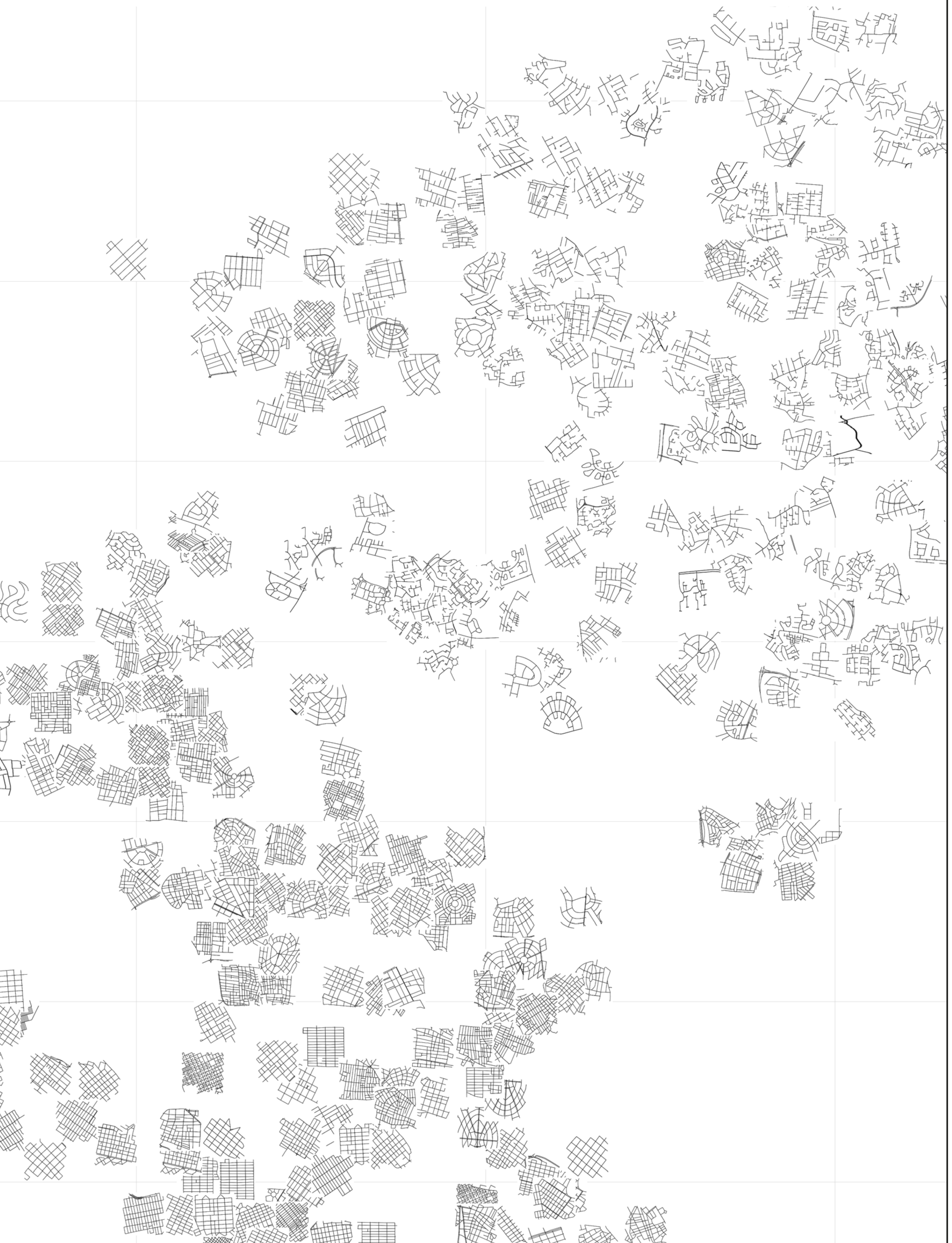
0	1	2	3	4
0.16	0.41	0	0.13	0.30



0	1	2	3	4
0.54	0.22	0	0.24	0



0	1	2	3	4
0.32	0.25	0	0.17	0.26



04. GENERATING

some remarks on how to encode graphs for applying generative models.

04. GENERATING

Three of the most popullar approaches of deep generative models would be: variational autoencoders (VAEs), generative adversarial networks (GANs), and autoregressive models. These generative techniques can be also combined with one-another. [1]

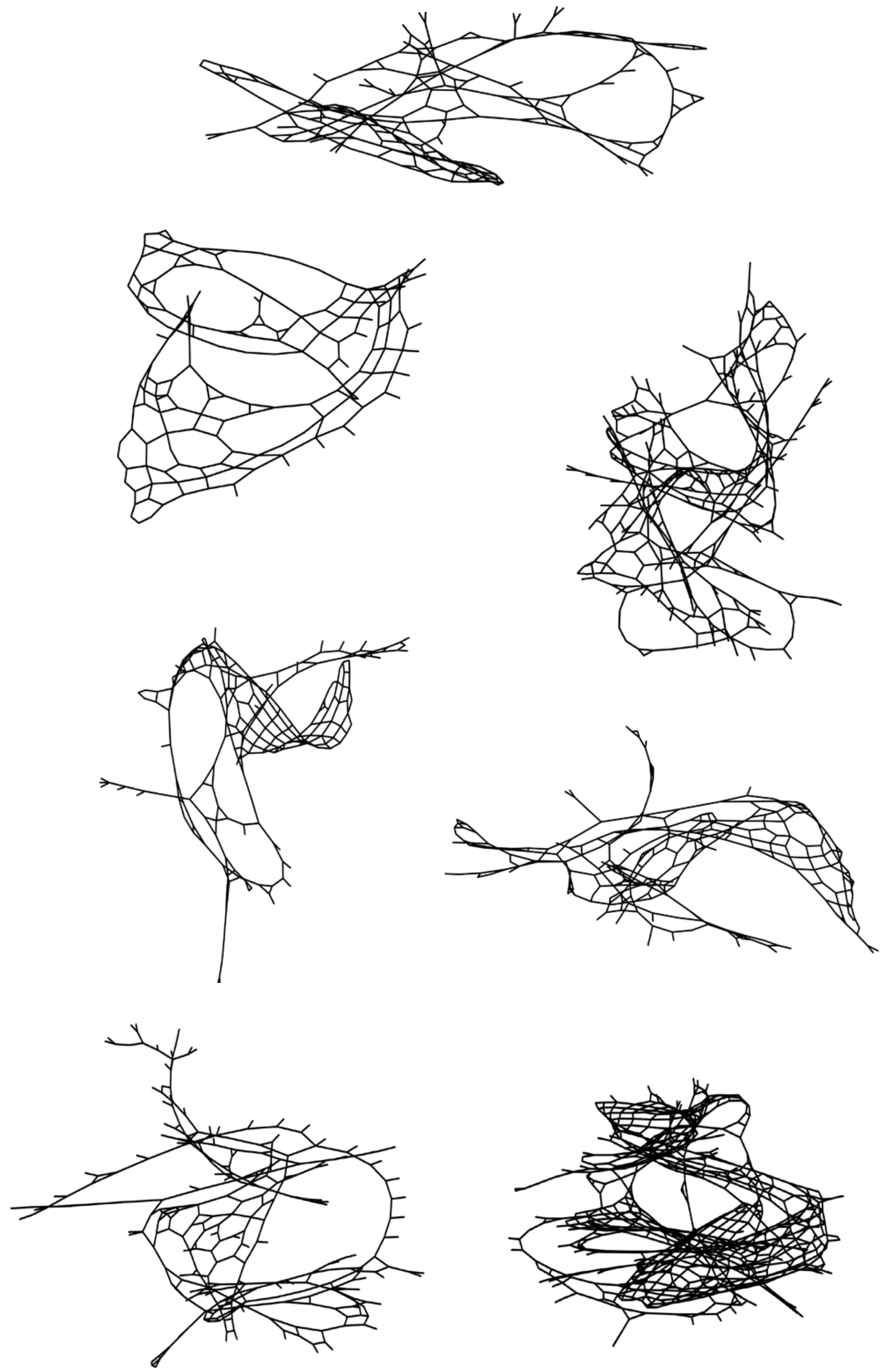
VAEs have serious limitations and can be applied to small graphs with hundred nodes or less.

GAN-based approaches to graph generation have so far received less attention and success than their variational counterparts. This is likely due to the difficulties involved in the minimax optimization that GANbased approaches require, and investigating the limits of GAN-based graph generation is currently an open problem.

Auto regressive models are more sophisticated models and can also be incorporated with VAEs and GAN for better performance. Some of the most popular ones include GraphRNN[2] and GRAN[3].

Both of these models can be applied for generating streetnetworks. In order to do so, however both these models should be tweaked for including node features, since they currently only are able to generate a graph topology.

1 (Hamilton 2020)
2 (You, et al. 2018)
3 (Liao, et al. 2020)



Some of the samples generated with the GRAN model, trained on a set of 120 graphs. These are simple spring layout representations of the graphs.

04. GENERATING

(Chu, et al. 2019) were succesful in generating street networks that resemble certain cities. [1]

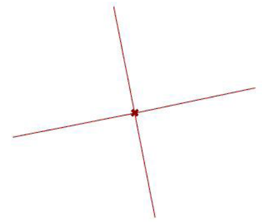
They have encoded the street networks as primal graphs where each node has as features its latitude and longitude.

Our proposed dual representation provides an easy way of encoding and decoding back street networks, and can be used as an input for graph generative models.

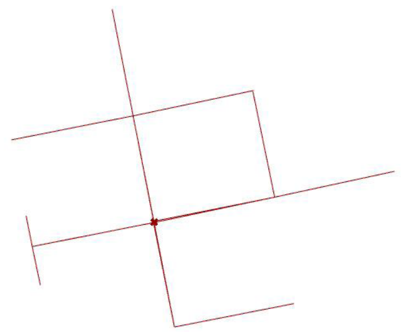
Now, its up to tweaking available generative models and see how it would compare to the encoding strategy by (Chu, et al. 2019)!

1 Neural Turtle Graphics for Modeling City Road Layouts

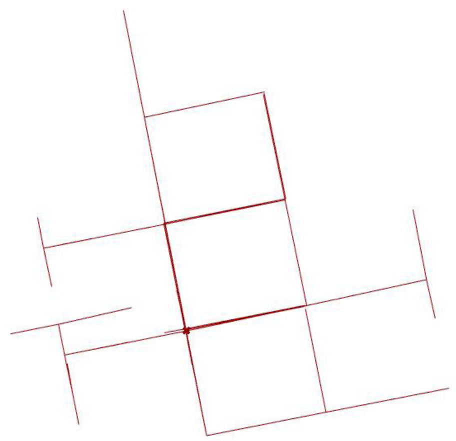
01



02

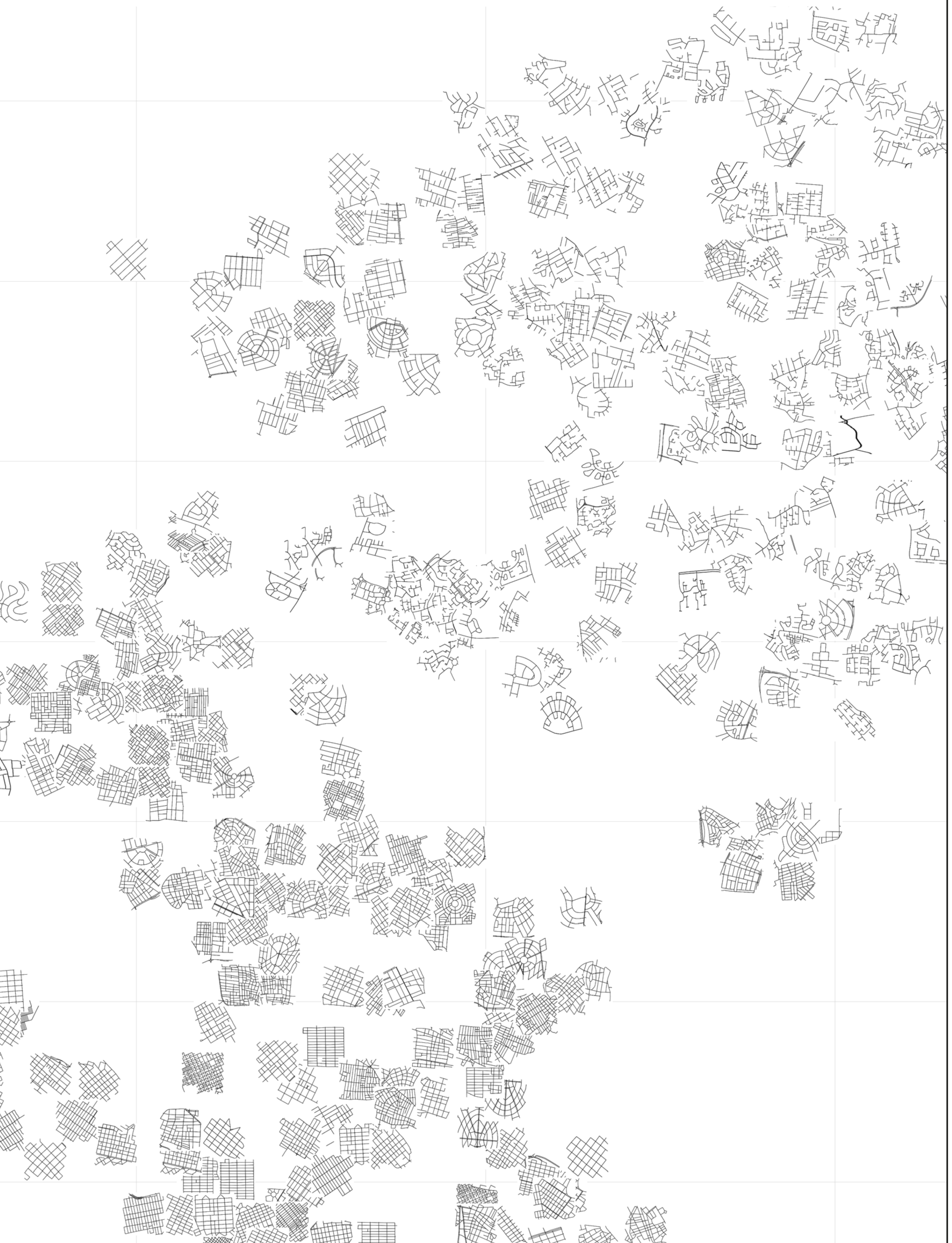


03



Through some scripting street networks can be generated back from the edge lists and node feature files.

In our case we have extracted the angle as a relation to the geodesic north, due to the ease of extracting this information from osmnx, but every other direction can be also chosen.



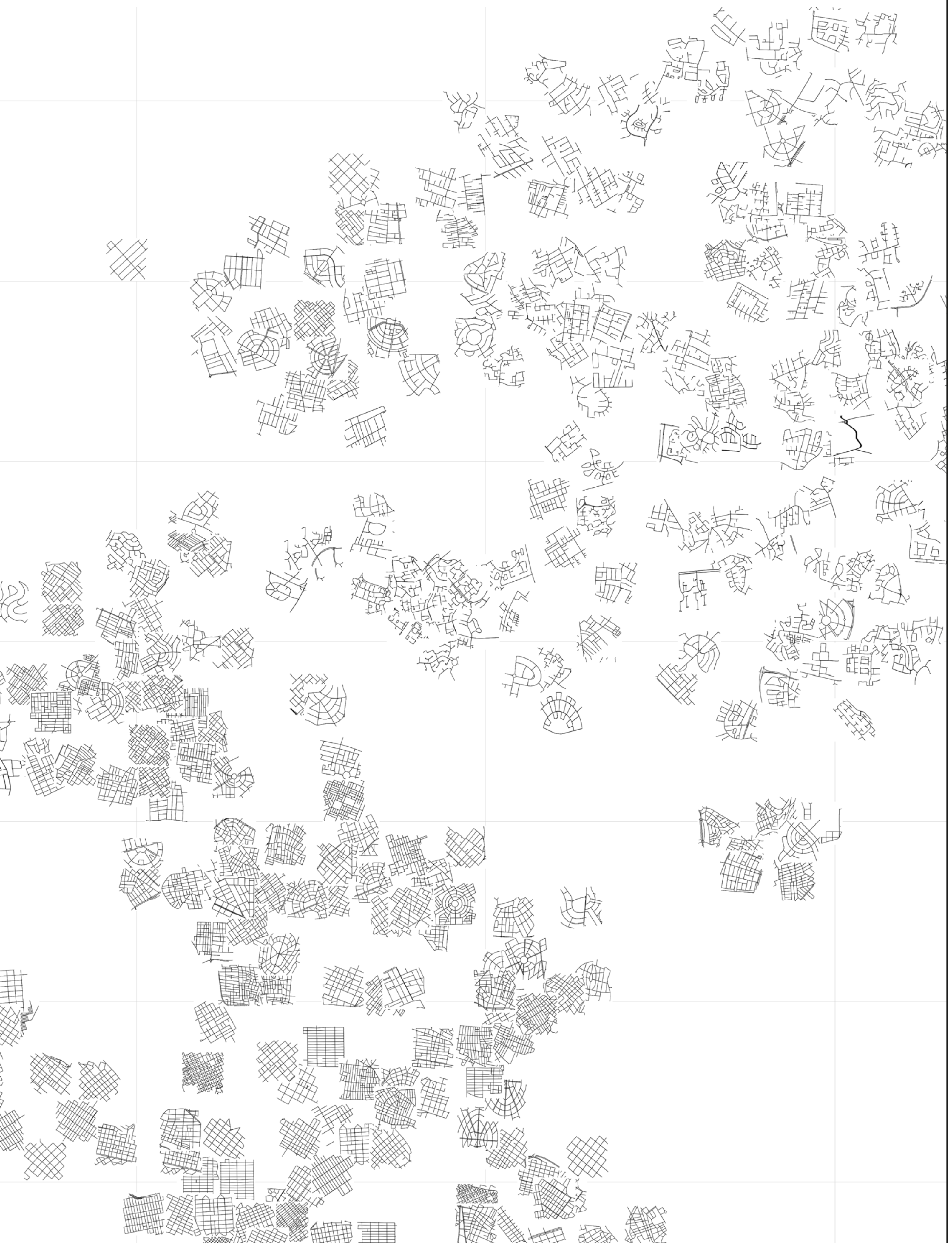
05. CONCLUSIONS

This study has been a exploration of graph machine learning methods applicable to the geometric analysis of street networks.

It proves that these models can be accesible even to non AI professionals due to the existence of open source libraries and demos.

It is up to architects and urban planners to think about how these methods can be used in their practise. In my opinion the analysis of street networks should continue in the chicken-egg dynamic, thinking about how street networks affect and can be affected by different factors, be them environmental, economic and even pragmatic ones such as building area or function.

In this sense both encoding strategies might be useful in further studies, be that analysis or generation of street networks.



06. REFERENCES

AlHalawani, Sawsan, Yong-Liang Yang, Peter Wonka, and Niloy Mitra. "What Makes London Work Like London." Eurographics Symposium on Geometry Processing , 2014.

Bai, Yunsheng, et al. "Unsupervised Inductive Graph-Level Representation Learning via Graph-Graph Proximity." 2019.

Chu, Hang, et al. "Neural Turtle Graphics for Modeling City Road Layouts." 2019.
DeMichele, Patrick, Pablo Santos, and Isaac Scheinfeld. "Graph Embeddings for Street Network Analysis." December 2019.

Fang, Zhou, Tianren Yang, and Ying Jin. "DeepStreet: A deep learning powered urban street network generation module." 27 02 2020.

Hamilton, William. "Graph Representation Learning." Synthesis Lectures on Artificial Intelligence and Machine Learning,, 2020: 1-159.

Kempinska, Kira, and Roberto Murcio. "Modelling urban networks using." Applied Network Science , 29 11 2019.

Liao, Renjie, et al. "Efficient Graph Generation with." 2020.

Marshall, Stephen, Jorge Gil, Karl Kropf, Martin Tomko, and Lucas Figueiredo. "Street Network Studies: from Networks to Models and their Representations." Networks and Spatial Economics, 2018: 1-15.

Porta, Sergio, Paolo Crucitti, and Vito Latora. "The Network Analysis of Urban Streets: A Dual Approach." 2004.

Southworth, Michael, and Eran Ben-Joseph. Streets and the Shaping of Towns and Cities. 1997.

Xue, Jiawei, et al. "Quantifying the spatial homogeneity of urban road networks via graph neural networks." Nature Machine Intelligence, 2022: 246-257.

You, Jiaxuan, Rex Ying, Xiang Ren, William Hamilton, and Jure Leskovec. "GraphRNN: Generating Realistic Graphs with Deep Auto-regressive Models." 2018.

Zhang, Muhan, Zhicheng Cui, Marion Neumann, and Yixin Chen. "An End-to-End Deep Learning." The Thirty-Second AAAI Conference, 2018.