Bouygues SpaceXYZ Floor Plan Detection : Room Type Detection

Chi Zhang, Qitong Wang, Kaihong Wang, Jiangshan Luo

Boston University Department Of Computer Science

Introduction

The goal of this project is to develop a machine learning model to estimate each type of the room in a given floor plan. The floor plans samples, which provided by Bouygues, are 272 PDF files, without a set of well-defined "ground truths". Our strategy mainly consists of two steps. First, we extracted discriminative features and labeled the ground truth from the training floor plans. Second, we developed two supervised classification models: Backpropagation Neural Network (BPNN) and Support Vector Machine (SVM), then analyzed their performances with model improvements.

Overview

Bouygues Group provided 133 training floor plans, which contain 2 buildings "Eko" and "Equation". The testing dataset contains 139 floor plans from the "Oxygen" building. We first created a converter that reads a PDF floor plan and convert it into an image (Fig 1.), then draw its segmentation image (Fig 2.).

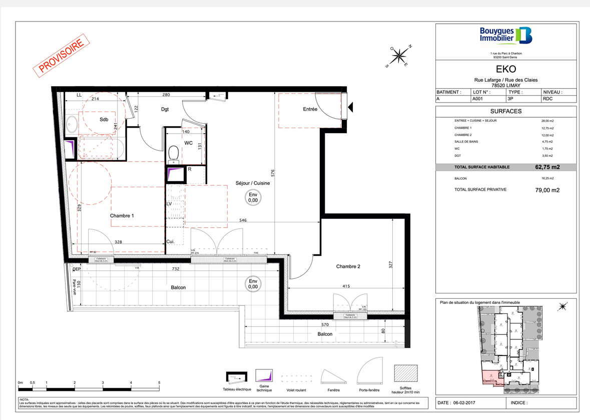


Figure 1.

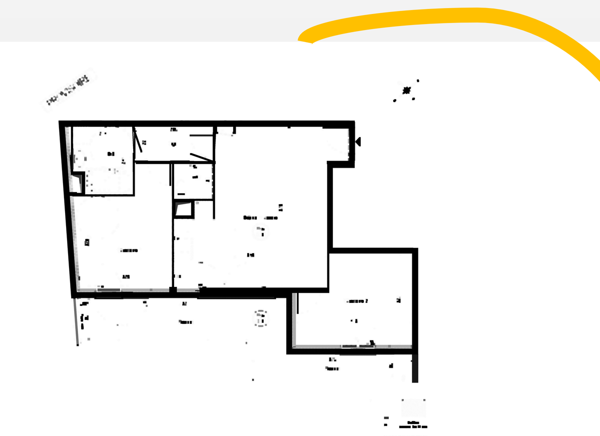


Figure 2.

Then we implemented a features generator, where inputs are a floor plan image and its segmentation image. Our features generator first finds all contours (rooms) in a segmentation floor plan, then calculates 3 features , with one additional feature for each room in this floor plan. In addition, we decided to define 6 labels for 6 type of main room that we want to detect in our models: living room, bedroom, bathroom, WC, intersection, and kitchen. We believe these 6 types of room can be well discriminated by our defined features. After running our features generator, we collected 686 training samples, and 185 testing samples, from "Eko", "Equation", and "Oxygen".

For metrics, we calculate the correct number of predictions divided by total number of predictions as our model accuracy score, and draw the confusion matrices.

Features Definition

Note that our target is to estimate the type of rooms given a floor plan. After analyzing these floor plan images and each main room's configurations, our hypothesis is these main rooms can be easily detected using several discriminative features, instead of running a computationally expensive Convolution Neural Network on the whole image, without compromise on the performance. Therefore, we decided to extract 4 discriminative and explainable features for each main room, which are relative room area, compactness of the room, density of the room, and adjacent room numbers.

1. Relative Area

The definition of one room's relative area is (see F1.).

$$Relative\ Room\ Area = \frac{Room\ area\ (pixel)}{total\ room\ area\ in\ this\ floor\ plan\ (pixel)}$$

F1.

We believe different type of rooms should have different relative areas, for example, the size of living room is often larger than the size of the bathroom and WC.

2. Compactness

For a room compactness (see F3.), we compute it by using the squared of relative room perimeter (see F2.) divided by the relative area.

$$Relative\ Perimeter = \frac{Room\ perimeter\ (pixel)}{total\ room\ perimeter\ in\ this\ floor\ plan\ (pixel)}$$

F2.

$$Room\ Compactness = \frac{(Relative\ Perimeter)^2}{Relative\ Room\ Area}$$

F3.

3. Density

For each room contour, we pick it pixel information from its original floor plan. We define the density of the room (see F4.)

$$Room\ Density = \frac{Number\ of\ nonwhite\ pixels}{Total\ pixels\ in\ this\ room}$$

F4.

Bathrooms usually have high density than other type of rooms (Fig 3.)

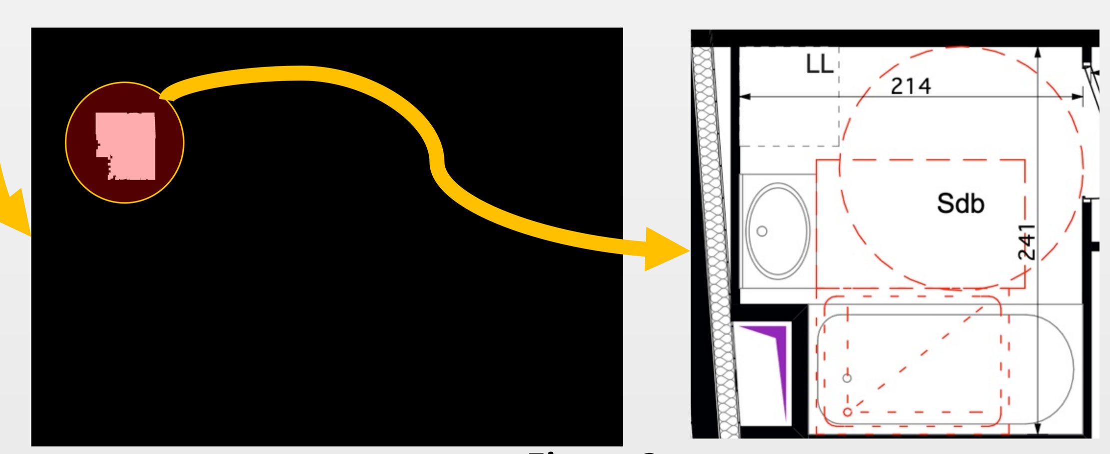


Figure 3.

4. Number of Adjacent Rooms

After analyzing the results of our first version models, with the first 3 features, we decided to created a fourth feature. We found that for some type of rooms, such as intersection, have many adjacent rooms. In addition, number of adjacent rooms can tell a room's relative location in a floor plan. For example, the intersection in Fig 1, has 4 adjacent rooms.

Models

We tried BPNN and SVM models to learn the type of room on 686 training samples, with 3 features: relative room area, compactness of the room, density of the room.

1. BPNN

Our BPNN model uses a multiclass SoftMax regression (F5.) as the output layer and its loss layer uses cross entropy loss (F6.).

1.1 Parameters Tuning:

- Activation function: Sigmoid, Relu
- Layers: 2, 3, 4
- Number of neurons: 5, 10, 30, 50, 100

$$P(y^i = j | x^i; \theta) = \frac{e^{\theta_j^T x^i}}{\sum_{l=1}^{k} e^{\theta_l^T x^i}}$$

F5.

$$J(\theta) = -\frac{1}{m} \sum_{m=1}^{m} \sum_{j=1}^{k} [1\{y^i = j\} \log(\frac{e^{\theta_j^T x^i}}{\sum_{l=1}^{k} e^{\theta_l^T x^i}})]$$

F6.

2. SVM

2.1 multi-class SVM:

Initialize K one vs. rest SVM models, reconstruct label of data: only the corresponding class is labeled as positive and others are labeled as negative classes

Negative class should be labeled to -(1/K) to avoid data imbalance

2.2 Tuning parameters:

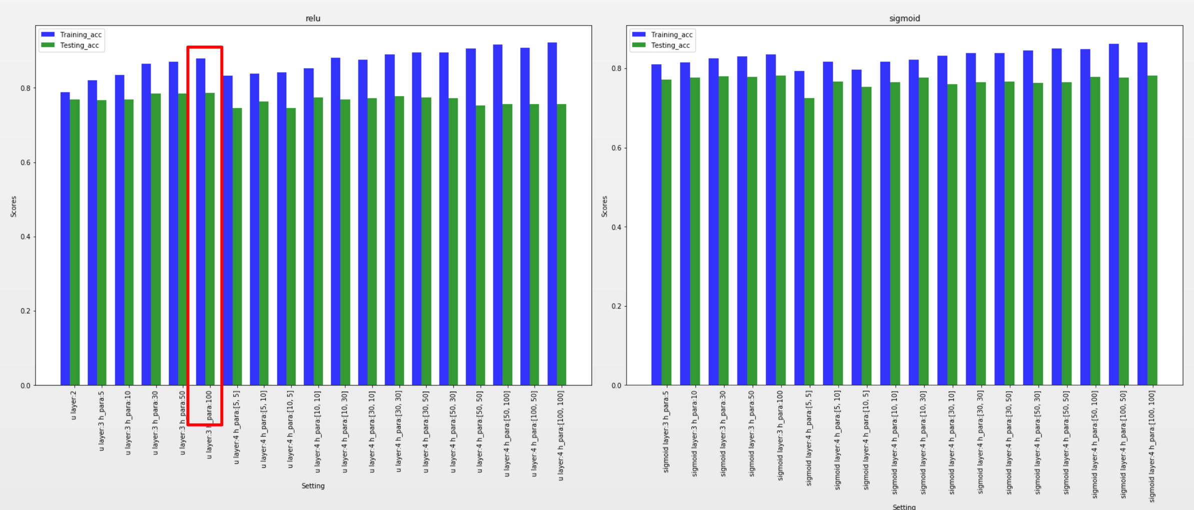
- Kernel: Linear, RBF
- C: 1,3,5,10
- Sigma:0.001,0.01,0.1

$$min_{w,\xi,b} \frac{2}{2} \|w\|^2 + C \sum_{n=1}^{N} \xi_n$$
$$subject\ to\ y_n \hat{y}(x_n) \geq 1 - \xi_n, n = 1, \ldots, N$$
$$\xi_n \geq 0, \qquad n = 1, \ldots, N$$

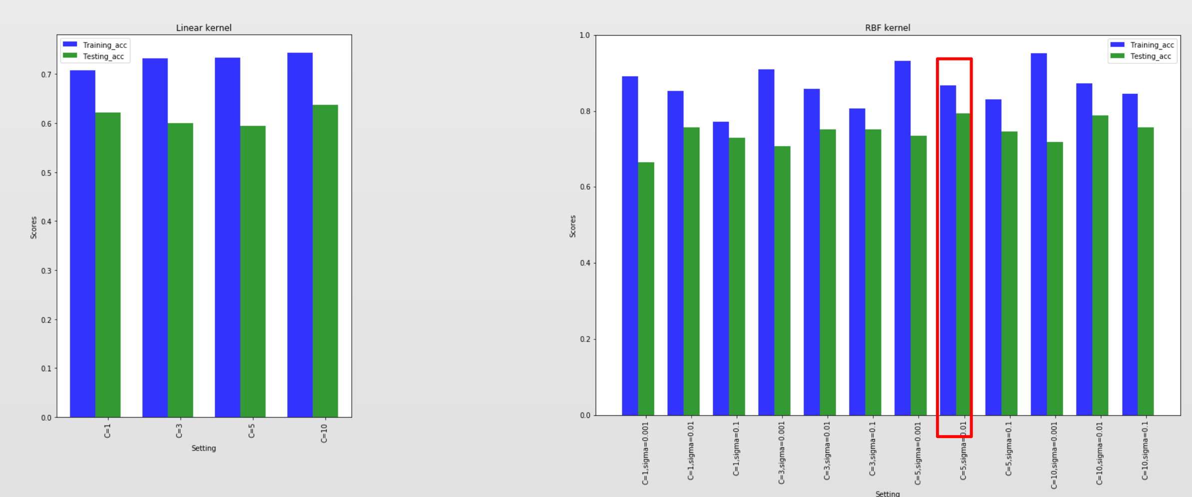
F7.

Analysis

3 features BPNN with Activation Relu and Sigmoid:



3 features SVM with Kernel Linear and RBF:



Noted that all models performances are testing accuracies, and all BPNN accuracies are based on 4 iterations average scores. We selected 2 optimal models from BPNN and SVM, respectively. Based the confusion matrices (Fig 4, 5.) from these 2 model, we found that most of bathrooms were mislabeled as intersections. So we created a fourth features to prevent this confusion and improved our model performance.

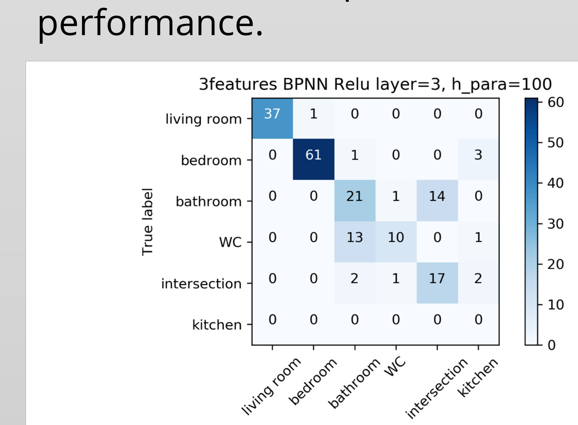


Figure 4.

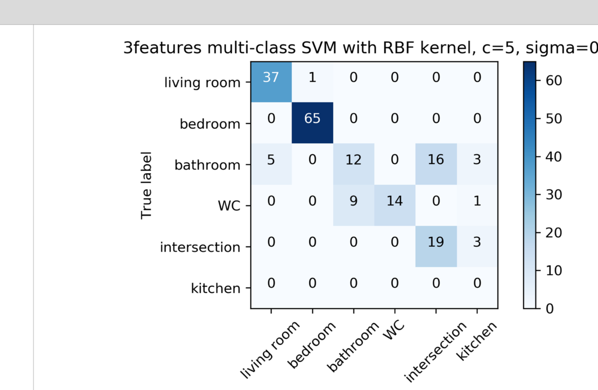


Figure 5.

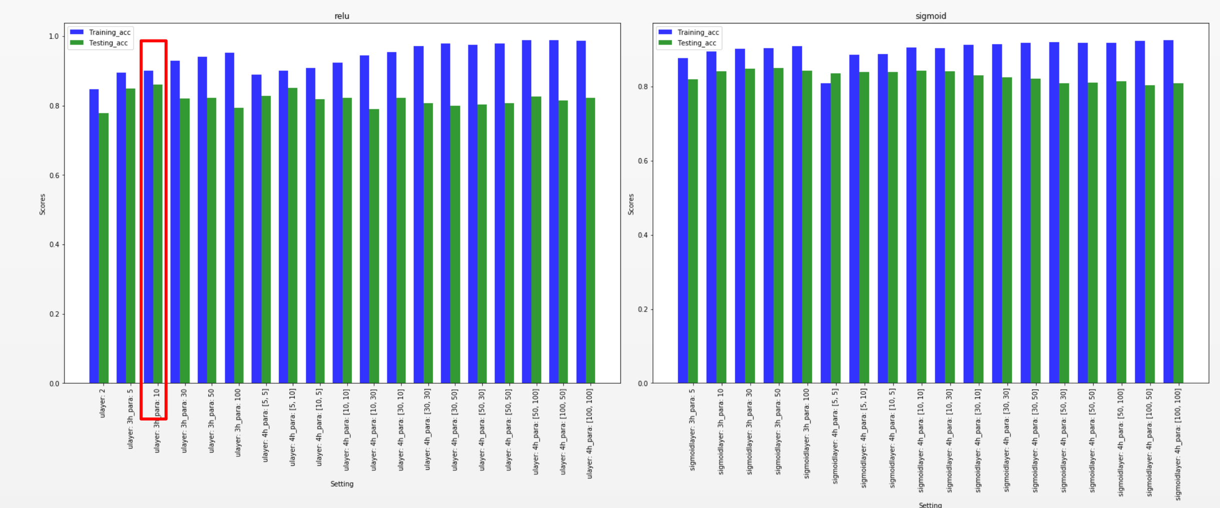
Noted that each entry in the confusion matrix represents number of rooms.

Models Improvement & Results

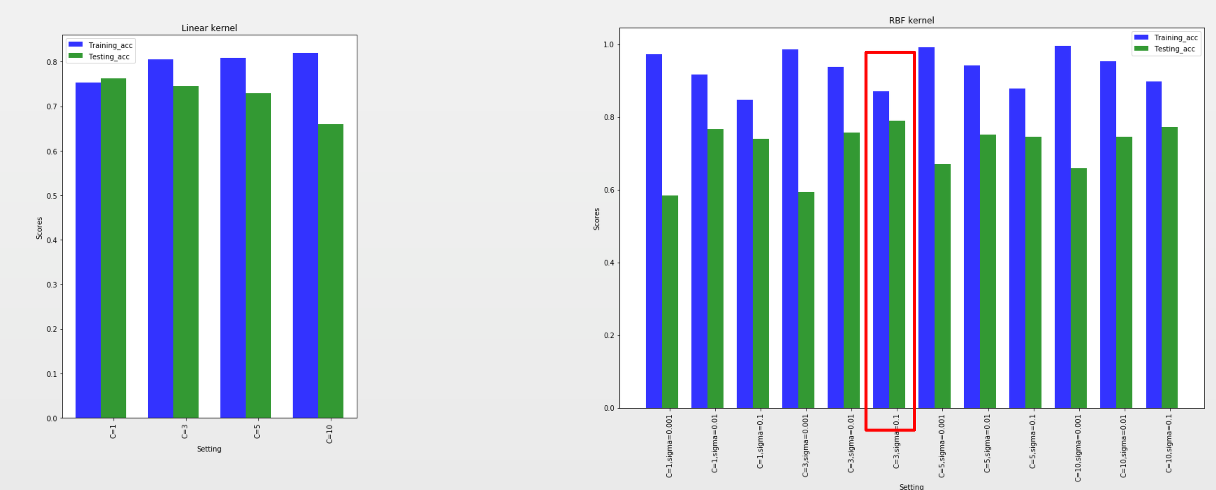
In the analysis section, both optimal SVM and BPNN models have relatively low performances, where models testing accuracies are both 79 %, below 80%.

We added the fourth features: number of adjacent rooms., because we testified this feature can help models to distinguish bathroom and intersection, without influence other rooms.

4 features BPNN with Activation Relu and Sigmoid:



4 features SVM with Kernel Linear and RBF:



We also selected 2 optimal models from BPNN and SVM, where their accuracies are 88% and 79%. Based on the confusion matrices (Fig 6, 7), the 3 layers BPNN model is able to solve the intersection-bathroom mislabeling issue. The WC-bathroom mislabeling issue still exists, but we think it is excusable, since bathroom and WC are both belong to restroom.

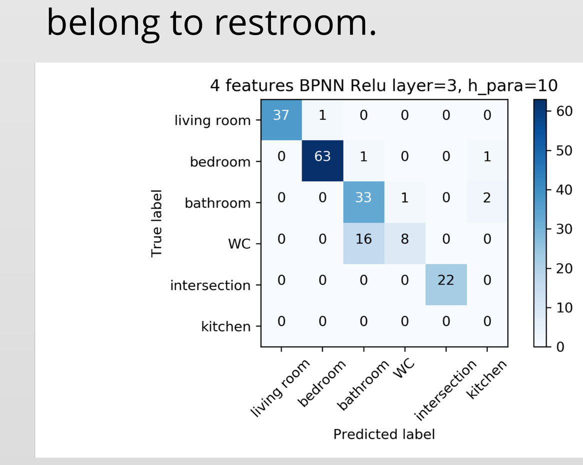


Figure 6.

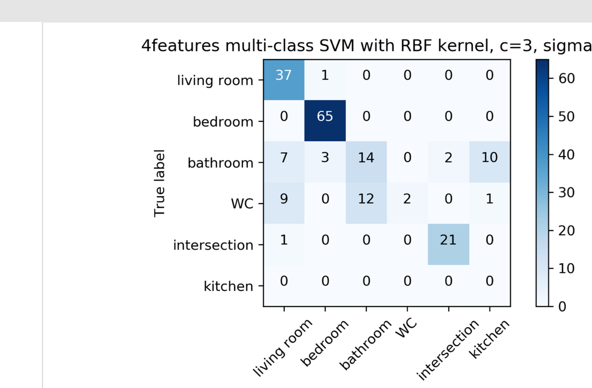


Figure 7.

Conclusion

We proposed automatic models that extracted the features from the original and segmented floor plan and predict type of room using BPNN and SVM. We observed and optimized their performance by setting different experiments and finally our models successfully detected the room types in most cases.