SpaceXYZ – Analyzing Floor Plan Images

Group22: Qinjin Jia, Rongyu Wang, Robert LeBourdais, Chi Yu Yeh

CS542 Fall 2018, Graduate School of Arts & Sciences, Boston University

Abstract

Architectural floor plans are scaled drawings of building layouts. Floor plan analysis has been an active research topic and has a number of applications (e.g. similarity search).
The aim of the project is to parse such floor plans using wall segmentation, object detection and optical character recognition (OCR). More specifically, we aim to recognize the number of main rooms, the type of each room and the surface area of the room from floor plans with different resolutions.

Introduction

Overall:

The project involves four techniques: wall segmentation, object detection, optical character recognition (OCR) and Merging and Splitting.

- We constructed a fully convolutional networks (FCN) for wall segmentation, so we can extract the wall from the floor plan.
- We trained a Faster R-CNN for object detection to detect the doors and windows, so we can get the locations of these elements.
- OCR is used for obtain the the room and size information from the floor plan.
- Finally, we merged wall mask, door/window mask together and then split into room units and calculated sizes.

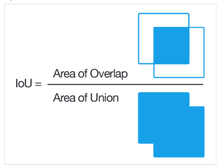
Dataset:

CVC-FP:

The collection consists of 122 scanned floor plan documents divided in 4 different subsets regarding their origin and style. It contains documents of different qualities, resolutions, and modeling styles, which is suitable to test the robustness of the analysis techniques.
The dataset is fully groundtruthed in .svg format for the structural symbols: rooms, walls, doors, windows, parking doors, and room separations

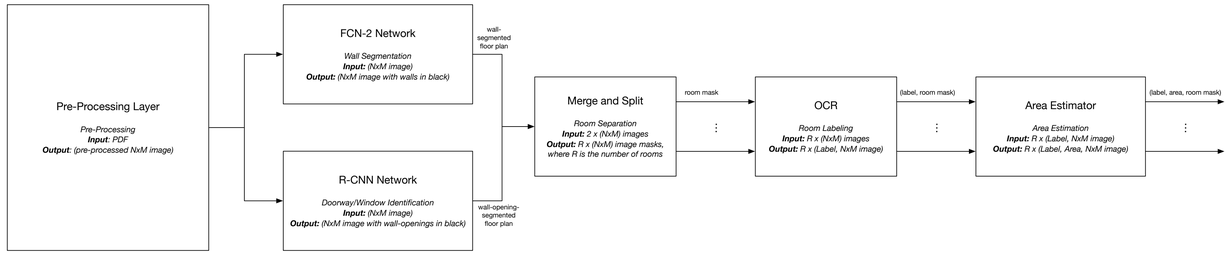
Evaluation Criteria:

Intersection Over Union(Jaccard index): The Jaccard coefficient measures similarity between finite sample sets, and is defined as the size of the intersection divided by the size of the union of the sample sets.



For **wall segmentation**, area of overlap of ground truth wall area and predicted wall area / area of union of ground truth wall area and predicted wall area.
For **door/window detection**, area of overlap of ground truth bounding box and predicted bounding box / area of union of ground truth bounding box and predicted bounding box.

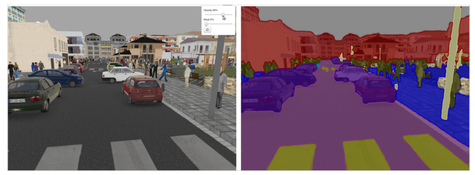
System Diagram

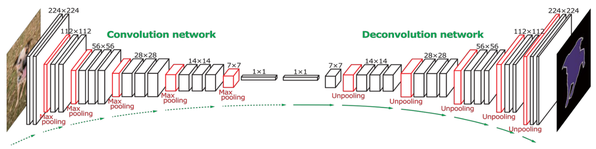


Methodology

Fully Convolutional Network (FCN):

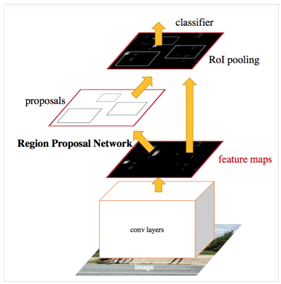
Semantic segmentation example:





- "Fully convolutional" networks are very deep model that take input of arbitrary size and produce correspondingly-sized output with efficient inference and learning.
- FCN can efficiently learn to make dense predictions for per-pixel tasks like semantic segmentation.
- However, as the group only has ~100 training samples, the pretrained VGG16 model is employed to initialize the FCN32 model and the training points will be used for fine tuning the VGG16 model.

Faster Region Convolutional Neural Network (Faster RCNN):



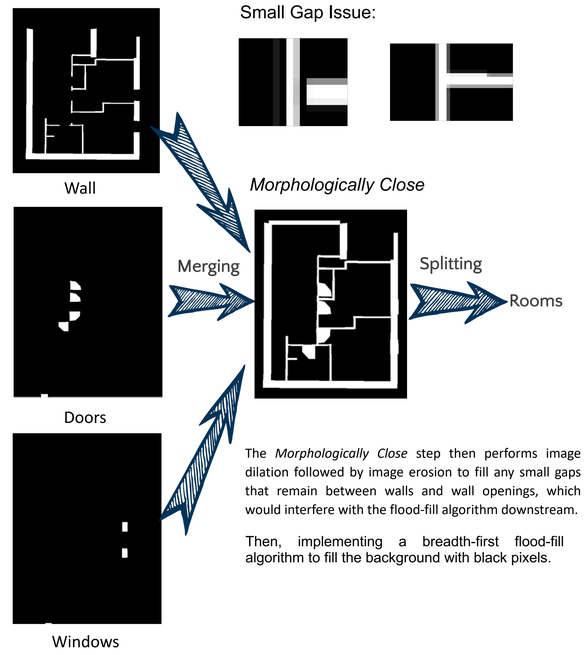
Faster RCNN is State-of-the-art object detection networks .
The image is provided as an input to a convolutional network which provides a convolutional feature map. Instead of using selective search algorithm on the feature map to identify the region proposals, a separate network is used to predict the region proposals. The predicted region proposals are then reshaped using a RoI pooling layer which is then used to classify the image within the proposed region and predict the offset values for the bounding boxes.

Optical Character Recognition (OCR):

Google Cloud

Extracting room names and sizes from the floor plan images with OCR API provided by Google Cloud.

Merging and Splitting:



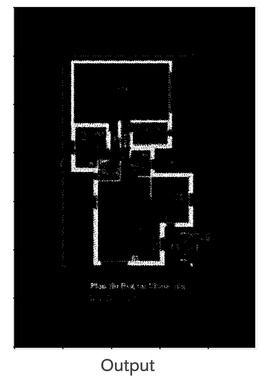
The *Morphologically Close* step then performs image dilation followed by image erosion to fill any small gaps that remain between walls and wall openings, which would interfere with the flood-fill algorithm downstream.

Then, implementing a breadth-first flood-fill algorithm to fill the background with black pixels.

Results

Wall segmentation with FCN:

Input



Output

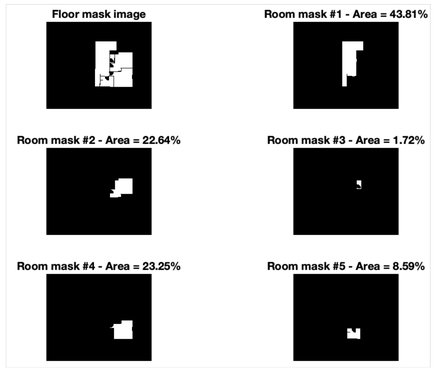
Intersection Over Union (IoU):

State-of-Art IoU	Our IoU
89.7%	60.5%

OCR:

SURFACES	
ENTREE + CUISINE + SEJOUR	28,00 m2
CHAMBRE 1	12,75 m2
CHAMBRE 2	12,00 m2
SALLE DE BAINS	4,75 m2
WC	1,75 m2
DGT	3,50 m2

Merging and Splitting:



Conclusion

In the future, the group will try to improve the performance of FCN by employing FCN32, FCN16, FCN8, FCN4, FCN2 deep networks. In sequential training each model is initialized with the parameters of the previous one.