PolisNet: Designing Future Cities with Deep Neural Networks

Gianni Barlacchi∗
gianni.balarach@gmail.com
Fondazione Bruno Kessler
Trento, Italy

Marco De Nadai∗
denadai@fbk.eu
Fondazione Bruno Kessler
Trento, Italy

Yahui Liu∗
yliu@fbk.eu
Fondazione Bruno Kessler
Trento, Italy

Bruno Lepri∗
lepri@fbk.eu
Fondazione Bruno Kessler
Trento, Italy

Luca Pappalardo∗
luca.pappalardo@isti.cnr.it
ISTI-CNR
Pisa, Italy

ABSTRACT

Cities are undoubting the protagonists of world migration of the last decades. People have increasingly moved to urban centers attracted by a greater pool of opportunities, jobs and education at eased reach thanks to shorter distances and public transportation. Recently, the advent of new data sources and methods has acted as a real game-changer in urban studies. The digitalization of society has allowed to observe, understand end even predict many aspects of human mobility and social behavior. However, despite the increasing ability to describe human behavior in cities, how to exploit big data to improve the design of urban areas is still unclear. Starting from the data provided by the Future Cities Challenge, we propose PolisNet, an AI-aided approach to design and modify urban areas to make them successful. This framework is based on Deep Neural Networks (DNNs) that exploits urban imagery and Point Of Interests (POIs) to propose what can be improved in an urban zone, how, and where.

1 INTRODUCTION

Urban planners, sociologists and policymakers have always been concerned about improving distressed districts and create the conditions to spur urban life. Neighbourhoods, in particular, are the most fundamental unit of city life that shape the crime [16], diversity [8] and even the individual’s economic success in life [18]. However, there is still little guidance on designing them and what improve on existing neighborhoods, how, and where.

Here, we propose PolisNet, an AI-aided approach to design and modify neighborhoods to make them successful. This framework is based on a Deep Neural Network (DNN) that exploits urban imagery and Point Of Interests (POIs) data to propose what can be improved in a neighborhood, how, and where. In particular, we empirically describe each neighborhood in a city through an OpenStreetMap ¹ aerial image, which describes the built environment, its POIs and the metrics of urban success, which is based on the well-established vitality of a neighborhood [3, 8, 21]. This metric was previously used to prescribe planning principles and sustainable human activity and urban development [21]. By leveraging aerial image, POIs and success data, PolisNet perturbs the aerial image of each urban zone to propose where it might be modified and what POI to add or eliminate to make the urban zone more successful. For example, given an aerial image of south-central Los Angeles, PolisNet suggests the type and the geographical position of POIs that, added or deleted, increase the success of south-central Los Angeles.

To the best of our knowledge, this is the first research effort towards the creation of AI-aided design systems that could be a valid support for policymakers and city architects, and help reduce the considerable costs related to urban design and simulation. We anticipate that our network will be a groundbreaking system to improve urban life and design neighborhoods with lower crime rate, cultural and economic distress.

The rest of the paper is organized as follows. Section 2 provides an overview on urban related works, whereas Section 3 presents an in-depth description of system. In Section 4, we characterize the employed datasets and explain our experimental setup, together with the obtained results. Finally, conclusions and future research are drawn in Section 5.

2 RELATED WORK

Our article primary speaks to two streams of literature. The first one is the community of urban computing, which studies human behavior in cities through massive data about human displacements, socio-economic and geographical data. The second one is the stream of literature that tries to understand, recommend and design cities that work by collaborating with local authorities, communities and policymakers.

In Urban planning, many approaches tried to study neighborhoods and their revitalization. Notably, Jane Jacobs with her The Death and Life of Great American Cities [8] influenced many urban plans in the world. In her book she praised diversity as an ingredient to have convenient social interactions, face-to-face encounters, and a spontaneous sense of community that spurs economy, security and urban life. More recently, some guidelines and policies have been shared to revitalize neighborhoods [10, 19] based on high-tech entrepreneurship, walkability, and a sense of community similar to the old-fashioned Jacobs ideas.

In Urban computing, many approaches tried to model and predict the activities and characteristics of urban areas. Notably, Noulas et al. [13] proposed a model for predicting an urban zone’s prominent activity using Foursquare and mobile phone data. Zhang et al. [22] proposed an online, semi-supervised, and multimodal embedding method for geo-located information with space, time and text.

Several works have also targeted land use classification and functional area detection. For example, Yuan et al. [20] propose a framework to classify functionalities of an area for the city of Beijing.

¹Authors equally contributed to the manuscript.
²https://www.openstreetmap.org
using POIs and trajectories of taxis. Barlacchi et al. [2] propose a novel machine learning representation based on the encoding of Foursquare POIs to classify the most prominent land use of an urban area.

Another strand of research focuses on defining measures of success for a city or urban zone. Yue et al. [21] define urban vitality as the capacity of an urban environment to boost social activities. They discover that urban vitality showed a positive correlation with phone usage density, and that the urban structure of Shanghai plays a crucial role in its urban vitality. De Nadai et al. [3] use the average number of mobile Internet connections throughout a typical business day, divided by the an urban zone’s area, as a proxy for urban vitality.

3 POLISNET SYSTEM

Our framework, PolisNet, is an AI-based framework that uses digital data and Deep Neural Networks (DNNs) to design a neighborhood and propose what improve, how, and where. The architecture of PolisNet is illustrated in Figure 3.

3.1 Urban Success Metric (USM)

Defining urban success is not an easy task and many scholars have tried to define it through extensive studies [14]. Although urban success can manifest itself in different ways, we straightforwardly define what is not: an area without visitors. Thus, we build upon the sociological theory of Jane Jacobs, which defines vitality and diversity as an essential factor for urban success. Recent work in urban computing empirically defined vitality by measuring the activity of people in a neighborhood from mobile phone data [3, 4]. Similarly, we here define as vital a place where people geo-localize through Foursquare data. Thus, vitality in a neighborhood \( i \) is formalized as:

\[
\text{vitality} = \sum_{p \in \mathcal{P}_i} C(p)
\]

where \( \mathcal{P}_i \) is the set of POIs in the neighborhood and \( C(p) \) is the number of check-ins for a particular POI. Similarly, we define the economic vitality counting only the vitality of the Shops and Food POIs. Note that our framework might be used with different metrics such as the number of crimes and the number of cultural events in a neighborhood.

3.2 Urban Success Evaluator (USE)

As shown in Figure 1, we design an Urban Success Evaluator (USE) to distinguish the zone to be successful or not, which can be formulated as a binary classification problem, where “0” and “1” refer to unsuccessful and successful, respectively. Recently, deep convolutional neural networks (CNNs) have led to a series of breakthroughs for image classification [6, 11]. Thus, we build our model based on two state-of-the-art methods, including Inception-v3 [17] and ResNet-v2 [7], which are widely used as basic networks in various computer vision tasks.

Specially, we define the training set as \( \mathcal{S} = \{ (I_n, M_n, C_n), n = 1, 2, \ldots, N \} \), where the \( I_n \) denotes the original map image, \( M_n \) denotes the mask of POIs, and \( C_n \in \{0, 1\} \) denotes the binary classes. To train our model, we calculate the cross-entropy loss:

\[
\mathcal{L}_{cls} = \frac{1}{N} \sum_{n=1}^{N} C_n \log(p(C_n)) + (1 - C_n) \log(1 - p(C_n)),
\]

where \( p(C_n) \) refers to the probability of zone \( (I_n, M_n) \) to be correctly classified. For simplicity, we set equal loss weights for both classes.

3.3 Urban Success Designer (USD)

We then design an Urban Success Designer (USD) to model the distribution of POIs in the spatial space in different cities, which shows the potential ability that can be applied to propose ideal new POIs to transform an unsuccessful zone in a successful one. We re-formulate such kind of problem into a location regression problem. Inspired by fast object detection methods [12, 15], we assume that each POI can be enlarged with a buffer area (formed as bounding box in object detection). After that, our approach discretizes the output space of POI locations into a set of default boxes over different aspect ratios per feature map location. At prediction time, the network generates scores for the presence of each POI category in each default box and produces adjustments to the box to better match the shape of the POI buffer.

Figure 2 shows the pipeline of our USD, in which the training objective is derived from the MultiBox objective [12] but is specified to handle the regression of multiple POI categories. In contrast with previous object detection methods, we try to predict location and category of the removed POI. Thus, we define the training set as \( \mathcal{D} = \{ (I_n, M_n, p^\dagger, C^\dagger), n = 1, 2, \ldots, N \} \), where the \( I_n \) denotes the original map image, \( M_n \) denotes the mask of POIs with a randomly removed POI (located in \( p^\dagger \) with a category \( C^\dagger \in \{1, 2, \ldots, K\} \), \( K \) refers to number of POI categories). The overall objective loss function is a weighted sum of the localization loss (loc) and the confidence loss (conf):

\[
\mathcal{L}_{reg} = \frac{1}{N} (\mathcal{L}_{conf}(x, C^\dagger) + \alpha \mathcal{L}_{loc}(x, l, g))
\]

where \( x^\dagger \in \{0, 1\} \) be an indicator for matching the \( t \)-th default box to the \( j \)-th ground truth box of category \( C \). The localization loss is a Smooth L1 loss [5] between the predicted box (\( l \)) and the ground truth box (\( g \)) parameters. The confidence loss is the softmax loss over multiple classes confidences (\( C \)) [12].

3.4 PolisNet Framework

First, we define a neighborhood as a uniform and non-overlapping regular cell of approximately 600 \times 600 meters based on the Slippy Map standard\footnote{https://wiki.openstreetmap.org/wiki/Slippy_Map} at zoom 15. Then, we associate each cell with the corresponding OpenStreetMap aerial image, the metrics of urban success, and the map mask built over the characteristics of the POIs contained in the cell. For each unsuccessful urban zone, i.e., an urban zone with the success metric lower than 0.5, PolisNet proposes several changes in order to make the zone successful.

The design process of PolisNet is divided into two phases: in the first phase, the USE evaluates whether an urban zone needs to be redesigned or not. In the case the evaluation reveals that the zone is successful, the system outputs the new generated successful area indicating the required changes: (i) where it requires changes,
and (ii) what POIs should be added in such locations. Differently, if the zone needs refinements, the urban zone representation is forwarded to the second step. In this step the USD, by relying on the location regression model, proposes a new list of possible POIs, with their corresponding geographical locations, that can be added in order to turn the zone into a successful one. The two steps are repeated in a loop for a maximum number of iterations or until when the urban zone does not turn into a successful zone.

Figure 1: Schema of the USE classifier.

Figure 2: Schema of the USD model: given an input area, it decides where and what to add to make the input area successful.

Figure 3: Schema of PolisNet: it proposes where an urban zone should be modified, and what POI to add or eliminate to make the zone successful.

4 EXPERIMENTS

Our experiments aim at demonstrating the effectiveness of our models on two phases: (i) the urban success evaluation and (ii) the urban success design. We use a database of cities composed by 8 kinds of POIs that belong to 9 different cities around the world.

4.1 Experimental Setup

As aforementioned, we represent each neighborhood through non-overlapping square cells based on Slippy Maps of zoom 15. For each cell we download the aerial map of OpenStreetmap removing any visual information connected to the POIs. This was done through a custom map style of Mapbox \(^3\) and it ensures higher independence between the POI mask and the Map. Then, we use the Future Cities Challenge [1] dataset to obtain the Foursquare venues and the movements information. From the venues we extract the category, latitude and longitude, while from the movements we aggregated the number of check-ins in each venue. We aggregate the categories of the venues to the highest hierarchy of the Foursquare category tree \(^4\) and assign a color to each category. We discard the event and residents categories. For each cell we build a POIs Mask where each venue is represented with a color based on its category, and located in a \((x, y)\) pixel depending on its geographical location in the map. Finally, we represent each neighborhood through a triplet (aerial image, POIs Mask, label), where the label is binarized into successful and unsuccessful. The label is successful when the vitality (Equation (1)) and the economic vitality are above the 20th percentile of the distribution of vitality for that city, unsuccessful otherwise. We build our data for nine cities namely Chicago, Istanbul, London, Los Angeles, New York, Paris, Seoul, Singapore, Tokyo. This resulted on 20,060 images of neighborhoods.

We train our model using Adam [9] with \(\beta_1 = 0.5\) and \(\beta_2 = 0.999\) and batch size 4 for evaluator model and batch size 1 for designer model. For data augmentation we flip the images horizontally with a probability of 0.5. The initial learning rate is set to 0.0002. We train for 30 epochs and linearly decay the rate to zero over the last 20 epochs. It takes about several hours for each model with a single GeForce GTX 1080 Ti GPU. All these models are implemented using the deep learning framework PyTorch. In particular, the weight term \(\alpha\) is set to 1.0 in the designer model. To measure the impact of our models as well as the baselines, we used well known metrics for assessing the accuracy of DNNs.

We use the accuracy score to assess the performances on the binary classifier in the USE. To evaluate the USD we rely on two metrics: (i) the top-k distance error, which is the minimal distance among the top-k predicted POIs and the target POI in the image coordinate, and (ii) the closest accuracy, which measures whether or not the label of closest predicted POI is equal to the target label.

4.2 Preliminary Results

In this section we present some preliminary results for the two steps that compose our PolisNet system. Table 1 shows the accuracy of the USE classifier. The Table shows that conditioning the model with the city name does not provide significant improvements. The Inception-v3 and Resnet networks with the city information have a 0.4% and -4.8% of absolute points improvement, respectively.

In Table 2 we report preliminary results for the very challenging task of proposing changes to apply to unsuccessful urban zones. The accuracy and distance of the top-k predicted results show that it is very difficult to guess the correct POI category and image

\(^3\)https://www.mapbox.com/gallery/

\(^4\)https://developer.foursquare.com/docs/resources/categories
coordinates of the removed POI. However, when more POIs candidates are considered, the model shows promising results both in terms of accuracy and distance error. This suggests that potential improvements can be applied to further develop this idea. The low performances with \( k < 50 \) might be due to the very basic input representation we are providing to our network. Indeed, some improvements can consider (i) to design a more expressive dense representation of POIs instead of using the simple one-hot encoding representation and the POI macro-category; (ii) to better balance the dataset both in terms of city map and categories of POIs; (iii) to provide a more informative visual input than a simple map mask with colored pixels.

\[
\begin{array}{ccc}
\text{k} & \text{Accuracy} & \text{Distance} \\
10 & 0.074 \pm 0.078 & 155.37 \pm 37.51 \\
20 & 0.112 \pm 0.074 & 140.51 \pm 35.48 \\
50 & 0.149 \pm 0.014 & 125.77 \pm 37.50 \\
\end{array}
\]

Table 2: Accuracy and distance error of the USD in proposing changes for a given unsuccessful urban zone.

## 5 CONCLUSION AND FUTURE IMPACT

In this paper we presented PolisNet, an AI-based system that recommends how to improve the success of urban zones in a city. Since the underlying idea is rather flexible, PolisNet can be improved and extended in several directions.

First, while we use in this paper a fixed threshold for defining an urban zone as successful or unsuccessful (i.e., the median of the urban vitality), we may let the user choose the threshold to use for the discrimination. In this way, a policy maker may decide, according to the available economic resources, how much to change a zone to achieve the desired level of urban success.

Second, with slight modifications, our system can be adapted to generate an optimized street network given the current mobility fluxes observed in an urban zone. In this case, PolisNet would perturb the underlying street network or the position of existing POIs, given certain constraints, in order to reduce indicators of mobility success, such as travel times or distances, the emergence of traffic jams, or the number of car accidents.

## REFERENCES


