Yang, Jun; Fricker, Pia; Jung, Alexander

From Intuition to Reasoning

Published in:
Journal of Digital Landscape Architecture

DOI:
10.14627/537724008

Published: 01/01/2022

Document Version
Publisher's PDF, also known as Version of record

From Intuition to Reasoning: Analyzing Correlative Attributes of Walkability in Urban Environments with Machine Learning

Jun Yang[1], Pia Fricker[2], Alexander Jung[2]

[1]Aalto University, Espoo/Finland · jun.1.yang@aalto.fi
[2]Aalto University, Espoo/Finland

Abstract: A fundamental challenge in the planning and operation of modern cities is their walkability. Walkability is typically assessed using geo-information system (GIS) or real-time observations. These existing methods, however, are not suited to the task in several key aspects. GIS-based assessment is inherently limited in capturing the details of a space, and observation-based methods are time and resource consuming. To overcome these limitations, we introduce a novel machine learning (ML) based approach. Our main concept is to make walkability an ML problem, where sites or locations are defined as data points. The data points are characterised by features extracted from street images. The ultimate quantity of interesting aspects (or labels) of a data point determines its level of walkability. Our assessment of walkability is based on the perceived accessibility of sites as measured via survey. Roughly speaking, our ML approach learns correlations between the presence of specific objects such as trees, buildings, sidewalks, and the perceived walkability of a specific location. The main methodological contribution of our research is a novel feature extraction method based on semantic segmentation techniques. The extracted features are fed into different off-the-shelf supervised ML methods and compared. The results demonstrate the usefulness of our approach to predict the walkability of an urban location based on an ML analysis of street image content.

Keywords: Walkability, machine learning, sustainable urban and landscape design, urban digitization

1 Introduction

1.1 Objectives under Global Urban Challenge and Urban Digitization

Most cities in the world are currently confronting increasingly complex problems (e. g., urban heat island effect (UHI), urban flooding, health challenges) caused by global warming and urbanization. Urban digitization is a tool that has the potential of grasping the complexity by converting urban information into a language that computers understand. It can provide enormous data to help deepen our understanding of cities and their problems. Machine learning (ML) is an application where computers predict future events based on learning algorithms (JUNG 2022). It is a new approach for analyzing data effectively and accurately, which can be used to support design strategies to help meet the continually growing challenges cities face. The study supports the Sustainable Development Goals 11: *"Sustainable cities and communities"* and 9: *"Good health and well-being"*, because walkability can be viewed as a key measure in sustainable urban strategies. Furthermore, as discussed by LITMAN (2003) and SPECK (2012), walkable neighbourhoods deliver substantial benefits in terms of health, air quality, sustainability and the economy. The research is developing ML methodology to understand firstly how attributes (e. g., trees, roads, buildings) in street image datasets affect walkability and secondly how to transform subjective visual perception into quantitative data. This shift in understanding quantifiable attributes will support urban landscape planners' research in the qualities of the street level based less on intuition and more on evidence-based

Journal of Digital Landscape Architecture, 7-2022, pp. 71-81. © Wichmann Verlag, VDE VERLAG GMBH · Berlin · Offenbach. ISBN 978-3-87907-724-3, ISSN 2367-4253, e-ISSN 2511-624X, doi:10.14627/537724008.

practice. The long-term objective of this research is to use the advantages of ML, namely continuous learning new data with improved effectiveness and accuracy, to maximize the understanding of correlations between data so that this method can support urban and landscape planning with more efficiency and acuity.

1.2 Filling a Gap Based on the Past Studies

Since the 1960s, city planners have identified different attributes like sidewalks (JACOBS 1961, LYNCH 1960), facades (GEHL & KOCH 2011), traffic noise (APPLEYARD & LINTELL 1972) and road networks (JACOBS 1995) as influencing the quality of streets and public spaces. These studies relied on intuitive observation, contributing empirical knowledge for developing walkable environments. However, it is still a challenge to accurately quantify each attribute's degree of influence on walkability. Furthermore, when cities grow larger and more complex, traditional non-digital methods are more cumbersome because of the difficulty quantifying wider ranges of perceived attributes.

Parallelly, researchers have begun to focus on a more measurable approach to a macro-level of walkability using GIS-based techniques or space syntax (HILLIER 1976), where numeric attributes (e. g., population density, land use mix) are easily calculable. Yet, these attributes usually do not take into account features at street level. For instance, software like Walk Score count the number of available amenities and Walkshed customize paths using decision tree algorithms to evaluate walkability. Nevertheless, they all have their limits on understanding the characteristics of streets like the quality of sidewalks (D'ORSO & MIGLIORE 2020, AGAMPATIAN 2014).

To help mitigate the pressing and complex urban challenges mentioned earlier, our paper aims to discuss novel ways for re-thinking the evaluation methods of perceived walkability. When cities grow larger and denser, the quality of urban spaces becomes extremely important in their ability to provide healthy and long-term sustainable user-friendly environments. Detailed data at street-level more precisely reflects the diverse aspects determining the quality of the spatial environment.

1.3 The Role of Machine Learning in Walkability Studies

Computer vision is a field of computer science where computers are programmed to distinguish patterns from visual data such as images or videos. This learning process is usually based on deep learning algorithms. The past decade has witnessed significant progress in computer vision using deep learning techniques, which provides people with new ways to understand and analyse visual data (VOULODIMOS et al. 2018). Current deep learning methods can detect objects (e. g., sidewalk, building, trees) in an image with high accuracy (SALOHEIMO et al. 2021). Image classification and clustering methods are commonly applied to urban-related studies (YIN et al. 2017). In the past five years, a small number of studies focused on more complex questions, such as analyzing the relationship between sky enclosure (YIN & WANG 2016) and walkability, detecting pedestrians on street images (YIN et al. 2017) or visual beauty prediction (WANG et al. 2019, JOGLEKAR et al. 2020) in cities. Although these studies focused only on a few aspects of walkability, the results seem promising for future development of the ML approach.

2 Methods

In our study, ML plays various roles in the data analysis. Firstly, ML is defined as an indirect observer. The process of observation is similar to human perception, though it is not a direct perception of the real world. In the process, ML can understand images at the pixel level, converting semantic image information to numbers of pixels that reflect different objects in image data. Essentially, a survey representing opinions of walking quality is first linked to image data, which is then converted into measurable attributes by using a semantic segmentation model. The objects in an image are considered as attributes of walkability at street level and quantified. Therefore, this approach transforms intuitive perception into measurable and quantifiable information. Following that, the ML model distils the relationship between the transformed data and subjective opinions about the walking environment itself. In the analysis of association, ML model generates the weight of each attribute correlated to walkability levels. These weights can then be used as a coefficient of the same attributes derived from new data to predict walkability levels at a new location. The purpose of converting image data to numeric data is not only to anticipate walkability levels but also to help recognize the correlation between each selected attribute and the predicted results. The study targets the significance of intuitive observation at a large scale in order to analyse and understand the working of complex urban layers in a quantifiable way.

2.1 Data

The location of the survey data is used to acquire street images through Street View Static API which is a service to receive image data based on certain parameters. Then the images are processed by the semantic segmentation library to get the ratios of each attribute in individual images. The ratios of each attribute in each location are represented as datapoints marked by labels representing correlating opinions in the survey data (Figure 1).

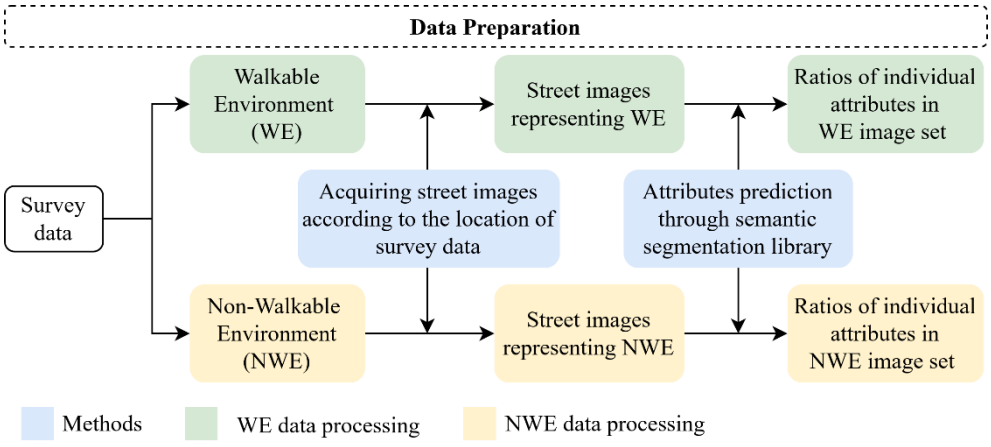

Fig. 1: Data Preparation

2.1.1 Image Acquisition

In the initial phase, the location of the survey data gathered by the City of Helsinki in 2018 is used to acquire street images. The survey collected the residents' opinions about where they considered a pleasant or unpleasant environment for walking. The survey location points are classified into two categories: walkable or non-walkable. Survey data coordinates are generated with GIS software. With coordinates as a single parameter, we cannot always attribute street view images to the actual directions the cameras are facing, namely headings, as the default heading (heading=0°) in Google Street View is true north. To be more specific, streets in a city do not always run in the north-south direction. Without an accurate heading, one may encounter an unexpected image facing the wrong direction. To solve this problem, we calculate the closest road from our coordinates and then we can compute the initial heading using Overpass-turbo, which is a map data mining tool containing road vector data. Then the heading and coordinates are used as parameters to acquire street images (image size, $640 \times 640\ pixels$) through Google Street View Static API. As a single image cannot represent the environment fully, four street images with different headings (heading1 = start degree, heading2 = start degree+90°, heading3 = start+180°, heading 4 = 270°) are acquired from every location (Figure 2).

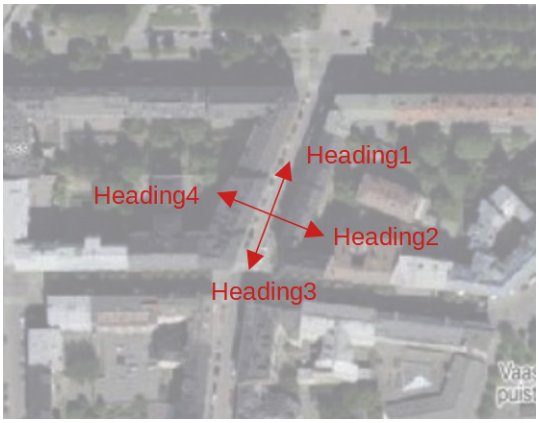

Fig. 2:
Image acquisition based on headings in four directions

2.1.2 Selection of Semantic Segmentation

A python library called Pixellib is selected to extract objects from the pixel level of the image data because of its low-resource consumption, a wide range of predictability, the flexibility of input data and its reliability. Compared to many pre-trained semantic segmentation models usually comprising hundreds of codes and requiring a GPU running environment, the Pixellib library can process one image with a few lines of codes and a common CPU environment. Second, while many pre-trained models are trained on data sets like CITYSCAPES, which include only twenty types of objects, the pre-trained model in Pixellib library is trained on data sets such as ADE20K containing as many as 150 different types of objects (e. g., sky, road, building, water, vegetation). As the urban environment is quite complex, 150 parameters can cover most of the important attributes of a street space. Thirdly, many pre-trained semantic segmentation models have strict input image size, whereas new data should always be resized accordingly for a prediction unless models are retrained. With the Pixellib, the

input image size is not required and comparatively flexible. Lastly, the Pixellib is among the top 10% on the Python Packages index, which makes it a reliable resource.

2.1.3 Calculate Attributes

Using the Pixellib library, the number of pixels is calculated for each attribute in a particular image. As each location comprises four street view images with different headings, we can calculate the total percentage of an individual attribute i at one location as below:

$$PCT^{i}_{n\ total} = \frac{PX^{i}_{n1} + PX^{i}_{n2} + PX^{i}_{n3} + PX^{i}_{n4}}{4}$$

where the $PCT^{i}_{n\ total}$ is the total percentage of the *ith* attribute in four images with different directions at the same location *n.* The PX^{i}_{n} is the pixel ratio of the *i*th attribute in four individual images (*n1 ,n2, n3, n4*) at location *n*. All the attributes occurring in one location are represented as one datapoint (Figure 3).

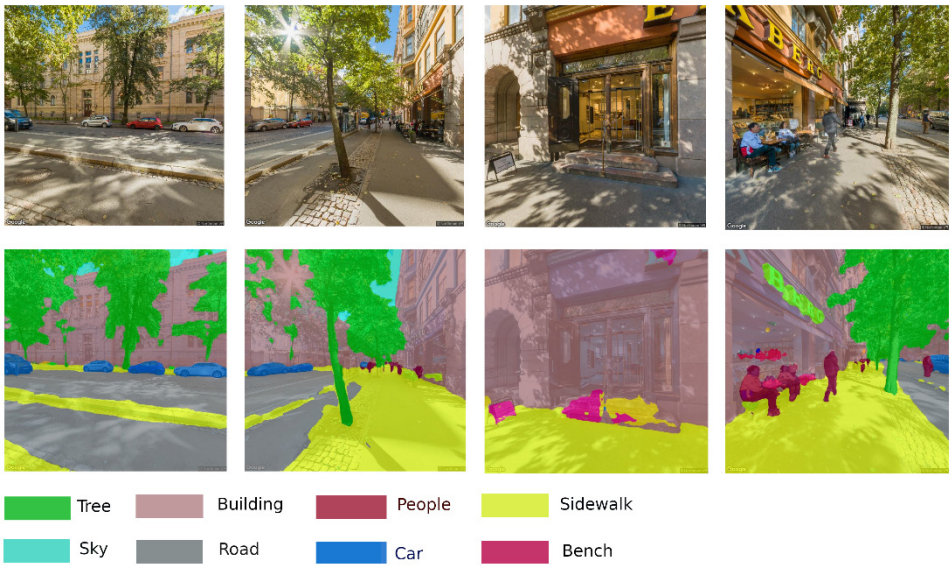

Fig. 3: Example of one datapoint representing a specific urban location. We use four street images of such a location which are then fed into an image segmentation method. The results of the segmentation are then used to construct the characteristic attributes (features) that characterize the datapoint (urban location).

2.2 Predicting Walkability from Attributes

Once the ratios of attributes on all locations are calculated, a process of attributes reduction is made to select the most relevant attributes as features for model training. The labels of training data are taken from the opinions of the survey, namely walkable or non-walkable. The weights of individual features are computed through the training process. Subsequently,

the trained logistic regression model can be used to predict the walkability on the street level when new data is given (Figure 4).

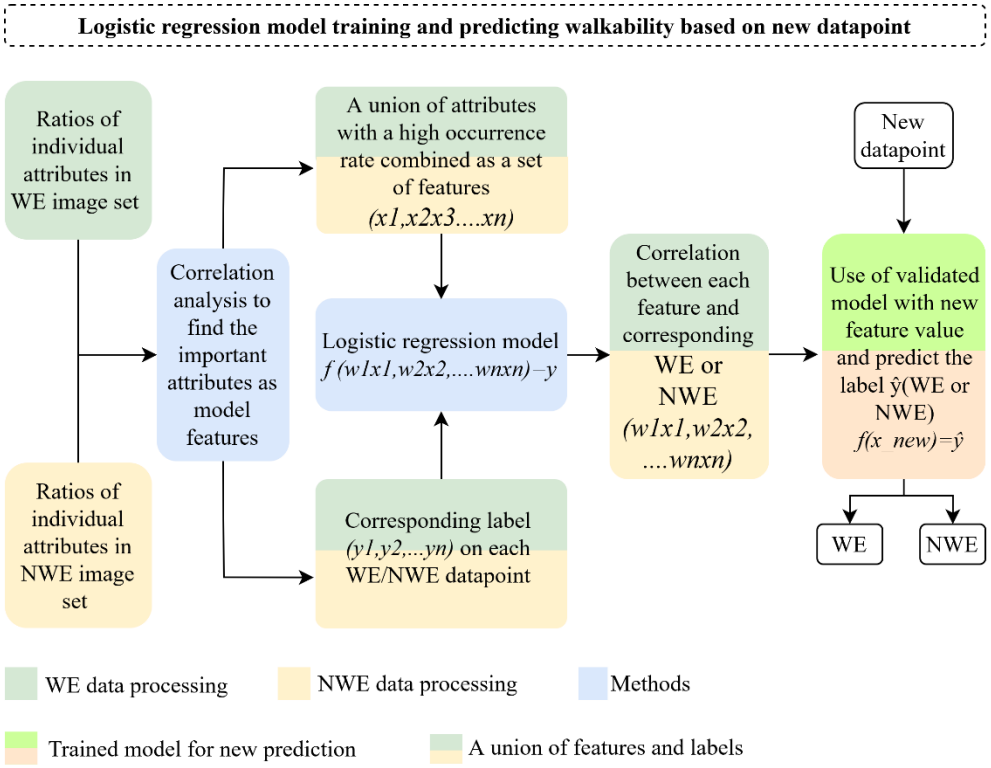

Fig. 4: Model training processing

2.2.1 Selecting Relevant Attributes for Logistic Regression Model

With the earlier equation, each location as a data point returns a set of numerical values representing the proportion of pixels that each attribute occupied on this location, which is demonstrated below:

$$X_n=\{x_1=\boldsymbol{PCT}^{\mathbf{1}}_{\boldsymbol{n\ total}},\ x_2=\boldsymbol{PCT}^{\mathbf{2}}_{\boldsymbol{n\ total}},\ x_3=\boldsymbol{PCT}^{\mathbf{3}}_{\boldsymbol{n\ total}},\ \ldots\ x_i=\boldsymbol{PCT}^{\boldsymbol{i}}_{\boldsymbol{n\ total}}\}$$

$$y_n = 0 \text{ or} y_n = 1$$

where X_n is the dataset at the *nth* location and *x* is a proportion of pixels of one attribute at the *nth* location. The y_n represents opinions of walkable (equal to one) / non-walkable (equal to zero) labels. We processed 1618 datapoints containing 6472 street images using the semantic segmentation model, which predicted a total of 142 classes, or types of attributes, where 125 classes in the unpleasant walking environment and 141 classes in the pleasant walking environment respectively.When the data set included a large number of attributes as input features with limited data points, there was a risk of overfitting, which means the model

performed with minimal errors on the training set but relatively poorly on the test set.To prevent this during the training, we omitted the classes, which were not relevant to our tasks by using the feature selection method. We used the L^1 regularization for the feature selection because it is used extensively as an approach for feature selection and can remove irrelevant features safely (Goodfellow et al. 2016). Subsequently, 50 classes as attributes were selected from the total 141 classes as the input features for the logistic regression model.

The scikit-learn library is used for building the logistic regression model represent as $f(X) = y$, where X is the set of all selected attributes $(x_1,x_2x_3....x_n)$ including n locations and y are the corresponding labels (walkable equals to one or non-walkable equals to zero). The data set X is split to three sets: 70% of data set is used for training and the two third of the rest 30% data is for model validation and one third for model testing respectively. After the training and validation, each feature x contains a learned weight value w that shows the impact degree of feature x on the walkability as below:

$$f(x_1w_1,\ x_2w_2,\ x_3w_3,\ldots x_nw_n) = y_n$$

$$y_n = 0 \text{ or} y_n\ = 1$$

The confusion matrix is a summary describing the competence of the classification model. The diagram below demonstrates the performance of our model on the test data (Figure 5). On non-walkable data points, the model predicted 38 times correctly and 7 times incorrectly. On walkable data points, the model predicted 39 times successfully and failed 14 times. To conclude, the accuracy of prediction of our model is approximate 80.2% on the training set and about 78.57% on the test set respectively.

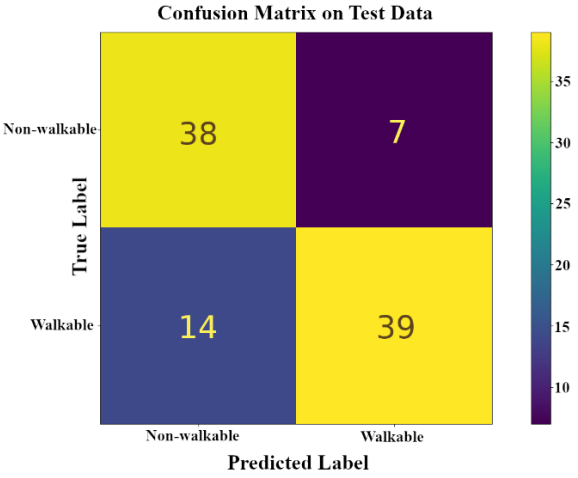

Fig. 5:
Confusion matrix of the classification model on test data

2.2.2 Correlation between selected Attributes and Walkability

After the model training, validating, and testing, the coefficients of each attribute are represented as weights related to walkable/ non-walkable environments (Figure 6). When the weight of a feature is greater than zero, the feature (light red to blue) represents a relatively

high correlation to walkable areas. In contrast, when the weight of a feature is below zero, the feature (dark red) possesses a relatively low correlation to walkable areas but a high correlation to non-walkable areas. The weight values are log odds which are not easy to understand. Thus, we can convert the weight values to normal odds with the exponential function:

$$e^{wi} = odds$$

where e is Euler's number and *wi* is the weight of *i*th feature and the *odds* means the likelihood of the *i*th feature in the targeted label representing a walkable environment.From the graph, the sea feature with its weight value (ca. 0.21) has the highest correlation value to a walkable environment. With the formula above, we can calculate that the sea feature is about 1.23 ($e^{0.21} = 1.23$) times more likely to occur in a pleasant walking environment. Other natural elements including water, field and vegetation (e. g., grass, plant, tree) also have a high link to a walkable environment. In contrast, features like signboard, bicycle, and fence have an insignificant correlation to pleasant walking environments. Lastly, feature like sky is neither related to walkable nor non-walkable environments.

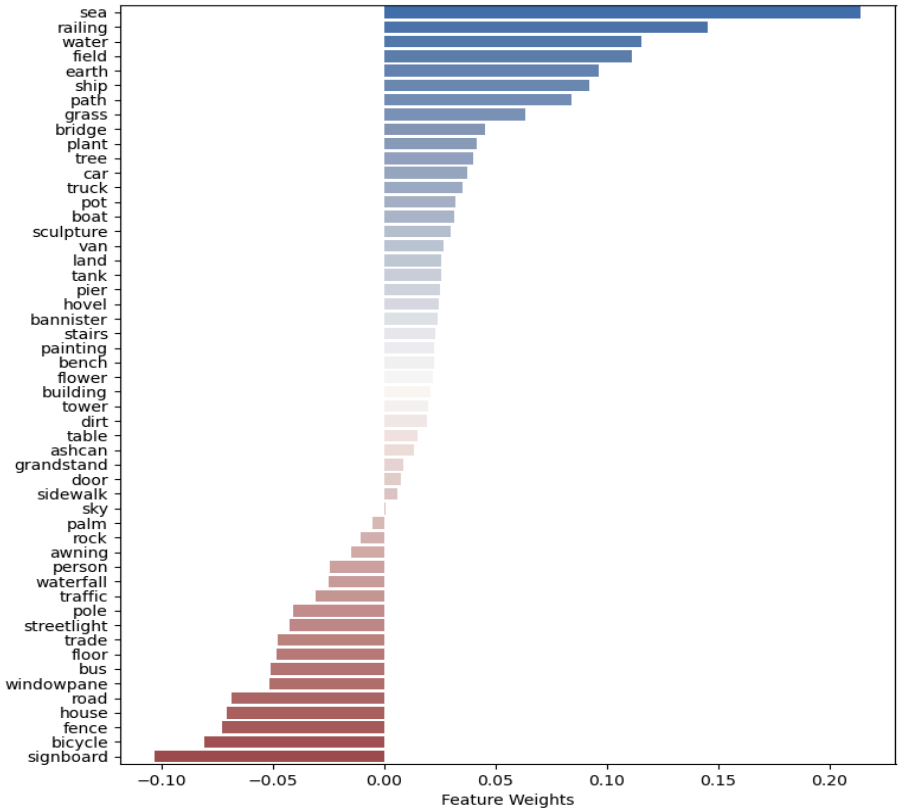

Fig. 6: Graph representing the weights of all selected features related to walkable or non-walkable environments

Finally, with the weights trained from the model, the model can be used to predict the level of walkability on a street level when individual attribute x is given, which is the proportion of related attributes from new image data taken from a new location.

3 Discussion

3.1 Advantages of the ML Method Compared to Traditional Methods

As global warming becomes inevitable and climate change worsens, cities need effective and precise reactions to mitigate the complex impacts in order to maintain livable, high-quality neighbourhoods. Due to the size and complexity of data set, traditional tools need to be expanded to AI-enhanced workflows. Our study has at least three advantages compared to traditional methods. First, our approach quantifies the attributes in the perceived built environment effectively, which, due to their complexity, cannot be achieved through an intuitive-based observation. Second, semantic segmentation is considered as a tool not to replace human observation, but to enhance the effectiveness of observation and expand towards environmental datasets. As street images are widely available, it is possible to apply our method to evaluate street space quality at a large scale. Meanwhile, we can easily reapply the trained walkability evaluation model to other urban locations by adjusting the desired parameters. This is called transfer learning, which is a common method used in the ML field when the model solves similar kinds of tasks with limited resources. On the contrary, traditional studies on the street level can rarely cover a large area and it is occasionally difficult to reproduce in other locations because of limited time and resources. Last, when intuitive data becomes measurable, the urban design at the street level can be adjusted through street views, which are generated by 3D modelling. With the same approach in our study, the street views can be evaluated to check if walkability improved based on the new design.

3.2 Challenges and Limitations

This study encountered a few challenges that will need to be addressed in subsequent trials. First, segmentation prediction is not one hundred percent accurate. For instance, one of the most recent models called Qualcomm AI Research (BORSE et al. 2021) received the third rank in the CITYSCAPE dataset with a mean intersection over Union (mIoU) with a score equal to 85.6%. Factors like input image quality or the data used to train the pre-trained model can affect the accuracy of the results. These should therefore always be considered when applying this method. In our study, the prediction of sidewalk pixels from the road is rather challenging, as the sidewalk is often similar in colour to vehicle roads, which decrease the ratios of the sidewalk on the image. This reduced the importance of the sidewalk feature related to the walkable environment (Figure 6). Similarly, streetlights are generally considered as important features in walkable areas (ZHANG 2019). However, when an image shows a walkable area with trees, it is more difficult for a computer to detect the streetlight from this image than from an image representing a non-walkable area like a motorway with clear background.Second, this paper only focused on the visual aspect of walkability; attributes related to other senses like auditory or haptic perception are not evaluated due to relatively limited data resources available. Although features like the number of available amenities and destinations are not considered in this paper, we will integrate them into our model in the future to gain a more comprehensive prediction for walkability. Finally, finding the right pre-

trained model for the specific tasks related to urban planning remains challenging. Moreover, retraining a new model which fits all the attributes of walkability requires not only expertise in the ML field but well-annotated data of high quality, which is still currently limited.

However, the study provides an example and foundation of utilizing ML for walkability evaluation. If data is collected with the required attributes, retraining the model based on existing environments with transfer learning, many urban challenges can be mitigated through data analysis and making informed predictions using ML.

4 Outlook

As urban digitization grows rapidly, novel ML methods may become a crucial tool in the field of urban and landscape planning (FRICKER et al. 2020). In order to extend the use of ML in city planning and design, new ML models designed for specific place-based applications need to be developed and critically evaluated (SALOHEIMO et al. 2021). Using our study as a foundation, a similar approach can be applied to areas like urban landscape quality evaluation or UHI effect evaluation (YAO & FRICKER 2021) at street level. In the near future, we will study additional attributes affecting walkability to refine the model and test the common attributes of UHI effects and walkability to understand their interrelationship. Lastly, the ML model can be also used in the urban design process to help predict walkability in 3D models. When 3D models generate street views as input to semantic segmentation models, the ratios of the individual attributes can be computed, and the trained retrogression model can predict to what extent and why the new design increases walkability. As a result, walkability will be able to become a key parameter in the computational design workflow of sustainable urban and landscape environments.

References

AGAMPATIAN, R. (2014), Using GIS to Measure Walkability: A Case Study in New York City. Master Thesis, Royal Institute of Technology, Stockholm, Sweden.

APPLEYARD, D. & LINTELL, M. (1972), The Environmental Quality of City Streets: The Residents' Viewpoint. Journal of the American Planning Association, 38 (2). https://doi.org/10.1080/01944367208977410.

BORSE, S., WANG, Y., ZHANG, Y., PORIKLI, F. & AI RESEARCH (n. d), InverseForm: A Loss Function for Structured Boundary-Aware Segmentation.

D'ORSO, G. & MIGLIORE, M. (2020), A GIS-Based Method for Evaluating the Walkability of a Pedestrian Environment and Prioritised Investments. Journal of Transport Geography, 82 (June 2018), 102555. https://doi.org/10.1016/j.jtrangeo.2019.102555.

FRICKER, P., KOTNIK, T. & BORG, K. (2020), Computational Design Pedagogy for the Cognitive Age. In: WERNER, L. & KOERING, D. (Eds.), Anthropologic: Architecture and Fabrication in the cognitive age. Proceedings of the 38th eCAADe Conference, Vol. 1, TU Berlin, Berlin, Germany, 16-18 September 2020, 685-692.

GEHL, J. (2011), Life between Buildings : Using Public Space. Island Press.

GOODFELLOW, I. et al. (2016), Deep Learning. MIT Press. https://www.deeplearningbook.org/contents/regularization.html.

HILLIER, B. et al. (1976), Space Syntax. Environment and Planning B: Planning and Design, 3 (2), 147-1. https://doi.org/doi:10.1068/b030147.

JACOBS, A. B. (1995), Great Streets. The MIT Press.

JACOBS, J. (1961), The Death and Life of Great American Cities. Random House, New York.

JOGLEKAR, S., QUERCIA, D. & REDI, M. et al. (2020), Facelift: A Transparent Deep Learning Framework to Beautify Urban Scenes. R. Soc. open sci., 7 (1).

JUNG, A. (2022), Machine Learning: The Basics. Springer, Singapore. https://doi.org/978-981-16-8193-6.

LITMAN, T. A. (2003), Economic Value of Walkability. Transportation Research Record, no. 1828, 3-11. https://doi.org/10.3141/1828-01.

LYNCH, K. (1960), The Image of the City. The M.I.T. Press Massachusetts Institute of Technology Cambridge, Massachusetts, London, England.

SALOHEIMO, T., KAOS, M., FRICKER, P. & HÄMÄLÄINEN, P. (2021), Automatic Recognition of Playful Physical Activity Opportunities of the Urban Environment. In: ACM International Conference Proceeding Series. https://doi.org/10.1145/3464327.3464369.

SPECK, J. (2012), Walkable City: How Downtown Can Save America, One Step at a Time. Farrar, Straus and Giroux.

VOULODIMOS, A., DOULAMIS, N., DOULAMIS, A. & PROTOPAPADAKIS, E. (2018), Deep Learning for Computer Vision: A Brief Review. Computational Intelligence and Neuroscience, 2018. https://doi.org/10.1155/2018/7068349.

WANG, R., LIU, Y. & LU, Y. (2019), The Linkage between the Perception of Neighbourhood and Physical Activity in Guangzhou, China: Using Street View Imagery with Deep Learning Techniques. International Journal of Health Geographics. https://doi.org/10.1186/s12942-019-0182-z.

YAO, C. & FRICKER, P. (2021), How to Cool down Dense Urban Environments? A Discussion on Site-Specific Urban Mitigating Strategies. Journal of Digital Landscape Architecture. https://doi.org/10.14627/537705007.

YIN, L., CHENG, Q. & SHAO, Z. (2017), "Big Data": Pedestrian Volume Using Google Street View Images, 461-69. Springer Geography. https://doi.org/10.1007/978-3-319-40902-3_25.

YIN, L. & WANG, Z. (2016), Measuring Visual Enclosure for Street Walkability: Using Machine Learning Algorithms and Google Street View Imagery. Applied Geography, 76 (2016). https://doi.org/10.1016/j.apgeog.2016.09.024.

ZHANG, X. (2019), How Street Features and Lighting Affect Neighborhood Walkability. Dissertation, The Ohio State University, Ohio. https://etd.ohiolink.edu/apexprod/rws_etd/send_file/send?accession=osu156338804759 3407&disposition=inline.