Wall polygon retrieval from architectural floor plan images using vectorization and deep learning methods

Pablo N. Pizarro
pablo@ppizarror.com
Department of Computer Sciences, University of Chile
Santiago, Chile

ABSTRACT

In this paper, we propose a model that retrieves the wall polygons from architectural floor plan images of Chilean multi-unit residential buildings. The model, assembled from the combination of wall segmentation and vectorization algorithms developed from comparing the state-of-the-art approaches, solves the wall retrieval for more detailed and complex plans than related work. Data preprocessing, train, and evaluation procedures will be studied for improving results. Furthermore, a novel dataset comprising 954 floor plans is proposed, considering the image and polygon data for each basement, first, and typical floor from 165 different buildings.

CCS CONCEPTS

• **Computing methodologies → Image segmentation**; **Scene understanding**.

KEYWORDS

wall retrieval, architectural floor plan, segmentation, image vectorization, polygon retrieval, raster-to-vector

ACM Reference Format:
Pablo N. Pizarro. 2021. Wall polygon retrieval from architectural floor plan images using vectorization and deep learning methods. In *CC7910 – Investigación en Ciencia de la Computación.* ACM, New York, NY, USA, 7 pages.

1 INTRODUCTION

Architectural floor plans are documents that result from an iterative design, planning, and engineering process to define the layout, distribution, and usage of a structure, playing a crucial role while designing, understanding, or remodeling indoor spaces [36]. Plans are 2D complex drawing that conveys geometric and semantic information from a 3D scene [64]. Usually, they consider the walls, windows, furniture, dimension lines, grids, text, or icon styles that define room types, making analysis and information recovery a challenging and open task [36].

Although plans are designed and built using advanced vector software such as AutoCAD[1], these are frequently stored as raster format images in the application process [64], discarding semantic and geometric information as only human post-processing is considered (e.g., home buyers, renters, or engineers). Thus, recovering information from rasterized floor plan images is a long-standing open problem [36, 39], posing three fundamental challenges. First, there is no standard notation among architectural and engineering firms, where colors, line thickness, and symbols differ [39]. Second, the plan structure must satisfy high-level geometric and semantic constraints. Finally, this high-level model structure varies across examples (e.g., different houses have different numbers of bedrooms) [36].

From the data recovery problem, a critical task is to identify the walls because these objects define the building's main layout and convey essential information to detect other structural elements [10]. This information is also helpful in the whole spectrum of architecture, engineering, and construction, providing data for design, analysis, cost estimation, among others [53]. Traditionally, the problem has been solved using low-level image processing methods that exploit heuristics to locate wall notations in floor plans using shape recognition, text filtering, line scanning, and pixel classification [66]. However, relying on hand-crafted features is an insufficient approach, as it lacks generality to handle diverse conditions [67]. Extensive effort is required to choose proper low-level processing operations, tune parameters, and craft rules and grammar based on drawing styles or architectural regularity [65].

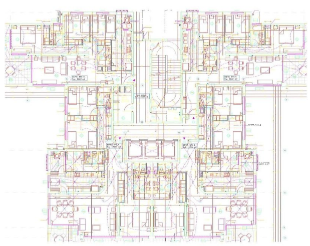
Recently, several deep learning methodologies have been proposed to identify the walls, mainly through the application of convolutional neural networks (CNNs) [23] and generative adversarial networks (GANs) [24] to segment and vectorize the structural objects, improving the performance with respect to a manual solution while keeping a general approach for handling different input styles [62]. Nevertheless, obtaining the wall's geometry from the segmented images, process known as vectorization, and the reconstruction of the wall topology, that is, the connection and neighborhood relationship, is a problem with great room for improvement. Studies such as the one conducted by Li *et al.* [35] make it possible to obtain a representation of the polygon contour by using an energy-refined approach. Macé *et al.* [39] detect walls and rooms from images using a coupling of the Hough Transform and image vectorization. Another study conducted by Feltes *et al.* [19] is capable of finding the object's corners in wall-line drawing images by filtering out unnecessary points without changing the overall structure.

[1]https://www.autodesk.com/products/autocad/

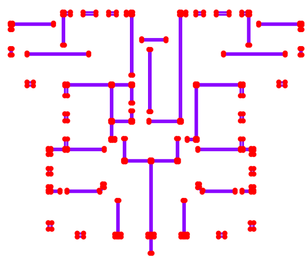
CC7910, Investigación en Ciencia de la Computación, 2021

Given the latest research in deep learning and image vectorization techniques, is it possible to automatically obtain the wall polygons from Chilean architectural floor plan images of residential buildings? That is, a model which considers as input an image of the floor plan (from a given drawing style of the national reality), and as output the wall polygon data.

In order to answer this question, the present work will implement and compare a selection of discriminative and generative-based deep learning wall segmentation models and evaluate raster-to-vector methods to obtain the wall polygons from the segmented images. Both segmentation and vectorization algorithms will be combined to assemble the proposed model, represented in Figure 1. The model input (Figure 1a) corresponds to an architectural floor plan image, which embeds the walls, doors, windows, grids, text, dimension lines, and furniture. Conversely, Figure 1b illustrates the output, composed of the wall's polygons with proper connectivity, length, thickness, and angles.



(a) Input – Architectural floor plan image from a Chilean residential building



(b) Output – Wall polygons, discretized and properly connected

Figure 1: Inputs and outputs of the proposed model.

2 RELATED WORK

Several previous studies have been conducted to retrieve the wall information from architectural floor plans, ranging from low-level image processing methods that recognize the objects considering manual features to recent deep learning procedures which train models to extract the walls in an automated fashion.

2.1 Low-level image processing

Among low-level image processing, there are studies where a parallel pair shape is searched to recognize walls, comparing the line style and orientation from a binarized image [8, 21, 38, 46, 61] or a set of prior database knowledge rules [55]. Several works applied pre-filtering methods for text and symbol removals, finding the walls through line recognition algorithms [1–4, 10, 14, 39, 45, 59] . Others performed manual binarization and classification from a specific group of rules [11, 12, 25]. Pixel proximity heuristics were also used to segment the walls [52], and other studies such as [5] approximated the wall polyline with a sequence of geometric primitives. Furthermore, graph-based models, which create a graph representation of the walls, have also been employed [9, 16].

Geometrically-based methodologies have also been proposed to retrieve the wall polygon from architectural floor plan images; for instance, Li *et al.* [35] obtained a representation of the polygon contour by using a generic energy-refined approach. Feltes *et al.* [19] work is capable of finding the object's corners in wall-line drawing images by filtering out unnecessary points without changing the overall structure; also, wall-gap filling is possible while performing a heuristic criterion. Finally, Mewada *et al.* [42] proposed a framework based on the α-shape algorithm to extract room shapes from binarized images, calculating and classifying their properties using a linear regression model.

2.2 Deep learning-based methods

Several deep learning models have been developed to recognize and vectorize the walls from architectural floor plan images in recent years. Among these, the wall object segmentation is one of the main tasks, which can be formulated as a classification (semantic segmentation) or partition problem (instance segmentation). Semantic segmentation performs pixel-level labeling with a set of object categories for all image pixels; by contrast, instance segmentation extends the classification scope further by detecting and delineating each object of interest in the image [43].

Within deep learning, image segmentation models can be discriminative or generative-based. Discriminative models learn the conditional probability distribution of the pixel classes (e.g., wall or background class), that is, the decision boundary. Conversely, generative models learn the joint probability distribution, that is, the distribution of the individual classes.

Among discriminative-based models, the semantic segmentation FCN [54], U-Net [51], DeepLab [6], and instance segmentation RCNN [22] have been used. FCNs, or Fully Convolutional Networks, are composed of two main sections: encoder (contraction) and decoder (expansion). The encoder section is used to capture the context from the image, comprised of several convolutional and max-pooling layers [23]; in opposition, the decoder section comprises many feature channels that enable precise localization through the transposed convolutions, propagating context information to higher resolution layers, giving the segmented output. Similar to FCNs, in U-Net the decoder also combines the feature and spatial information through a sequence of up-convolutions and concatenations with high-resolution features from the encoder, improving localization and reconstruction of the segmented output image; therefore, the expansive path is symmetric to the contracting part and yields an u-shaped architecture [51]. Likewise, DeepLab is a state-of-the-art semantic segmentation model which employs a pre-trained CNN to get encoded feature maps from the input and a decoder to reconstruct the segmented output image. Finally, the instance segmentation RCNN is a family of models which produce a set of bounding boxes for each object in the image, named regions of interests (ROI), where the position and category (e.g., wall) are inferred using neural networks.

Concerning the wall segmentation problem, Dodge *et al.* [15] was the first to propose an FCN model with different pixel-strides to

Table 1: Representative research in floor plan analysis concerning deep learning methods.

Title (year)	Model	Dataset (plans)	Parsing	Segmentation	Vectorization	Metrics
[15] (2017)	FCN, Faster-RCNN	R-FP (500), CVC-FP (122)	✓	✓	–	IoU, JI, ACC, *AP*
[36] (2017)	CNN	R2V (870) from LIFULL	✓	–	✓	*AP*, *RE*
[26] (2018)	U-Net	LIFULL (NA, *private*)	–	✓	✓	–
[63] (2018)	FCN	LIFULL (5000, *private*)	✓	✓	–	ACC
[64] (2018)	U-Net	EAIS (325, *private*)	–	✓	–	IoU, ACC
[66] (2019)	CNN, DeepLabV3+, PSPNet, RCF	R3D (232), R2V (815)	✓	✓	–	IoU, ACC
[58] (2020)	U-Net, Faster-RCNN	BTI (700, *private*), CubiCasa5K (5000)	–	✓	✓	IoU, *mAP*
[62] (2020)	Mask-RCNN	CVC-FP (122)	✓	✓	✓	*MS*
[30] (2020)	DeepLabV3+	EAIS (319, *private*), R-FP (500)	–	✓	✓	IoU
[67] (2020)	GAN	R3D (232), R2V (815)	✓	✓	–	ACC
[69] (2020)	FCN, DeepLabV3+	CubiCasa5K (540)	–	✓	–	IoU
[18] (2020)	U-Net, ANN	PFP (1554, private)	–	✓	–	<u>hD</u>, <u>P</u>
[32, 33] (2021)	GANs, Pix2Pix	EAIS (450, *private*), CVC-FP (122)	–	✓	✓	ACC
[17] (2021)	EdgeGAN	ZSCVFP (10800, *private*)	✓	–	✓	ACC

segment walls and Faster-RCNN [50] for object detection. Yamasaki *et al.* [63] also used an FCN model to segment the floor plan objects and a graph structure to assemble the apartment layout. Hanme *et al.* [26] segmented the walls employing U-Net. Similarly, Yang *et al.* [64] used U-Net along pixel deconvolutional layers (PixelDCL [20]) to avoid checkerboard artifacts, work extended by Surikov *et al.* [58] who detected the wall objects with the Faster-RCNN model and applied the Ramer-Douglas-Peucker [49] algorithm to simplify the polygons. Wu *et al.* [62] used Mask-RCNN [27] to segment the walls by finding a region proposal that represents its width, thickness, and location. Jang *et al.* [30] segmented the walls using DeepLabV3+ and later vectorized them using a node-edge graph. Finally, Zhu *et al.* [69] compared different training strategies to parse complex floor plans considering FCN-2s and DeepLabV3+, with VGG16 [56] as a backbone.

CNN-based models have also been developed to obtain the wall objects without segmenting the image. Liu *et al.* [36] used a CNN model to detect the wall edges (ResNet-152 [28]) and applied integer programming on the extracted corners to construct the vector data that represented the walls. Zeng *et al.* [66] used a shared VGG network for feature extraction and a Conv+ReLU subnet to find the wall edges, comparing the results against the RCF edge detection network [37], DeepLabV3+, and PSPNet (Pyramid scene parsing network [68]) segmentation networks. On the other hand, not only walls are the target. Egiazarian *et al.* [18] obtained the line primitives from generic technical plans, using U-Net for image pre-processing, later splitted into patches to estimate the line primitives with a feed-forward artificial neural network (ANN), each patch encoded with ResNet-based feature estimator and decoded using Transformer blocks [60].

Among generative-based approaches, Kim *et al.* [32, 33] studied style transfer to transform plan styles to a unified format using the image-to-image conditional generative adversarial network Pix2Pix [29], and a GAN [24] to vectorize the floor plan, generating candidates for the wall junctions which are later connected to assemble the vector representation through integer programming. Zhang *et al.* [67] created manually-defined direction-aware additive kernels to optimize the recognition and generation of walls with a GAN, comparing the implementation against [66]. Finally, Dong *et al.* [17] developed an edge extraction GAN (EdgeGAN), based on the image-to-image translation Pix2Pix model, to detect the wall primitives.

Table 1 details the deep learning related work within floor plan analysis, considering the dataset used, the tasks performed such as floor parsing (subdivision and categorization of rooms), object segmentation and vectorization, and finally, the metrics used. The segmentation results were evaluated in terms of the intersection over union (IoU) [54], pixel accuracy (ACC), and the Jaccard Index (JI) proposed by de Las Heras *et al.* [9]. By contrast, works that detected objects (e.g., walls, doors, windows) used the mean average precision (*mAP*), the average precision (*AP*), recall (*RE*), and the match score (*MS*) [9] metrics. Finally, to measure the difference in skeleton structures of two vector images, Hausdorff distance (<u>dH</u>) and the number of primitives (<u>P</u>) were used. Table 2 details the dataset properties.

Although previous work considered segmentation to retrieve the structural objects from architectural floor plan images, they do not aim to obtain a high-level representation of the wall polygons; usually, the segmentation was an intermediate step to assemble the floor structure (parsing). Also, the dataset's floor plans belong to single apartments or houses, as illustrated in Figure 4, which are different from the dataset to be used in this work, composed of multi-unit floor plans from Chilean residential buildings (including all apartments, halls, or perimeter walls). Moreover, in Chile, the floor layout is usually composed with a non-uniform distribution of many walls with complex cross-sections [34, 40, 41] compared to other countries such as the U.S., where the architectural floor plan usually consists of a few rectangular walls [44].

For the reasons mentioned above, ideas can be gathered from previous work to segment and generate the vector representation of the wall polygons, finding the consideration and thresholds needed to process large and complex plans.

3 PROBLEM STATEMENT

In this investigation, the problem to be solved is obtaining the wall polygons from a rasterized architectural floor plan image in an automatic procedure, without the need for human input, in such a way it can handle several input styles and the resulting polygon adequately represents the semantics underlying the plan drawing. In previous work, researchers have used low-level image processing methods that exploited manual heuristics to find the objects; however, as these methods lack generality to handle diverse conditions imposed from the highly variable input plans, a deep learning approach will be used.

We aim to build a model that uses a deep learning method to segment the wall objects from architectural floor plan images and a vectorization algorithm to retrieve the polygons (Figure 1). The dataset to train and test the proposed model comprises 165 Chilean residential building projects [47, 48] designed by 53 different offices, yielding 954 high-quality plan images for the basement, first floor, and typical floor of the buildings. The images are stored in PNG format of 9000 pixels wide, obtained from each AutoCAD source file. For each plan, the walls were labeled and stored as a rectangular-discretized polygon in a graph structure; thus, the connection topology is retained. Figure 2 illustrate a selection of three different floor plans from the dataset, considering the wall polygon graph and the image in different drawing styles; from these examples, it is possible to observe that the walls have different thicknesses and line colors, with multiple orientations, complex connection layouts, and discontinuities due to other elements such as notations or grids. The assembly process of the floor plan dataset is detailed in Figure 3. First, the image of each plan was collected from the digital drawings (Figure 3a), alongside the wall contour polygon (Figure 3b); each polygon was processed, discretizing the complex cross-section into a connected graph, where vertices belong to wall joints, and the edges to the wall segments (Figure 3c). Finally, the polygons were located on each plan's correct position (Figure 3d).

To solve the problem, first, a comparison between a selection of deep learning-based wall segmentation models from related work will be performed, using both discriminative and generative-based approaches, to find the one that obtains the best results in terms of the intersection over union (IoU) [54], which is a standard evaluation metric widely used within floor plan analysis (Table 1). Then, a vectorization algorithm will be studied to obtain the wall polygon shape from the segmented images [35, 39] while applying several heuristics for filtering and simplifying the polygons [19]. Finally, the segmentation and vectorization algorithms will be combined to assemble the proposed model, comparing the results between the output and the real wall polygons from the Chilean dataset.

3.1 Research question

From the proposed problem, the following questions arose:

- Which state-of-the-art deep learning model, from discriminative to generative-based, has better performance for segmenting the wall objects considering a dataset of Chilean architectural floor plans?
- Is there room for improvement for these deep learning segmentation models since new methods have been discovered to improve training and input pre-processing in recent years?
- Is there a better metric than IoU to characterize wall results from architectural floor plan analysis?
- How do the classical raster-to-vector algorithms perform while processing the segmented wall images? Could previous knowledge of the polygons be used to improve results?
- Can the proposed model, assembled from the selected segmentation and vectorization algorithms, obtain the wall polygons directly from the floor plan image, with equal or better results than the state-of-the-art approaches?

3.2 Hypothesis

The proposed model, assembled with the best deep-learning and raster-to-vector segmentation models that emerge from the comparison of related work, will allow obtaining the wall polygons from Chilean architectural plan images, with results similar to the state-of-the-art approaches that solve the problem for less complex and detailed plans.

3.3 General objective

In this investigation, we will develop a model (Figure 1) that considers as input an architectural floor plan image (Figure 1a) and returns a highly accurate and refined wall polygon as output (Figure 1b). The model will be assembled from a couple of deep learning-based wall segmentation and raster-to-vector algorithms that emerged from comparing the state-of-the-art floor plan analysis works.

3.4 Specific objectives

O1) Compare discriminative and generative-based deep learning models for wall segmentation, which have been proved to have a better performance against low-level image processing methods that rely on manual heuristics.

O2) Process the Chilean architectural floor plan dataset (954 floor plans from 165 different residential buildings) to find the best data structure to handle the wall segmentation and vectorization considering semantic and memory restraints.

O3) Evaluate different methodologies to improve the segmentation model results regarding the intersection over union (IoU) between the ground-truth and the output images. In addition, other metrics will be studied to determine if it is possible to characterize the wall polygon results better.

O4) Implement and evaluate a vectorization algorithm that obtains the wall polygon from the segmented images.

O5) Implement the proposed model that automatically obtains the wall polygon from the floor plan image, using the developed segmentation and vectorization algorithms, comparing the results to the ground-truth for each building from the Chilean dataset.

3.5 Methodology

3.5.1 Research. The first step of the research is to explore the state-of-the-art methods that solve obtaining the wall objects from an architectural floor plan image, considering both classical low-level image processing and the latest deep learning approaches.

3.5.2 Wall polygon retrieval model. The model development considers three steps. First, a comparison between different state-of-the-art wall segmentation deep learning algorithms will be made (**O1**), which has been proved to better perform against low-level image processing methods that rely on manual heuristics [62]. The comparison will include discriminative-based models such as U-Net [26, 58, 64] and DeepLabV3+ [30, 66, 69], and the generative-based Pix2Pix [17, 32]. Models will be trained and tested using the Chilean architectural floor plan dataset, which will be processed to handle the segmentation considering semantic and memory restraints (**O2**) or studying data-augmentation procedures from related work [18, 26, 30, 32, 33, 36, 57, 58, 62]. Studies such as [69] will also be applied to enhance the current deep learning models (**O3**) in terms of the intersection over union (IoU), a standard metric used within floor plan analysis research (Table 1).

As a second step, a vectorization algorithm will be implemented and evaluated to find the wall polygon from the segmented images (**O4**). This vectorization can be based on the Hough Transform [39], or energy-based polygon refined method [35]. The output polygon will be post-processed using geometrical heuristics [19] and prior knowledge of the wall polygon domain to improve the result's quality.

Finally, the deep learning model for segmenting the image and the vectorization algorithm for obtaining the wall polygons will be combined, assembling the proposed model, which will be compared to the groud-truth included for each floor plan in the dataset (**O5**).

3.5.3 Experimentation. Experiments will be performed throughout steps to check the deep learning segmentation models and the vectorization results. The aim is to compare the output polygons alongside the real solution (ground-truth), which is already included in the whole dataset of 954 Chilean residential building floor plans.

3.5.4 Technologies. The model will be implemented in Python, using Keras-TensorFlow[2] as the machine learning backend. For image processing, OpenCV[3] is considered. The front-end application will comprise a web interface for uploading the floor plan, execute the model considering its input parameters, and visualize the results.

3.6 Expected results

- A short literature review that includes both low-level image processing methods and the latest deep learning approaches to retrieve wall objects from architectural floor plan images.
- A novel dataset comprising 954 Chilean architectural floor plans of residential buildings to be used for wall retrieval research.
- A comparison among several state-of-the-art deep learning discriminative and generative-based wall segmentation models for Chilean floor plans.
- A vectorization algorithm that obtains the wall polygons from a segmented image output.
- A model that automatically obtains the wall polygon from Chilean architectural floor plan images of residential buildings, comprising a couple of the segmentation and vectorization algorithms.

[2]https://www.tensorflow.org/
[3]https://opencv.org/

REFERENCES

[1] C. Ah-Soon and K. Tombre. 1997. Variations on the analysis of architectural drawings. In *Proceedings of the Fourth International Conference on Document Analysis and Recognition, ICDAR*, Vol. 1. IEEE Comput. Soc, 347–351.
[2] Sheraz Ahmed, Marcus Liwicki, Markus Weber, and Andreas Dengel. 2011. Improved Automatic Analysis of Architectural Floor Plans. In *2011 International Conference on Document Analysis and Recognition*. IEEE, 864–869.
[3] Sheraz Ahmed, Marcus Liwicki, Markus Weber, and Andreas Dengel. 2012. Automatic Room Detection and Room Labeling from Architectural Floor Plans. In *2012 10th IAPR International Workshop on Document Analysis Systems*. IEEE, 339–343.
[4] Sheraz Ahmed, Markus Weber, Marcus Liwicki, Christoph Langenhan, Andreas Dengel, and Frank Petzold. 2014. Automatic analysis and sketch-based retrieval of architectural floor plans. *Pattern Recognition Letters* 35 (jan 2014), 91–100.
[5] Eugene Bodansky and Alexander Gribov. 2006. Approximation of a Polyline with a Sequence of Geometric Primitives. In *Image Analysis and Recognition. ICIAR 2006. Lecture Notes in Computer Science, vol 4142*. Springer, Berlin, Heidelberg, 468–478.
[6] Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, and Hartwig Adam. 2018. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. In *Computer Vision – ECCV 2018. ECCV 2018. Lecture Notes in Computer Science, vol 11211*. Springer, Cham, 833–851.
[7] Chenxi Liu, Alexander G. Schwing, Kaustav Kundu, Raquel Urtasun, and Sanja Fidler. 2015. Rent3D: Floor-plan priors for monocular layout estimation. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 3413–3421.
[8] Paramita De. 2019. Vectorization of Architectural Floor Plans. In *2019 Twelfth International Conference on Contemporary Computing (IC3)*. IEEE, 1–5.
[9] Lluís-Pere de las Heras, Sheraz Ahmed, Marcus Liwicki, Ernest Valveny, and Gemma Sánchez. 2014. Statistical segmentation and structural recognition for floor plan interpretation. *International Journal on Document Analysis and Recognition (IJDAR)* 17, 3 (sep 2014), 221–237.
[10] Lluis-Pere de las Heras, David Fernandez, Ernest Valveny, Josep Llados, and Gemma Sanchez. 2013. Unsupervised Wall Detector in Architectural Floor Plans. In *2013 12th International Conference on Document Analysis and Recognition*. IEEE, 1245–1249.
[11] Lluis-Pere de las Heras, Joan Mas, Gemma Sanchez, and Ernest Valveny. 2011. Wall Patch-Based Segmentation in Architectural Floorplans. In *2011 International Conference on Document Analysis and Recognition*. IEEE, 1270–1274.
[12] Lluís-Pere de las Heras, Joan Mas, Gemma Sánchez, and Ernest Valveny. 2013. Notation-Invariant Patch-Based Wall Detector in Architectural Floor Plans. In *Graphics Recognition. New Trends and Challenges. GREC 2011. Lecture Notes in Computer Science, vol 7423*. Springer, Berlin, Heidelberg, 79–88.
[13] Lluís-Pere de las Heras, Oriol Ramos Terrades, Sergi Robles, and Gemma Sánchez. 2015. CVC-FP and SGT: a new database for structural floor plan analysis and its groundtruthing tool. *International Journal on Document Analysis and Recognition (IJDAR)* 18, 1 (mar 2015), 15–30.
[14] Lluís-Pere de las Heras, Ernest Valveny, and Gemma Sánchez. 2014. Unsupervised and Notation-Independent Wall Segmentation in Floor Plans Using a Combination of Statistical and Structural Strategies. In *Graphics Recognition. Current Trends and Challenges. GREC 2013. Lecture Notes in Computer Science, vol 8746*. Springer, Berlin, Heidelberg, 109–121.
[15] Samuel Dodge, Jiu Xu, and Bjorn Stenger. 2017. Parsing floor plan images. In *2017 Fifteenth IAPR International Conference on Machine Vision Applications (MVA)*. IEEE, 358–361.
[16] B. Domínguez, Á.L. García, and F.R. Feito. 2012. Semiautomatic detection of floor topology from CAD architectural drawings. *Computer-Aided Design* 44, 5 (may 2012), 367–378.
[17] Shuai Dong, Wei Wang, Wensheng Li, and Kun Zou. 2021. Vectorization of Floor Plans Based on EdgeGAN. *Information* 12, 5 (may 2021), 206.
[18] Vage Egiazarian, Oleg Voynov, Alexey Artemov, Denis Volkhonskiy, Aleksandr Safin, Maria Taktasheva, Denis Zorin, and Evgeny Burnaev. 2020. Deep Vectorization of Technical Drawings. In *Computer Vision – ECCV 2020. ECCV 2020. Lecture Notes in Computer Science, vol 12358*. 582–598.

[19] Max Feltes, Sheraz Ahmed, Andreas Dengel, and Marcus Liwicki. 2014. Improved Contour-Based Corner Detection for Architectural Floor Plans. In *Graphics Recognition. Current Trends and Challenges. GREC 2013. Lecture Notes in Computer Science, vol 8746*. Springer, Berlin, Heidelberg, 191–203.

[20] Hongyang Gao, Hao Yuan, Zhengyang Wang, and Shuiwang Ji. 2017. Pixel Deconvolutional Networks. *arXiv* (may 2017). arXiv:1705.06820 https://arxiv.org/abs/1705.06820

[21] Lucile Gimenez, Sylvain Robert, Frédéric Suard, and Khaldoun Zreik. 2016. Automatic reconstruction of 3D building models from scanned 2D floor plans. *Automation in Construction* 63 (mar 2016), 48–56.

[22] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. 2014. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 580–587.

[23] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. 2016. Deep learning. *MIT Press* (2016). https://www.deeplearningbook.org/

[24] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative adversarial nets. In *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2*. MIT Press, Montreal, Canada, 2672–2680.

[25] Guanghui Pan, Jia He, and Rui Fang. 2017. Automatic floor plan detection and recognition. In *2017 2nd International Conference on Image, Vision and Computing (ICIVC)*. IEEE, 201–205.

[26] Jang Hanme, Yang Jong Hyeon, and Kiyun Yu. 2018. Automatic Wall Detection and Building Topology and Property of 2D Floor Plan. In *10th International Conference on Geographic Information Science (GIScience 2018)*, Winter Stephan, Griffin Amy, and Sester Monika (Eds.). Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 33:1–33:5.

[27] Kaiming He, Georgia Gkioxari, Piotr Dollar, and Ross Girshick. 2017. Mask R-CNN. In *2017 IEEE International Conference on Computer Vision (ICCV)*. IEEE, 2980–2988.

[28] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep Residual Learning for Image Recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 770–778.

[29] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. 2017. Image-to-Image Translation with Conditional Adversarial Networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 5967–5976.

[30] Hanme Jang, Kiyun Yu, and JongHyeon Yang. 2020. Indoor Reconstruction from Floorplan Images with a Deep Learning Approach. *ISPRS International Journal of Geo-Information* 9, 2 (jan 2020), 65.

[31] Ahti Kalervo, Juha Ylioinas, Markus Häikiö, Antti Karhu, and Juho Kannala. 2019. CubiCasa5K: A Dataset and an Improved Multi-task Model for Floorplan Image Analysis. In *Image Analysis. SCIA 2019. Lecture Notes in Computer Science, vol 11482*. Springer, Cham, 28–40.

[32] Seongyong Kim, Seula Park, Hyunjung Kim, and Kiyun Yu. 2021. Deep Floor Plan Analysis for Complicated Drawings Based on Style Transfer. *Journal of Computing in Civil Engineering* 35, 2 (mar 2021), 04020066.

[33] Seongyong Kim, Seula Park, and Kiyun Yu. 2018. Application of Style Transfer in the Vectorization Process of Floorplans (Short Paper). *10th International Conference on Geographic Information Science (GIScience 2018)* (2018).

[34] René Lagos, Mario Lafontaine, Patricio Bonelli, Rubén Boroschek, Tomas Guendelman, Leonardo M. Massone, Rodolfo Saragoni, Fabián Rojas, and Fernando Yañez. 2021. The quest for resilience: The Chilean practice of seismic design for reinforced concrete buildings. *Earthquake Spectra* 37, 1 (2021), 26–45.

[35] Muxingzi Li, Florent Lafarge, and Renaud Marlet. 2020. Approximating shapes in images with low-complexity polygons. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 8630–8638.

[36] Chen Liu, Jiajun Wu, Pushmeet Kohli, and Yasutaka Furukawa. 2017. Raster-to-Vector: Revisiting Floorplan Transformation. In *2017 IEEE International Conference on Computer Vision (ICCV)*. IEEE, 2195–2203.

[37] Yun Liu, Ming-Ming Cheng, Xiaowei Hu, Kai Wang, and Xiang Bai. 2017. Richer Convolutional Features for Edge Detection. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 5872–5881.

[38] Tong Lu, Huafei Yang, Ruoyu Yang, and Shijie Cai. 2007. Automatic analysis and integration of architectural drawings. *International Journal on Document Analysis and Recognition (IJDAR)* 9, 1 (feb 2007), 31–47.

[39] Sébastien Macé, Hervé Locteau, Ernest Valveny, and Salvatore Tabbone. 2010. A system to detect rooms in architectural floor plan images. In *Proceedings of the 8th IAPR International Workshop on Document Analysis Systems - DAS '10*. ACM Press, New York, USA, 167–174.

[40] Leonardo M. Massone. 2013. Fundamental principles of the reinforced concrete design code changes in Chile following the Mw 8.8 earthquake in 2010. *Engineering Structures* 56 (nov 2013), 1335–1345.

[41] Leonardo M. Massone, Patricio Bonelli, René Lagos, Carl Lüders, Jack Moehle, and John W. Wallace. 2012. Seismic Design and Construction Practices for R.C. Structural Wall Buildings. *Earthquake Spectra* 28, SUPPL.1 (jun 2012), 245–256.

[42] Hiren K. Mewada, Amit V. Patel, Jitendra Chaudhari, Keyur Mahant, and Alpesh Vala. 2020. Automatic room information retrieval and classification from floor plan using linear regression model. *International Journal on Document Analysis and Recognition (IJDAR)* 23, 4 (dec 2020), 253–266.

[43] Shervin Minaee, Yuri Y. Boykov, Fatih Porikli, Antonio J Plaza, Nasser Kehtarnavaz, and Demetri Terzopoulos. 2021. Image Segmentation Using Deep Learning: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2021), 1–1.

[44] National Institute of Standards and Technology (NIST). 2012. Comparison of U.S. and Chilean Building Code Requirements and Seismic Design Practice 1985–2010. *NIST GCR 12-917-18* (2012). https://nehrp.gov/pdf/nistgcr12-917-18.pdf

[45] Siu-hang Or, Kin-Hong Wong, Ying-kin Yu, and Michael Ming-yuan Chang. 2005. Highly Automatic Approach to Architectural Floorplan Image Understanding & Model Generation. *Proc. of Vision, Modeling, and Visualization 2005 (VMV - 2005)* (2005), 25–32. https://www.cse.cuhk.edu.hk/~shor/paper/vmv05.pdf

[46] Jaehwa Park and Young-Bin Kwon. 2004. Main Wall Recognition of Architectural Drawings Using Dimension Extension Line. In *Graphics Recognition. Recent Advances and Perspectives. GREC 2003. Lecture Notes in Computer Science, vol 3088*. Springer, Berlin, Heidelberg, 116–127.

[47] Pablo N. Pizarro and Leonardo M. Massone. 2021. Structural design of reinforced concrete buildings based on deep neural networks. *Engineering Structures* 241 (aug 2021), 112377.

[48] Pablo Nicolás Pizarro, Leonardo M. Massone, Fabián R. Rojas, and Rafael O. Ruiz. 2021. Use of convolutional networks in the conceptual structural design of shear wall buildings layout. *Engineering Structures* 239 (jul 2021), 112311.

[49] Urs Ramer. 1972. An iterative procedure for the polygonal approximation of plane curves. *Computer Graphics and Image Processing* 1, 3 (nov 1972), 244–256.

[50] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. 2017. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39, 6 (jun 2017), 1137–1149.

[51] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Lecture Notes in Computer Science* 9351 (2015), 234–241.

[52] K. Ryall, S. Shieber, J. Marks, and M. Mazer. 1995. Semi-automatic delineation of regions in floor plans. In *Proceedings of 3rd International Conference on Document Analysis and Recognition*, Vol. 2. IEEE Comput. Soc. Press, 964–969.

[53] Divya Sharma, Nitin Gupta, Chiranjoy Chattopadhyay, and Sameep Mehta. 2017. DANIEL: A Deep Architecture for Automatic Analysis and Retrieval of Building Floor Plans. In *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*. IEEE, 420–425.

[54] Evan Shelhamer, Jonathan Long, and Trevor Darrell. 2017. Fully Convolutional Networks for Semantic Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39, 4 (apr 2017), 640–651.

[55] Akio Shio and Yasuhiro Aoki. 2000. Sketch Plan: A prototype system for interpreting hand-sketched floor plans. *Systems and Computers in Japan* 31, 6 (jun 2000), 10–18.

[56] Karen Simonyan and Andrew Zisserman. 2014. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* (sep 2014). arXiv:1409.1556 https://arxiv.org/abs/1409.1556

[57] J. Song and K Yu. 2021. Framework for Indoor Elements Classification via Inductive Learning on Floor Plan Graphs. *ISPRS Int. J. Geo-Inf* 10, 97 (2021).

[58] Ilya Y. Surikov, Mikhail A. Nakhatovich, Sergey Y. Belyaev, and Daniil A. Savchuk. 2020. Floor Plan Recognition and Vectorization Using Combination UNet, Faster-RCNN, Statistical Component Analysis and Ramer-Douglas-Peucker. In *Computing Science, Communication and Security. COMS2 2020. Communications in Computer and Information Science, vol 1235*. Springer, Singapore, 16–28.

[59] Rui Tang, Yuhan Wang, Darren Cosker, and Wenbin Li. 2017. Automatic structural scene digitalization. *PLOS ONE* 12, 11 (nov 2017), e0187513.

[60] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention Is All You Need. *arXiv* (jun 2017). arXiv:1706.03762 http://arxiv.org/abs/1706.03762

[61] Qiao Wen and Rui-Guang Zhu. 2020. Automatic Generation of 3D Building Models Based on Line Segment Vectorization. *Mathematical Problems in Engineering* 2020 (oct 2020), 1–16.

[62] Yijie Wu, Jianga Shang, Pan Chen, Sisi Zlatanova, Xuke Hu, and Zhiyong Zhou. 2020. Indoor mapping and modeling by parsing floor plan images. *International Journal of Geographical Information Science* 35, 6 (jun 2020), 1205–1231.

[63] Toshihiko Yamasaki, Jin Zhang, and Yuki Takada. 2018. Apartment Structure Estimation Using Fully Convolutional Networks and Graph Model. In *Proceedings of the 2018 ACM Workshop on Multimedia for Real Estate Tech*. ACM, New York, NY, USA, 1–6.

[64] JongHyeon Yang, Hanme Jang, JiYeup Kim, and JungOk Kim. 2018. Semantic Segmentation in Architectural Floor Plans for Detecting Walls and Doors. In *2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*. IEEE, 1–9.

[65] Xuetao Yin, Peter Wonka, and Anshuman Razdan. 2009. Generating 3D Building Models from Architectural Drawings: A Survey. *IEEE Computer Graphics and Applications* 29, 1 (jan 2009), 20–30.

[66] Zhiliang Zeng, Xianzhi Li, Ying Kin Yu, and Chi-Wing Fu. 2019. Deep Floor Plan Recognition Using a Multi-Task Network with Room-Boundary-Guided Attention. *2019 IEEE/CVF International Conference on Computer Vision (ICCV)* (aug 2019).

[67] Yuli Zhang, Yeyang He, Shaowen Zhu, and Xinhan Di. 2020. The Direction-Aware, Learnable, Additive Kernels and the Adversarial Network for Deep Floor Plan Recognition. *arXiv* (jan 2020). arXiv:2001.11194 https://arxiv.org/abs/2001.11194

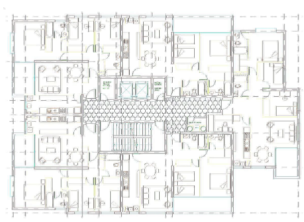
[68] Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, and Jiaya Jia. 2017. Pyramid Scene Parsing Network. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 6230–6239.

[69] Ruiyun Zhu, Jingcheng Shen, Xiangtian Deng, Marcus Walldén, and Fumihiko Ino. 2020. Training Strategies for CNN-based Models to Parse Complex Floor Plans. In *Proceedings of the 2020 9th International Conference on Software and Computer Applications*. ACM, New York, NY, USA, 11–16.

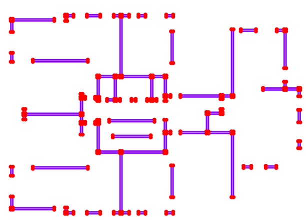
A APPENDICES

Table 2: Datasets used by prior floor plan analysis research.

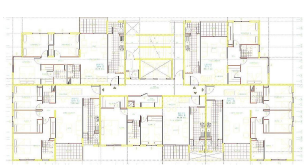
Dataset (year)	Annotation (quantity)
R3D – Rent3D, [7] (2015)	Walls, openings, room types (232)
CVC-FP, [13] (2015)	Walls, doors, windows, rooms without type (122)
R-FP, [15] (2017)	Walls (500)
R2V, [36] (2017)	Walls, openings, room types, icon types (815)
CubiCasa5K, [31] (2019)	80 object categories such as doors, windows and walls (5000)
EAIS, [30] (2020)	Walls, doors (319-400)
LIFULL HOME'S dataset (–)	None (–)



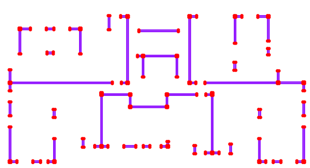
(a) Ex. 1 – Image



(b) Ex. 1 – Wall polygon

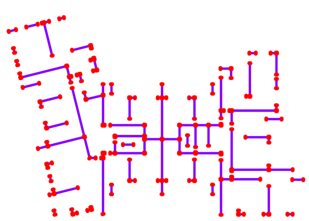


(c) Ex. 2 – Image



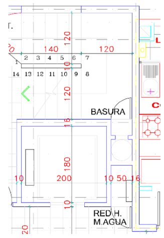
(d) Ex. 2 – Wall polygon

(e) Ex. 3 – Image

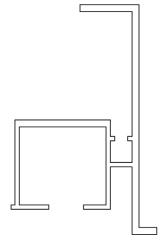


(f) Ex. 3 – Wall polygon

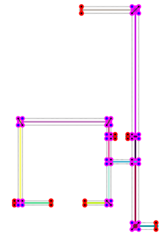
Figure 2: Example of three different architectural floor plans, considering the image and the wall polygons.



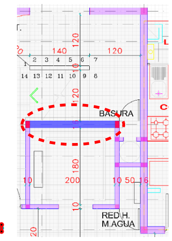
(a) Retrieval of the floor plan image



(b) Polygon of the wall contour



(c) Graph-based wall discretization

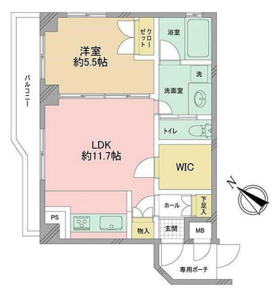


(d) Modeling of the wall

Figure 3: Example of the walls' assembly process.

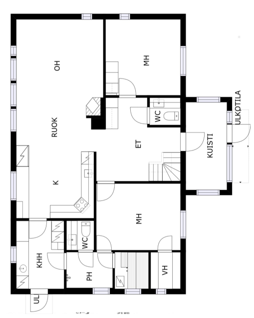
R3D – Rent3D

CVC-FP

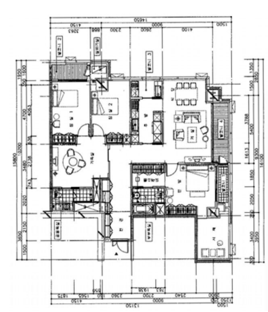


R-FP

R2V / LIFULL



CubiCasa5K



EAIS

Figure 4: Floor plan image examples from related work' datasets.