room_ID: An Architectonic Image Classifier Tool Correlating Machine Learning and the Domestic Space

Alex Nogueira[1], Luís Romão[2]

[1] Researcher at CIAUD, Research Centre for Architecture, Urbanism and Design, Lisbon School of Architecture, Universidade de Lisboa, Rua Sá Nogueira, Polo Universitário do Alto da Ajuda, 1349-063. Lisboa, Portugal, and Professor at Universidade Federal de Mato Grosso do Sul (UFMS), Curso de Arquitetura e Urbanismo. Cidade Universitária, Av. Costa e Silva, 79070-900, Campo Grande, Brazil
alex.nogueira@fa.ulisboa.pt; alex.nogueira@ufms.br
[2] Associate Professor at Universidade de Lisboa, Faculdade de Arquitetura, CIAUD, Research Centre for Architecture, Urbanism and Design, Lisbon School of Architecture, Universidade de Lisboa, Rua Sá Nogueira, Polo Universitário do Alto da Ajuda, 1349-063. Lisboa, Portugal
lromao@fa.ulisboa.pt

Abstract. There are still considerable gaps in the relationship between artificial intelligence (AI) and architecture, both in the field of computer-aided architectural design and the everyday spatial experience. Therefore, we sought to build a bridge through correlating machine learning (ML), one of the significant branches of AI, and the domestic space, probably the most experienced architectonic space. Thus, our utmost goal is to develop an architectonic image classifier tool that allows the computer to identify the rooms that generally compound an ordinary dwelling. Our approach includes a brief theoretical background, the development of the *room_ID* app, and, afterward, discussions are presented. In the stage of the *room_ID* application development, a five-phase framework is proposed: a) Class definition; b) Database criteria; c) Image collection; d) Tool development; and e) Preliminary validation. Therefore, the *room_ID* tool provides us with a possible way to recognize the specificities of architectonic spaces rendered as computational data.

Keywords: Artificial Intelligence, Machine Learning, Domesticity, Architectonic image classifier tool, Room labels.

1 Introduction

It is a common-sense state that artificial intelligence (AI) is present in a large set of knowledge domains, with several implications for society (Oliveira, 2017).

Since its very first moment, a multidisciplinary approach was required, including philosophy, mathematics, economics, linguistics, computer engineering, and cybernetics, among others (Russell & Norvig, 2010). In architectural spatial experience, AI rarely is a focal point. In a broad view, meaningful examples show the house (or other architectonic programs) as a shelter (or bunker) for technology (Furtado & Moreira, 2001) and not just as part of its proprieties.

In a correlatable approach, Ted Krueger (2006) suggests: *Rather than having an architecture that contains smart bits, we might consider an architecture that possesses intelligence* (p. 82). It no longer causes surprise or astonishment when a home is equipped with a personal assistant which uses natural language processing (NLP), for instance. Nevertheless, this sort of device has limited spatial comprehension or, in some aspects, is nonexistent. Such spatial understanding, as input, could enrich the device outputs and expand its functions.

Rodney Brooks (1991), around works on robotics, introduces new concepts for AI, namely: *situatedness*, *embodiment*, *intelligence*, and *emergence*. For our proposes, the idea of situatedness is especially relevant. Nevertheless, while Brooks proposes that agents should not use an abstract model of the world, but one must use the real world as a reference, being able to situate themselves within the very world (aids by sensors); in this context, we propose an architectonic spatial recognition, using the apparatus of machine learning (ML), as an alternative way of situatedness.

Differently from Brooks, our approach regresses to the symbolic world, one labeled world from taking as reference the real one and being able to inform the agent in which class of space it is. Hence, we justify this research as a first effort to transform particular aspects of dwellings into computational data. In a second moment, starting from this recognition, we could glimpse the emergence of systems that seek to assist specific domestic functions and activities.

Our general goal is to promote a sort of computer *situatedness* concerning the dwelling context, where the intelligent agent can classify which part of the dwelling one is through image recognition. Consequently, the main goal is to develop an architectonic image classifier tool, an application named *room_ID*, based on machine learning (ML) apparatus, that allows the computer to be able to classify room labels appropriately, using a mobile device.

Beyond an initial theoretical background and the posterior discussion about the results, the methodology has a framework divided into five main phases: a) Class definition: general analysis of domestic needs and dwelling compartmentalization, in order to set adequate labeling of the rooms; b) Database criteria: establish criteria for selecting the images that shall make up the database; c) Image collection: data collection and shaping of parameters to training and test the algorithm; d) Tool development: configuration and processing of the stages of training, test, programming and, installation of the app, supported by *MIT App Inventor* (n.d.); e) Preliminary validation: first use of the tool in real situations.

As related works, we can mention Wang et al. (2019) by pointing out similarities in the legibility of architectural indoor space between humans and a Deep Convolutional Neural Networks (DCNN) model, with a detailed description of both the technical apparatus and the computational architecture, having as case studies two Parisian train stations. Sheraz Ahmed et al. (2012) developed *an automatic system for analyzing and labeling architectural floor plans* (p. 339) concerning architectural graphic representation. Wu et al. (2018) describe the development of a tool capable of understanding and reconstructing the 3D structure of an object based on a single image (real or synthetic image).

In general, this paper presents a tangency between architecture and AI, by correlating domestic spaces and ML using image collection of existent rooms compounding a learning database, which allows the computer to statistically predict where it is in the house. We understand that allowing the computer to process this type of architectural information, as delivered by *room_ID*, can be an essential initial step in developing future applications that seek in this relationship a way to meet demands and explore potentialities around domesticity.

2 Theoretical Background

To establish the tangency between architecture and AI, we are following an approach where *Artificial Intelligence is fundamentally a statistical approach to architecture* (Chaillou, 2019, p. 15). Nevertheless, at the same time, in this paper, we have in mind an overall idea, as suggested by José Miranda (2006): *it is about regarding it as an experimental space on the limits of human inhabitation when this starts to confound itself with the mathematical matrix* (p. 106).

In the architectural domain, our focus regards the domestic space, in a slant based on the user needs and space requirements, as a means to properly establish classes of rooms (labels). In the AI universe, we follow the ML methods with neural networks (NN), more specifically Convolutional Neural Networks (CNN), and supervised learning techniques in order to model an image classifier. The sum of both disciplines, supported by several authors, allows us to launch the theoretical framework for what we call an architectonic image classifier tool.

2.1 Domesticity

Briefly, we can say that in Western society, the idea of domesticity has, during the Eighteenth-Century, undergone significant changes (for several reasons). By that time, differently from the prior centuries, both public and private dimensions began to be differentiated (Aries & Duby, 2009). Furthermore, the Eighteenth-Century was marked by bourgeois individualism (Furtado & Moreira, 2001), which brings modifications to domestic spatiality, with the emergence of more specialized intimate spaces (Rybczynski, 2009).

From this context comes the proposal of French architects Viollet-Le-Duc and César Delay, who seek to organize the domestic interior based on three different sets of activities: 1) activities related to the public dimension of domestic life; 2) activities related exclusively to the family nucleus; 3) activities related to domestic services (Pereira, 2004).

These correlations between activities and space are a fundamental aspect of the configuration of the dwelling, and, consequently, they guide our approach to selecting the room labels. To João Branco Pedro (2000), *space* is an area designed to develop a function (or a set of functions). Nevertheless, the author still shows us that there are other ways to group domestic activities and spaces. Besides that, Nuno Portas (1969), in the late 1960s, proposed a relevant study grounded on the analysis of the functions and activities that take place in the Portuguese dwelling. In this study, Portas sought to establish the minimum areas necessary for dwelling. For us, the list of the 16 main functions and activities proposed by Portas becomes particularly important. With some considerations, the list becomes the reference base to establish a register of the most representative rooms of a dwelling.

Therefore, 16 activities are grouped into 1 – sleep: personal rest; 2 – food: preparation; 3 – food: ordinary meals; 4 – food: formal meals; 5 – living: meeting and free time; 6 – living: receive; 7 – recreation: children; 8 – recreational study: young people; 9 – recreational work: adults; 10 – clothes treatment: ironing and sewing; 11 – clothes treatment: washing; 12 – clothes treatment: drying; 13 – personal hygiene; 14 – permanence in external space; 15 – communication, separation; and, 16 – storage: clothes and various items (Portas, 1969).

2.2 Machine Learning

Before we talk about ML, it is significant to contextualize AI, a branch of Computer Science, and, officially, had its term created in the summer of 1956, with the conference held at Dartmouth College (Coelho, 1996; Oliveira, 2017; Russell & Norvig, 2010). On the whole, we can say that AI is centered on mathematical logic, seeking solutions to problems that require some kind of intelligence to be performed (Coelho, 1996). ML is one of the main fields of AI (Marcus, 2020), which still includes NLP, robotics, planning, and games, among others.

ML is a set of methods (and algorithms) based on Statistics, which can learn from preview examples. Some authors, such as Max Welling (2015), consider ML and Statistics complementary disciplines. For learning, we can understand that *an agent is learning if it improves its performance on future tasks after making observations about the world* (Russell & Norvig, 2010, p. 693). Many algorithms can do it, such as perceptron, learning decision trees, and probably approximately correct (PAC) learning, among others.

In essence, we have three main ways of learning: unsupervised learning, reinforcement learning, and supervised learning. We are following supervised learning methods, to Stuart Russell and Peter Norvig (2010), *in supervised learning the agent observes some example input-output pairs and learns a* *function that maps from input to output* (p. 695). Hence, when we create labels and train them with correct examples (matching images, in our case), we teach the algorithm what is right or wrong in our topology.

Concisely, this is the ML apparatus involved with the architectonic image classifier tool *room_ID*. This AI application is based on ML methods, where CNNs are trained and tested through supervised learning means. Thereby, *from a collection of input-output pairs, learn a function that predicts the output for new inputs* (Russell & Norvig, 2010, p. 693). The selected images, correlated with the established room labels, form the input-output pairs, through which it is made statistically possible to predict among these labels in new room samples (new inputs).

3 *room_ID* Development

The methodological framework demands that the application must have a formalized room label (a), criteria for selecting images (b), organize the training and test database (c), application development in four phases (d), and first validations (e). The framework is divided as follows:

a) Class Definition: To build an image classifier, the first step is to define the labels, or classes, in which the new inputs will be classified into the corresponding ontology. In our experience, the classes need to have enough similarities to belong to the same "family" and, at the same time, significant singularities that make one different from another; consequently, being subject to comparative classifications. Our "family" base encompasses the main dwelling compartments (see Figure 1), grounded on what was covered in "*2.1 Domesticity*." Therefore, we select four main activities groups: personal hygiene, sleep, food, and living. Safeguarding regional differences and even personal freedom of configuration and use of domestic spaces, the four categories are 1 – Bathroom, 2 – Bedroom, 3 – Kitchen, and 4 – Living room.

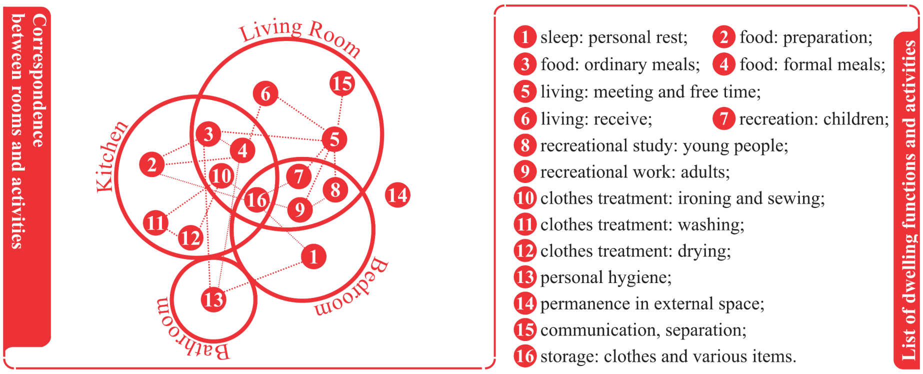

Figure 1. Room labels, based on the list of dwelling functions and activities, proposed by Nuno Portas (1969). Source: our authorship.

b) Database Criteria: Once we have the classes defined, we need to establish some criteria for selecting the images that will feed the databases of each room class, as can be seen in Figure 2. Other selection criteria were: only internal images of Lisbon dwellings (houses and apartments) available in real estate agency search engines available on the internet (Imovirtual, n.d.); the images had to be colored, without watermarks or logos; both positions, portrait or landscape were accepted, however much the landscape format prevailed; the images had their dimensions adjusted to a small size (to fit within 854 x 480 pixels). Based on the general hypothesis that the offer presented in the search engine of the real estate agency, to some extent, may reflect the diversity of types, ages, and patterns of the existing houses in the city, we try to make a selection of images that preserves this diversity.

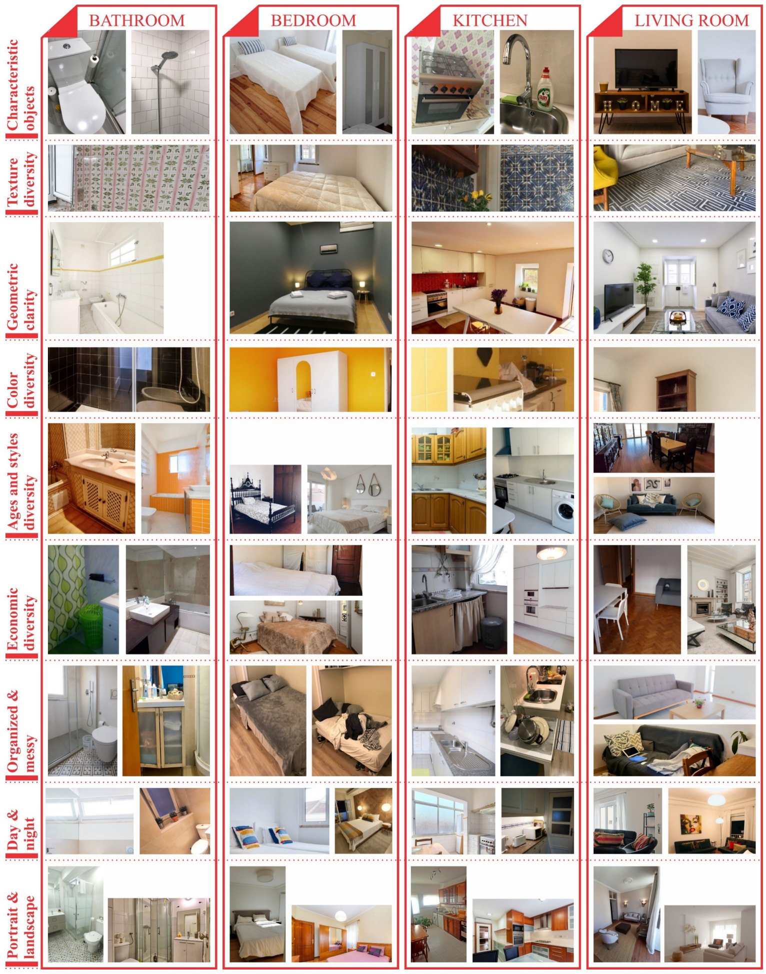

Figure 2. Illustrative examples of database criteria for room labels. Source: our authorship.

c) Image Collection: It is relevant to explain that the furniture, fixed parts (such as kitchen sink or sanitary pieces), and textures (tiles, print on fabrics, and details as colors) have a crucial role in characterizing the domestic rooms (Figure 2). The inclusion of similar types of furniture in different groups, (table and chairs set, for instance), present both in the living room and in the kitchen, is a conduct to avoid premature or false results. This is a strategy so that the tool does not associate these pieces of furniture to just one label (living room or kitchen), but as a possibility for both labels to be differentiated with the comparative analysis of other qualities. In the same sense, when organizing the collections, we were careful to preserve some ambiguities commonly found in the dwelling, such as wooden floors (more present in the bedroom and living room), decorative tiles (kitchen and bathroom), curtains (bedroom and living room), wooden cabinets (kitchen, bedroom and living room), among others.

d) Tool Development: This development is reported in four phases: *training*, *testing*, *programming*, and *downloading* (installation). The first two phases are developed through the *Personal Image Classifier* (PIC) extension (n.d.), while the third and part of the fourth ones use the *MIT App Inventor* (n.d.) platform. Figure 3 gives an overview of development. In the second phase – *testing*, we used a database with ten images per label and used the same test images in all scenarios (125, 250, 500, and 1000 samples). It is important to note that the images used for testing cannot be the same as those used in training to prevent addictive results. It is interesting to observe that the accuracy of forecasts has improved as the number of training images was increased, but this evolution has not always been linear.

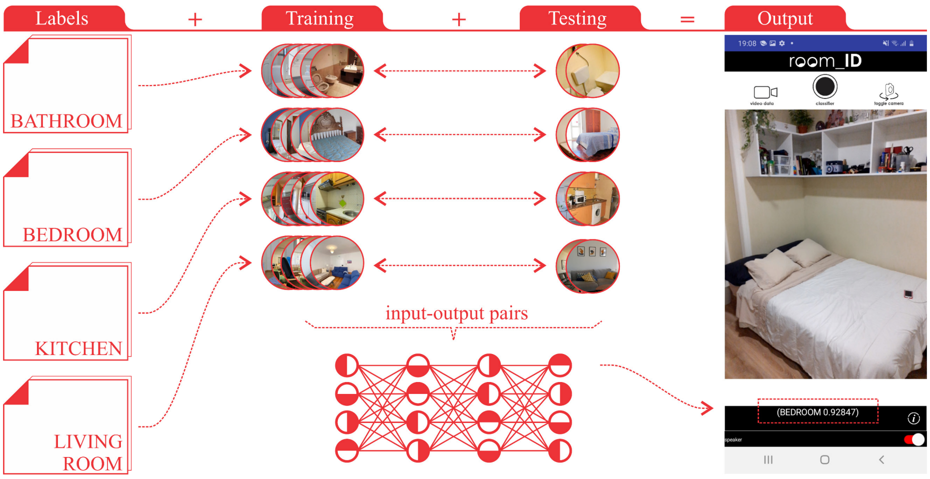

Figure 3. *room_ID* general schemata, with inputs-outputs provided, allowing to predict the output for new inputs. Source: our authorship.

The model trained with 1000 samples per label and tested with ten images per label was exported to the *MIT App Inventor* (n.d.) platform in ".mdl" format.

Accordingly, the third phase – *programming* begins. It is possible to import the extension mentioned on the platform, configure the layout, and add functions to the *room_ID* app. There is also a need to work on application programming using block language available on *MIT App Inventor*. After that, we configured how outputs should be displayed (see Figure 4). Finally, the fourth and final phase – *downloading* corresponds to availability for installation. The *MIT App Inventor* platform allows exporting the model in ".apk" format, which can be introduced in a compatible app store.

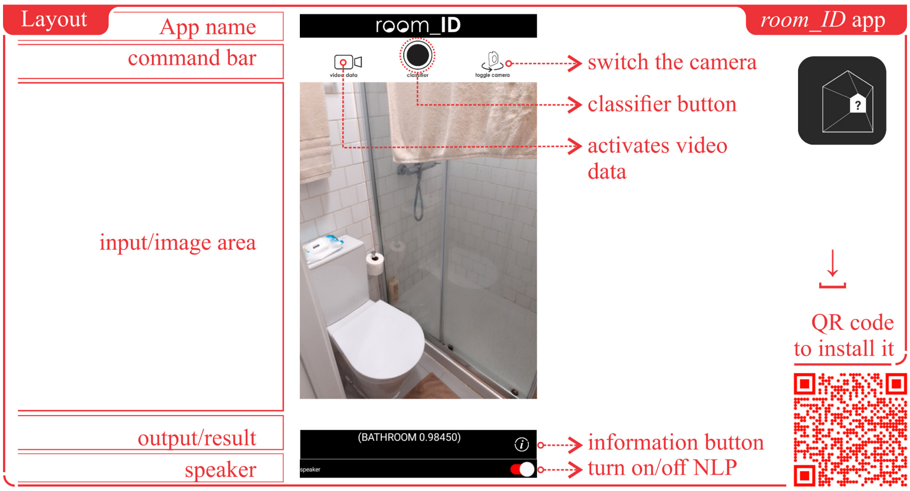

Figure 4. Layout and features of the *room_ID* application and QR code to install the app. Source: our authorship.

e) Preliminary Validation: Only with the application available for installation it was possible to verify its efficiency more broadly and pursue to validate it through these preliminary results. With the support of contributors, who, having accepted invitations, installed and tested the app in their own homes, we obtained feedback from ten cases. A small number, but more than enough to simply test whether the tool works or not. Half of the cases are in Lisbon, and the other half of the cases are spread among Campo Grande (Brazil), Luanda (Angola), and Amman (Jordan). The purpose of including different locations is to observe the application's performance in scenarios differing from that one where its databases have come from (see Figure 5).

The assessment made by the tool consists of state, based on statistical analysis, with which room label the new input (image seen through the mobile device) has higher correspondence. Therefore, as one can see in Figure 5, the five evaluation cases in Lisbon provide a total of 20 possible classifications (four per case), out of which the tool got 19 correct answers (95% of hits), while the other worldwide five cases also offered 20 possible classifications, with 17 correct answers (85% of hits). Considering all ten cases, we have 40 evaluation opportunities, out of which the application got 36 correct answers, corresponding to an overall hit rate of 90%.

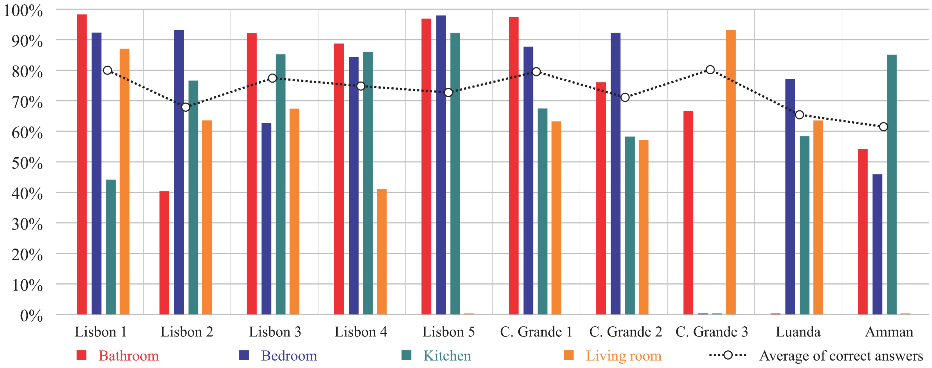

Figure 5. Hit rate per label in each case, and the average score of ratings. Source: our authorship.

As a general example of the data collected, we organized five assessments that sought to identify the Living room label (see Figure 6). The selection of these specific cases was made to show both the diversity of rooms that permeated our cases (including partially the diversity of countries) and examples of errors and successes. When analyzing the errors, we have furniture that led to ambiguity and probably led to the error. For instance, in the case of Lisbon 3, despite being in the living room, the great emphasis given to the table and chairs in the evaluation image made the app "think" it was in the kitchen; while in the case of Amman, the generous sofa was, probably, mistaken for a bed.

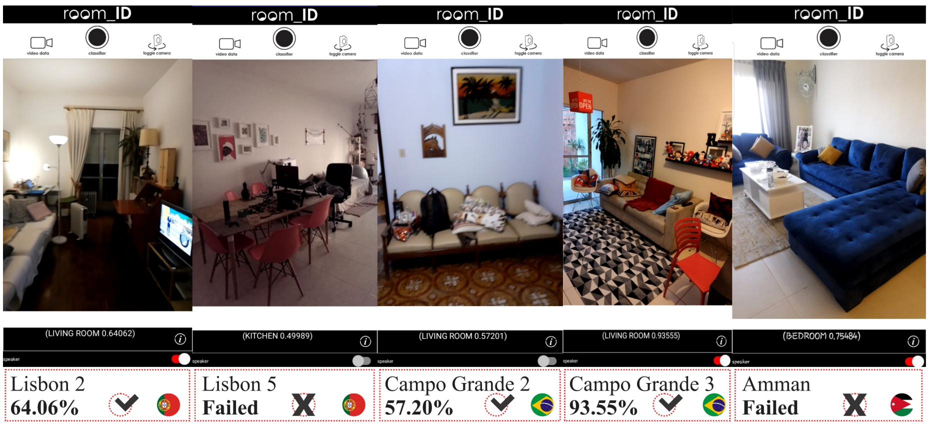

Figure 6. The *room_ID* application tries to label five cases of living rooms. Source: our authorship.

4 Results

We introduced *room_ID*, an application that mixes architecture and AI. An architectonic image classifier tool that uses the ML methods, with CNN, and supervised learning methods, to differentiate among four classes of domestic rooms. These room labels were established following a list of domestic functions and activities, among other considerations about domesticity.

The development of the tool observes steps and criteria that allow structuring the ML apparatus to meet specificities of architectural interest within the logic and limitations of an image classifier. In this way, and according to the preliminary data we reached, we can say that the tool effectively recognizes the four types of dwelling interiors aforementioned when analyzing images different from those used in its learning stage.

If, on the one hand, we conclude that both operation and performance of the application are satisfactory, on the other hand, we keep aware once the results presented are still restricted (there is no 100% accuracy, the classification has some inconsistencies, it has few room labels available, robust algorithms must be considered, among other things). However, we understand that the significance of works of such a nature is meant to pave the way for other studies to evolve from this sort of contribution. Potentially, we believe that many domestic functions and relationships can be improved from this perspective, which, broadly, motivates this investigation and future developments.

Acknowledgements. This work is financed by national funds through FCT - Fundação para a Ciência e a Tecnologia, I.P., under the Strategic Project with the references UIDB/04008/2020 and UIDP/04008/2020. We also thank to UFMS, which has always been an important support institution and, finally, our friends and colleagues who collaborated with this research.

References

Ahmed, S., Liwicki, M., Weber, M., Dengel, A. (2012). Automatic Room Detection and Room Labeling from Architectural Floor Plans. In: 10th IAPR (pp. 339–343). https://doi.org/10.1109/DAS.2012.22

Aries, P., Duby, G. (2009). História da vida privada 4: Da Revolução Francesa à Primeira Guerra. Companhia de Bolso.

Brooks, R.A. (1991). Intelligence Without Reason. In J. Mylopoulos & R. Reiter (Eds.), Proceedings of the IJCAI'91. (pp. 569–595). Morgan Kaufmann Publishers Inc. https://dl.acm.org/doi/10.5555/1631171.1631258

Chaillou, S. (2019). AI+ Architecture: Towards a new approach, Harvard GSD.

Coelho, H. (1996). Sonho e Razão: Ao lado do artificial. Círculo de Leitores.

Furtado, G., Moreira, I. (2001). Cartografias da Domesticidade. Jornal dos Arquitectos. N. 203, 96–103.

Imovirtual. (n.d.), Imovirtual website. Retrieved September 03, 2022, from https://www.imovirtual.com/arrendar/apartamento/

Krueger, T. (2006). Metadermis: Like a second skin. In: Architecture: Machine and body. pp. 87–98. Faculdade de Arquitectura da Universidade do Porto.

Marcus, G. (2020). The Next Decade in AI: Four steps towards robust artificial intelligence. ArXiv200206177 Cs.

Miranda, J.B.: The “House of the Future” as an experience: problems and perspectives. In: Architecture: Machine and body. pp. 99–107. Faculdade de Arquitectura da Universidade do Porto.

MIT App Inventor. (n.d.). MIT App Inventor Plataform. Retrieved May 23, 2023, from http://ai2.appinventor.mit.edu/?locale=en#6127387376222208

Oliveira, A.L. (2017). The Digital Mind: How science is redefining humanity. The MIT Press.

Pedro, J.B.O. (2000). Definição e Avaliação da Qualidade Arquitectónica Habitacional, FAUP.

Pereira, S.M. (2004). Pressupostos Ideológicos da Casa Actual: O espaço como veículo do ideário moderno. CCT.

Personal Image Classifier (n.d.). PIC extension. Retrieved May 10, 2023, https://classifier.appinventor.mit.edu/oldpic/

Portas, N. (1969). Funções e Exigências de Áreas da Habitação. MOP Laboratório Nacional de Engenharia Civil.

Russell, S.J., Norvig, P. (2010). Artificial Intelligence: A modern approach. Prentice Hall.

Rybczynski, W. (2009). La Casa: Historia de una idea. Nerea.

Wang, Z., Liang, Q., Duarte, F., Zhang, F., Charron, L., Johnsen, L., Cai, B., Ratti, C. (2019). Quantifying Legibility of Indoor Spaces Using Deep Convolutional Neural Networks: Case studies in train stations. Environ. 160. https://doi.org/10.1016/j.buildenv.2019.04.035

Welling, M. (2015). Are ML and Statistics Complementary? In: Data Science in the next 50 years. p. 3.

Wu, J., Xue, T., Lim, J.J., Tian, Y., Tenenbaum, J.B., Torralba, A., Freeman, W.T. (2018). 3D Interpreter Networks for Viewer-Centered Wireframe Modeling. Int. J. Comput. 126, 1009–1026. https://doi.org/10.1007/s11263-018-1074-6