THE VALUE OF DEEP LEARNING FOR LANDSCAPE REPRESENTATION COMPARISON BETWEEN SEGMENTATION IMAGES MAPS AND GIS

F. Bianconi [1], M. Filippucci [1], S. Ceccaroni [1], M. Seccaroni[1]*

[1] Department of Civil and Environmental Engineering, University of Perugia, 06125 Perugia, Italy
fabio.bianconi@unipg.it marco.filippucci@unipg.it ceccaroni.simona@gmail.com marco.seccaroni@unipg.it

Commissione IV

KEYWORDS: Image segmentation, Big data, Perception, Automation, Deep learning

ABSTRACT:

Landscape refers to the qualities of a place, the result of a structural, territorial and environmental component, and the attribution of meanings, which is certainly the fundamental issue of the interpretative process. Percepire etymologically derives from "per", which means "by means of, through", and "capere", which translates as "to take", "to collect" (information, sensory data), "to learn". Since images are derived from the territory, it is of first interest to propose a comparison between representations derived from automated processes on photographs and the synthetic data interpreting the territory inherent in the plans developed with GIS in order to obtain a more precise perceptual analysis. The emergence of new tools for the processing and reproduction of data offers new opportunities for the knowledge and representation of the landscape, in architectural and urban contexts, and the integrative support that these processes can bring to the representation of the qualities of a place have to be reinterpreted in a Spatial Information Dataset in order to make synthetic and intelligible information. Identifying specific themes by questioning these data through criteria and placing at the centre the capacity of the digital environment in its mathematisation to compare data, transforming them into information, in an automated process is aimed at the exploitation of Big Data and the full replicability of the procedure. In this way, it is possible to enter into the analysis of the quality of space, of that notion of landscape concieved as "that part of the territory perceived by the population that lives it".

1. INTRODUCTION

The present research aims to represent the landscape, concieved in its normative meaning as "that part of the territory as it is perceived by the populations". (D'Europa, 2000). The reasons for this study are to be found in the need to deepen synthetic representations of plans, which are concerned with the qualitative values of the territorial components but do not go into these relationships with visual aspects, with data relating to perception. The randomness of the concept of landscape is a fundamental cultural challenge for our society of images (Mitchell, 1980): "everything is landscape' (Kroll, 1999), from which, however, one can transpose the paradox highlighted by Munari for art (Munari, 1978) whereby 'if everything is landscape, nothing is landscape'. The risk is the association with picturesque aspects or 'ecological' issues (Greppi, 1991), which does not take into account the centrality of living (Pizzetti, 1991). We often simplify landscape by directly associating it with 'territory', which is certainly at the genesis of its birth, but does not lead us to consider 'the relationship of the iconic type interpreted in its most general and elementary sense, which corresponds to the subject of the cognitive process'. (Farinelli, 1991). These exemplifications coincide linearly with representations of the landscape understood as picturesque images, representations exemplified in childlike graphic forms, maps of the territory correlated with information systems that report aseptic data of the perceptive aspects that are fundamental to the very definition of landscape (Figure 1). Landscape refers to the qualities of a place, the result of a structural, territorial and environmental component, and of the attribution of meanings, which is certainly the fundamental issue of the interpretative process. However, underpinning it is the centrality of the vision (Kepes, 1944), with the eye (Gregory, 1970), at the centre of our culture for more than two centuries (Panofsky, 1927), which is responsible for over 80% of our stimulation. Perceiving is etymologically derived from "per", meaning "by means of, through", and

Figure 1. Different types of landscape

"capere", which translates as "taking", "collecting" (information, sensory data), "learning".

The derived "cognitive product" (Goldhagen, 2017; Seccaroni et al., 2021; Zube et al., 1982) is the result of a series of elaboration processes on the sensory information, which are carried out in a completely automatic and implicit way (Shin et al., 2015; Smardon, 1988). The identification of a visual stimulus involves two different stages of processing (Arnheim, 1986; Kepes, 1944; Palmer, 1999): a primary and a secondary stage. In the primary stage, the structural description of the sensory stimulus takes place (in general, the analysis of the form), while in the secondary stage, the recognition of the stimulus itself takes place through a comparison ("matching") of the result of this description with the traces deposited in the memory of the same or similar objects. These two processes (description and comparison) are the main stages of visual information processing. (Goldstein, 2007; Ramachandran, 1990; Snowden et al., 2012). The first (primary

* Corresponding author

stage), involves the sensory system (in this case the visual system) and provides a description of the stimulus independently of the meaning of the object, i.e. a structural description of the object is provided. Since the attribution of meanings is a process that can only be made explicit through language (Jencks, 1969), this analysis finds objective difficulties in the synthetic interpretation of the representation. On the other hand, it becomes possible to place in parallel with the representation of the components of the territory that generate the landscape, the definition of the components of the images generated by it, to be understood as the first processing of the information perceived. This research is based on these coordinates and is concerned with analysing the impact of these techniques for interpreting the landscape, using new methods of representation. This approach is the result of a parallel study of the stimulus of the visual through biosensors. (F Bianconi et al., 2021; Fabio Bianconi et al., 2021a, 2021b; Seccaroni et al., 2021). The method developed here is based on image analysis (Warburg et al., 2020), digitally understood as perceptual data sets. By means of deep learning processes (Dubey et al., 2016) it is possible to define an image segmentation (Zhang et al., 2018), which is then associated with entities (Johnson and Jozdani, 2018), and can be located on a map. The great challenge inherent in this process concerns the interpretation of the data, which does not have to be demonstrated but verified.

Since images are derived from the territory, it is of primary interest to propose a comparison between representations derived from automated processes on photographs and the synthetic data interpreting the territory inherent in GIS planimetric representations. The thesis that is argued is that the use of deep learning can deepen the identification of those perceptual qualities of a territory that the interpretation of spatial components cannot bring out. The capacity of the digital environment in its mathematisation to compare data is fundamental to transform them into information, in an automated process aimed at the exploitation of Big Data and the full replicability of the procedure. The study therefore intends to demonstrate the integrative support that such processes can provide in the representation of the qualities of a place, which must however be reinterpreted in a Spatial Information Dataset in order to make the information synthetic and intelligible.

2. METHOD

The image analysis process is based on the use of heterogeneous photos (Figure 2) and their semantic segmentation (Ye et al., 2019) offered by Mapillary (Neuhold et al., 2017), which allows a photo to recognise the elements present and their quantity in the image, which can be associated with the impact on vision. The use of automated tools has also been employed in other case studies (Meng et al., 2020) but unlike the proposed mode they use neither data nor open source programs (Salazar Miranda et al., 2021). In fact, the most common use of semantic segmentation for the analysis of urban spaces is through the use of satellite photos. (Geiss et al., 2020).

The interpretation of data by criteria is illustrative for understanding the value of datasets: the much debated topic (Badland and Pearce, 2019; Forsyth, 2015; Schmitz and Scully, 2006) is usually analysed in relation to infrastructure (Badland and Pearce, 2019), optimised for objective values, such as travel time or the shortest route, or taking into account the difference in altitude, but not usually studied by criteria that can analyse what is seen, using images through an interpretation of the present elements. Using a new request and imposing the key image as a parameter, Mapillary returns a json file. Within the file, for each photo we get the geographical location where the photo was taken and the value with the relative area of the semantic segmentation of the image.

Figure 2. Some images used for Pietrafitta case study.

The next step is to filter the data, i.e. if there is a dashboard value within the image this area is removed from the summation and the remaining areas are extended proportionally so that the sum is always close to 1. A list was created with a DataFrame for each photo analysed. The DataFrame consists of imagekey, geographical coordinates of the photo, value and segmentation area. The DataFrame thus created was used for the analysis of the average elements present and their interpretation against the criteria. From the segmentation and subsequent analysis of the images, it is possible to represent the distribution according to area of the main elements such as "Building" "Road" "Sky" and "Vegetation", and the relationships with all the other secondary elements that the semantic segmentation returns. Through a criterion that composes the main elements extracted from the image associated with a value, it is possible to define a summary value that we call "landscape value". With this evaluation, through a further script in pyton, it was possible to obtain a map representing in green scale the best routes from the landscape point of view and in black the worst ones. It is thus possible to have a summary data that from the interpolation of the data of the different images, linked to routes, and it is possible to identify the qualities perceived in a place (Figure 3). The implemented process enhances the medial value of the images, elements of connection between the territory and the vision, concieved as the foundation of interpretation. The resulting spatial data can be compared with different information collected independently of perceptual aspects. For this reason, the traces derived from the Machine Learnjng processes are reported in a GIS environment, using open source software (QGIS), and open source data (open street map), to develop a comparison between the information sets in order to bring out a knowledge of the qualities of the place. The process is tested in two case studies, one concerning an urban centre (Orvieto) and one concerning productive settlements (Pietrafitta). The experiment was divided into two phases: data acquisition and subsequent processing using the developed algorithm. In the case study of Pietrafitta, the data acquisition part was carried out by means of an on-site inspection, walking and driving routes, using a smartphone and the Mapillary database with the setting of a shot every 5 seconds to create a series of image sequences.

Figure 3. Representation of the slow mobility analysis applied to Pietrafitta.

These tracks concerned in particular the downstream part of the village of Pietrafitta in correspondence with the former Enel power station, an area deeply marked by productive settlements that have generated a significant fragmentation of the territory, leading to marginalisation, abandonment and degradation. During the exploration, the images acquired were first saved on the smartphone and then uploaded to the Mapillary database. The data acquisition phase resulted in approximately 1633 images of the area which were then processed by the developed algorithm. In the case study of Orvieto, on the other hand, this first phase was carried out using a different methodology that did not involve an inspection, but the approximately 5870 images were acquired using Mapillary.

The data were processed using python processes and the results imported into Qgis maps, which already contained spatial information on the structural elements of the landscape, divided into various categories: buildings, roads, water features and vegetation, which in turn was identified as trees - land with few trees and other vegetation.

3. RESULTS AND DISCUSSIONS

The results derived from the experimentation are described in correspondence with the objective of showing the validity of this process in the integration of the representation of the spatial components of the landscape. It is therefore a question of starting from the denotative values of a territory, which can be associated with the result of the landscape, that can however be deepened in their validity with such digital processes.

In the case study of Pietrafitta, it is immediately evident that there is a prevalence of agricultural areas with portions of built-up areas for productive use that are mostly abandoned, as well as the presence of infrastructures that are inserted within the area. From this impact one can certainly understand that the quality and perception of the place is not optimal. Comparing the GIS spatial data with the Deep Learning maps, it can be seen that the part of the area in which there is a prevalence of road infrastructures and disused industrial plants the landscape value has a low index, while in the part in which there is a prevalence of agricultural areas and the roads have less impact on the territory the perception takes on a scale of higher values.

In the case study of Orvieto, on the other hand, given its territorial configuration, large built-up areas surrounded by dense vegetation can be seen in the central area, while further downstream there are agricultural areas and important road infrastructures that fit into the territory.

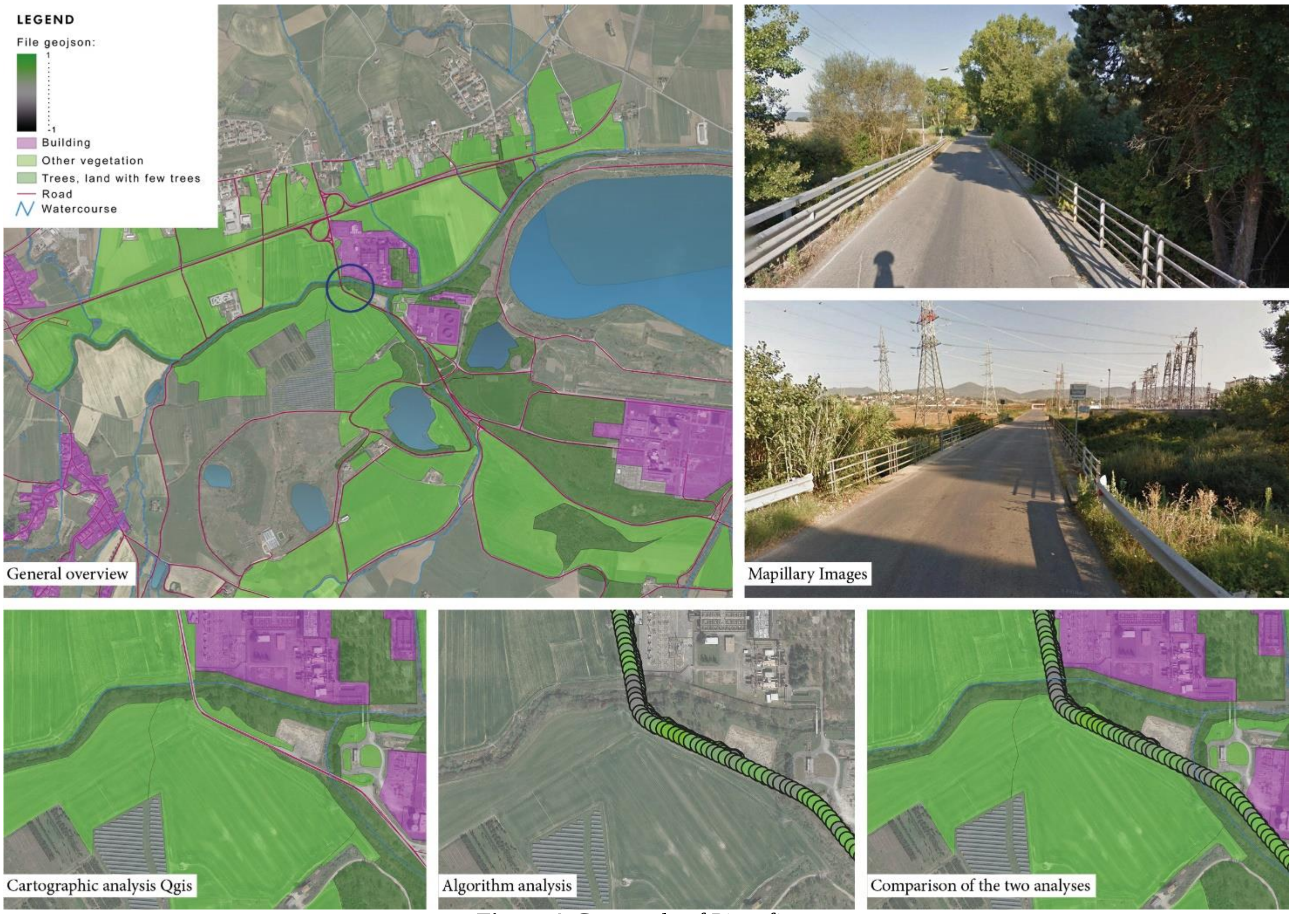

Figure 4. Case study of Pietrafitta.

Figure 5. Case study center of Orvieto

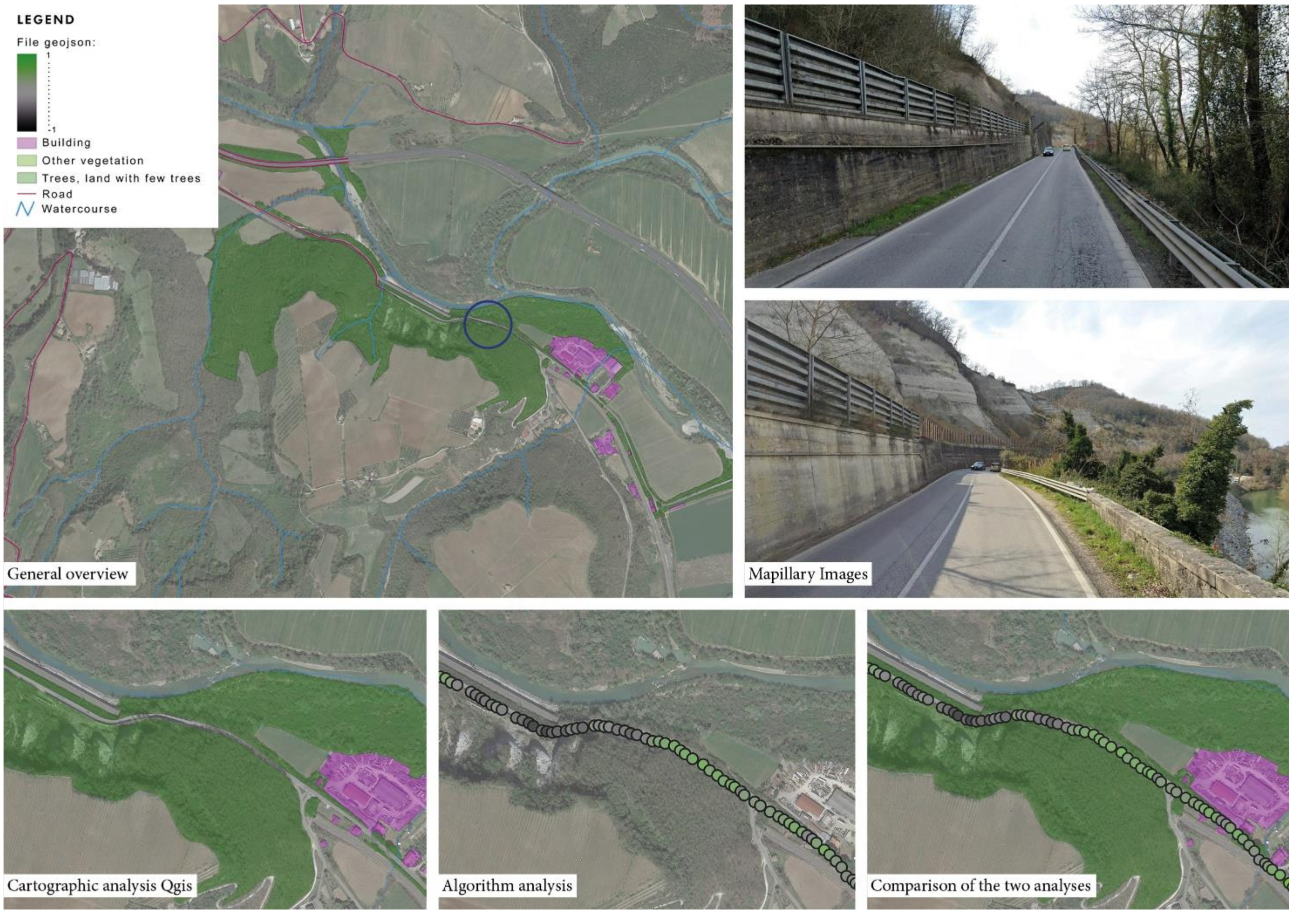

Figure 6. Case study of Baschi

The obtained perception is certainly positive, in fact when comparing the analysis obtained from the algorithm the "landscape value" takes on mostly positive values. Going into detail, however, in both cases there are small errors in the cartographic analysis of Qgis. In particular, for example in Pietrafitta (Figure 4) in the area in front of the diesel tanks belonging to the former Enel power station there is an area dominated by vegetation and plots of agricultural land, but in reality the analysis of the algorithm shows that the perception takes on a scale of negative values. In fact, the road is bordered by a high guardrail and the adjacent diesel tanks intrude into the area, so the impact is different. Similar cases occur in the case study of Orvieto, specifically in the city centre and near Baschi scalo. Starting from the old town centre (Figure 5), which is mainly made up of buildings and streets, the perception when observing the maps is of a place dominated by buildings where a sense of suffocation prevails.

On the other hand, when analysing the images through the algorithm, the perception of the place is totally different and assumes a higher scale of values since in reality the streets within the historic centre are not real streets characterised by asphalt and guardrails, but are streets and small squares made with cobblestones where people can walk freely and create points of aggregation. Finally, in the part of Baschi scalo (Figure 6) there is instead, a stretch similar to Pietrafitta; the road is surrounded by scattered buildings, of minor importance and there is a prevalence of vegetation for which there is a positive perception of the area. The algorithm, however, shows that the carriageway is bordered by guardrails and a retaining wall, so the vegetation is at a higher level than the road surface and the perception of the area changes radically.

4. CONCLUSIONS

The aim of the experiment was to apply research tools to provide a clear representation of the perception of the qualities of an area by comparing it with GIS data, which is potentially useful for understanding possible relationships between the elements of an area. The aim of this work was to discuss how the emergence of new tools for data processing and representation offers new opportunities for landscape knowledge and representation in architectural and urban contexts.

At the heart of the research is the value of information digitalization processes, which are based on Big Data and can easily be correlated with a territory. Once the data had been acquired, concerning image and vision, it was processed through an algorithm. The proposed strategy proved to be able to extract connotative information on the quality of places from the images, interpreting and evaluating the acquired data and graphically returning different results. Qgis, in fact, is a computer system that has made it possible to associate and identify objective data evaluated on a two-dimensional plane. Mapillary and the analysis through the algorithm, on the other hand, made it possible to obtain a new image dimension. Studies have revealed that the combination of both tools gives the possibility of obtaining a more precise perceptual analysis. The data that emerge, processed by means of the algorithm, are averaged over the results of the image analysis and the Qgis maps, then summarised by means of graphic representations. It is then possible to analyse some of the most significant zones/elements.

This process made it possible to analyse and edit spatial data and generate cartographies. The interpolation of the two systems made it possible to extract connotative information on the quality of places from the images. This methodology, susceptible to some improvements such as the use of Deep Learning with higher performance and a higher degree of recognition and accuracy, provides a clear representation of the perception of the qualities of an area and opens up a full comparison with GIS data, potentially useful for understanding possible relationships with elements of an area. Similarly, the data can be analysed through critical categories to which an interpretative and evaluative value can be attributed, demonstrating that it is a versatile and high-impact system for detecting urban qualities, with a much wider application potential.

Detecting perception, assessing the impact of a landscape, analysing which places create wellbeing, are complex interpretations, often theorised, sometimes intuited, but only few times quantified and qualified. The relationship between man and space has been studied through the creation of a replicable digital procedure which, starting from data, proposes interpretations of place values. The innovative approach to the representation of emotional states reveals how and how much places influence our wellbeing and health, and how the interpolation of innovative tools guarantees even more truthful analyses as well as the introduction of a new spatial dimension to the images.

REFERENCES

Arnheim, R., 1986. *New essays on the psychology of art*. University of California Press, Berkeley.

Badland, H., Pearce, J., 2019. Liveable for whom? Prospects of urban liveability to address health inequities. *Soc. Sci. Med.* https://doi.org/10.1016/j.socscimed.2019.05.001

Bianconi, Fabio, Filippucci, M., Magrini, G., Seccaroni, M., 2021a. DESIGNING WITH EMOTIONAL AWARENESS. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* XLVI-M-1–2, 55–62. https://doi.org/10.5194/isprs-archives-XLVI-M-1-2021-55-2021

Bianconi, Fabio, Filippucci, M., Seccaroni, M., 2021b. New Interpretative Models for the Study of Urban Space, in: DISEGNO. FrancoAngeli, pp. 235–240. https://doi.org/10.3280/oa-686.37

Bianconi, F, Filippucci, M., Seccaroni, M., Aquinardi, C.M., 2021. Urban parametric perception. The case study of the historic centre of Perugia. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* XLIII-B2-2, 839–846. https://doi.org/10.5194/isprs-archives-XLIII-B2-2021-839-2021

Convenzione Europea sul Paesaggio 2000.

Dubey, A., Naik, N., Parikh, D., Raskar, R., Hidalgo, C.A., 2016. Deep Learning the City : Quantifying Urban Perception At A Global Scale. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* 9905 LNCS, 196–212.

Farinelli, F., 1991. L’arguzia del paesaggio. *Casabella* 575–576, 10–12.

Forsyth, A., 2015. What is a walkable place? The walkability debate in urban design. *Urban Des. Int.* 20, 274–292. https://doi.org/10.1057/udi.2015.22

Geiss, C., Zhu, Y., Qiu, C., Mou, L., Zhu, X.X., Taubenbock, H., 2020. Deep Relearning in the Geospatial Domain for Semantic Remote Sensing Image Segmentation. *IEEE Geosci. Remote Sens. Lett.* https://doi.org/10.1109/LGRS.2020.3031339

Goldhagen, S.W., 2017. *Welcome to Your World. How the Built Environment Shapes Our Lives*. HarperCollins, New York.

Goldstein, E.B., 2007. *Sensation and perception*.

Thomson/Wadsworth, Belmont.

Gregory, R.L., 1970. *The Intelligent Eye.* McGraw-Hill Book Company, New York.

Greppi, C., 1991. Guardare con meraviglia. *Casabella*.

Jencks, C., 1969. *Meaning in architecture.* The Cresset Press, London.

Johnson, B.A., Jozdani, S.E., 2018. Identifying generalizable image segmentation parameters for urban land cover mapping through meta-analysis and regression tree modeling. *Remote Sens.* 10. https://doi.org/10.3390/rs10010073

Kepes, G., 1944. *Language of vision*. Paul Theobald, Chicago.

Kroll, L., 1999. *Tutto è paesaggio*. Testo & immagine, Venezia.

Meng, Y., Xing, H., Yuan, Y., Wong, M.S., Fan, K., 2020. Sensing urban poverty: From the perspective of human perception-based greenery and open-space landscapes. *Comput. Environ. Urban Syst.* 84. https://doi.org/10.1016/j.compenvurbsys.2020.101544

Mitchell, W.J.T., 1980. *The Language of images*. University of Chicago Press.

Munari, B., 1978. *Disegnare un albero.* Zanichelli, Bologna.

Neuhold, G., Ollmann, T., Rotabuì, S., Kontschieder, P., Research, M., 2017. The Mapillary Vistas Dataset for Semantic Understanding of Street Scenes.

Palmer, S.E., 1999. *Vision science: photons to phenomenology*. MIT Press, Cambridge.

Panofsky, E., 1927. Die Perspektive als “symbolische Form.” *Vor. der Bibl. Warbg.* xxii, 258–330.

Pizzetti, I., 1991. Luoghi della coscienza paesaggistica. *Casabella* 50–51.

Ramachandran, V.S., 1990. Visual Perception in People and Machines, in: Blake, A., Troscianko, T. (Eds.), AI and the Eye. Wiley, London.

Salazar Miranda, A., Fan, Z., Duarte, F., Ratti, C., 2021. Desirable streets: Using deviations in pedestrian trajectories to measure the value of the built environment. *Comput. Environ. Urban Syst.* 86. https://doi.org/10.1016/j.compenvurbsys.2020.101563

Schmitz, A., Scully, J., 2006. *Creating walkable places: compact mixed-use solutions.*

Seccaroni, M., Aquinardi, C.M., Bettollini, E., 2021. Urban Atmospheres: Representation of Intangible Relations, in: Lecture Notes in Civil Engineering. Springer Science and Business Media Deutschland GmbH, pp. 1111–1128. https://doi.org/10.1007/978-3-030-59743-6_52

Shin, Y. Bin, Woo, S.H., Kim, D.H., Kim, J., Kim, J.J., Park, J.Y., 2015. The effect on emotions and brain activity by the direct/indirect lighting in the residential environment. *Neurosci. Lett.* 584, 28–32. https://doi.org/10.1016/j.neulet.2014.09.046

Smardon, R.C., 1988. Perception and aesthetics of the urban environment: Review of the role of vegetation. *Landsc. Urban Plan.* 15, 85–106. https://doi.org/10.1016/0169-2046(88)90018-7

Snowden, R.J., Thompson, P., Troscianko, T., 2012. *Basic vision: an introduction to visual perception*. OUP Oxford, Oxford.

Warburg, F., Hauberg, S., López-Antequera, M., Gargallo, P., Kuang, Y., Civera, J., 2020. Mapillary Street-Level Sequences: A Dataset for Lifelong Place Recognition, in: Computer Vision and Pattern Recognition (CVPR).

Ye, Y., Zeng, W., Shen, Q., Zhang, X., Lu, Y., 2019. The visual quality of streets: A human-centred continuous measurement based on machine learning algorithms and street view images. *Environ. Plan. B Urban Anal. City Sci.* 46, 1439–1457. https://doi.org/10.1177/2399808319828734

Zhang, F., Zhou, B., Liu, L., Liu, Y., Fung, H.H., Lin, H., Ratti, C., 2018. Measuring human perceptions of a large-scale urban region using machine learning. *Landsc. Urban Plan.* 180, 148–160. https://doi.org/10.1016/j.landurbplan.2018.08.020

Zube, E.H., Sell, J.L., Taylor, J.G., 1982. Landscape perception: Research, application and theory. *Landsc. Plan.* 9, 1–33. https://doi.org/10.1016/0304-3924(82)90009-0