7 Conclusions

Efficient retrieval of similar floor plan images is a critical component of early design phase of construction as it is the primary reason of delay in the design/ conceptualization phase. Traditional CBIR approach is a straight forward and intuitive solution to this problem. The individual techniques proposed in literature, provide an incomplete answer to the question of representing a floor plan in a particular way to facilitate efficient search. Most of the proposed approaches in the literature are tailored for natural images, and some of them for text based document images. However, none of the known results are sufficient to tackle the problem of a unified framework that incorporates multiple modalities of query (sketch and image) for floor plan retrieval. In this chapter the contributions presented during each approach are summarized and the conclusions in the form of strengths and weaknesses of each technique are discussed. Finally, several working lines to be considered in the future are discussed.

In this thesis, it is shown that both the structural and semantic features are important for retrieval of similar floor plan images. It is also argued that hand crafted features are performing better that the learned features (using convolutional neural networks) in case of fine grained retrieval. However, for global similarity (i.e. similarity based on the overall shape of the floor plan), deep features are superior. This thesis also highlights that when sketch is used as a query mode, the traditional techniques (including deep features) fail to scale to this domain shift. Additional processing components need to be incorporated in the framework to tackle such a scenario.

The proposed system can also find application in scenarios where buyers are looking for specific features in house designs, while presenting their requirements to the architect. It can also be deployed in online real estate platforms, where the user can be directed to conveniently query their required floor plans using query as image or sketch and retrieve their required homes from the repository. The novelty of the contribution is that floor plans have been in the past used as graphical documents for analysis and symbol spotting purposes but retrieval based on significant content derived from a floor plan has not been researched. Extensive experimental evaluation on real-world datasets shows that combination of structural and semantic approaches for analysis of floor plans and also employing a combination of traditional and deep learning approaches for feature extraction add value to the overall retrieval accuracy. Both qualitative and quantitative comparisons were made with recent techniques for feature extraction to exhibit the superior performances of the proposed frameworks.

7.1 CONTRIBUTIONS

The key contributions of this thesis are:

- In Chapter 3, a novel approach to perform floor plan retrieval incorporating both symbol spotting and room level analysis is proposed by taking query by example as an image. Inclusion of the room decor arrangement during retrieval is a novel idea that can find application in scenarios where the user has specific choices in terms of the fully furnished floor plans. Furthermore, a spectral embedding approach to represent graphs obtained from the layout is proposed in the baseline retrieval framework proposed in Chapter 3. This technique helped in reducing the computation time taken during the graph matching stage to a considerable level.

- In Chapter 4, a deep learning framework is proposed to extract semantic features from floor plans and those features are used for matching and retrieval of similar floor plan images. Novelty lies in using convolutional neural networks for the task of feature extraction in floor plans which prove to increase the retrieval performance as compared to hand-crafted feature extraction strategies.

- In Chapter 5, a novel framework for fine grained retrieval of floor plans using features like area, number, type of furniture and also adjacencies between the rooms is proposed. Inclusion of the high level features during retrieval is a novel idea that can find application in scenarios where the user has specific choices in terms of the fully furnished floor plans. Also, weighted feature fusion is proposed which helps in weighing features when a user is querying and the user is fixated over certain features to have a mandatory presence while retrieval. This framework has an edge over the deep learning framework in a respect that priority to certain features can be given while retrieval which in not otherwise possible in a deep learning network where globally the features are extracted.

- In Chapter 6, the introduction of sketch as a query mode while retrieval is proposed. Sketch based query mode is very intuitive because a user can conveniently draw his/her requirements on an interface as if they are drawing with pen and paper. Sketch based querying helps in capturing the content better. Two deep learning approaches, one with the usage of Cyclic GANs and the other with the usage of autoencoders are applied for domain adaptation purposes aiding to sketch to image retrieval. The performance of both the systems were obtained to be impressive while retrieval on benchmark datasets.

- Two datasets namely, ROBIN with floor plan images and S-ROBIN with sketched floor plans have also been proposed to benefit the research community to carry out research in the area of floor plan analysis and retrieval. These two datasets help in overcoming the drawbacks that the existing publicly available datasets had in terms of less variation in samples as well as inadequate number of samples to carry out deep learning and retrieval tasks.

7.2 DISCUSSION ON THE RELATIVE PERFORMANCE OF THE TECHNIQUES PROPOSED

In this section we discuss the comparative quantitative performance of the techniques mentioned in Chapters 3, 4, 5 and 6 of the thesis.

In Chapter 3, handcrafted topological features were proposed to be extracted from floor plan images. The Mean Average Precision (MAP) value obtained during retrieval on the ROBIN dataset was observed to be 0.247 in case only room layout was considered while retrieval, on the other hand incorporating both room layout and room decor features made retrieval quite specific. This was due to the fact that the layout similarity in a category of the ROBIN dataset was higher than the combined layout and decor similarity. Thus, the system gave a MAP value of 0.232 in this case. We observed that although, topology and room decor features help in distinguishing two floor plans and retrieving similar ones, but inclusively are not a good enough measure to analyse a floor plan completely. It is very difficult to curate handcrafted features powerful enough to extract

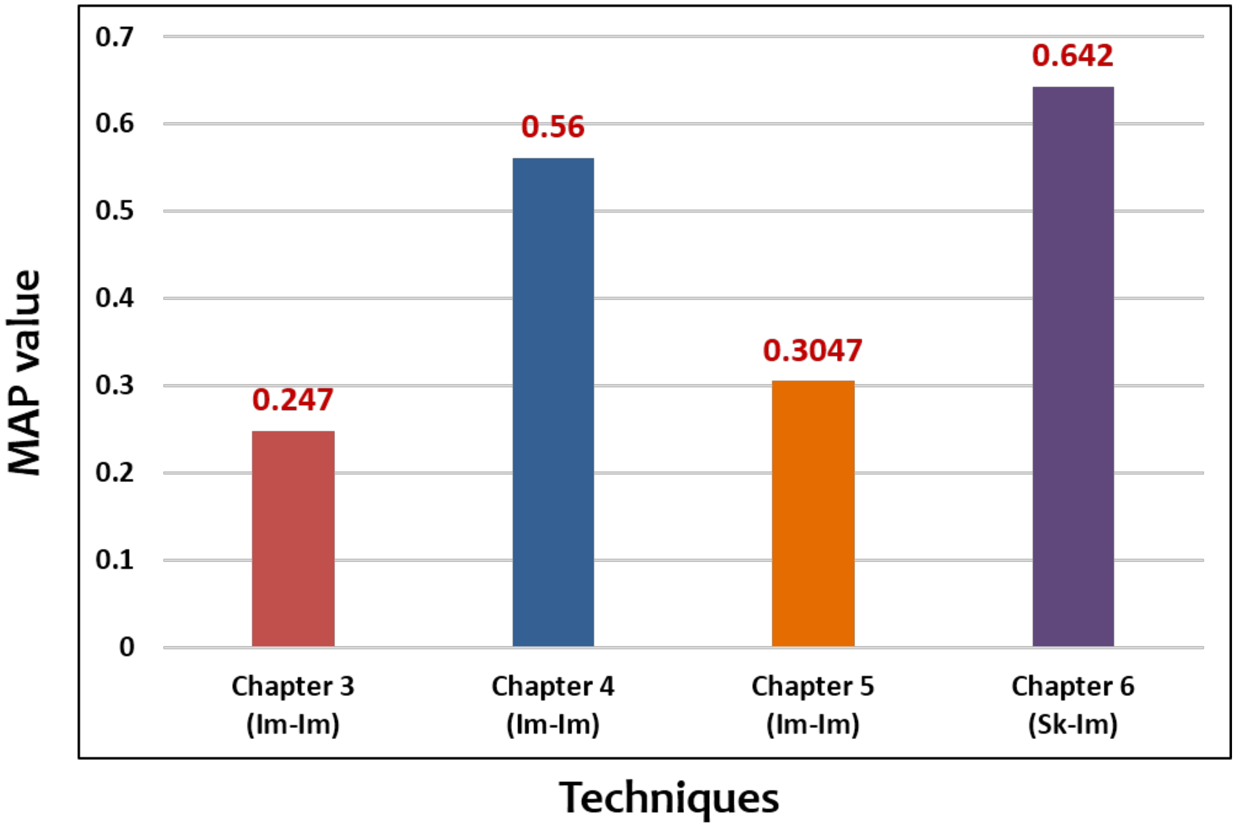

Figure 7.1. : Comparison of MAP values obtained by techniques mentioned in Chapters 3,4,5 and 6.

all the details from a floor plan. Therefore, we explored deep neural networks to extract structural and semantic information from a floor plan.

In Chapter 4, a deep learning framework using CNNs is proposed to extract semantic features from floor plan images and those features are used for matching and retrieval of similar floor plan images. As compared to Chapter 3, the deep learning framework, was able to enhance the retrieval performance considerably. We observed a MAP value of 0.56 on the ROBIN dataset. However, we observed that deep learning is convenient and bypasses the need of curating features manually, the disadvantage is that the layers in the deep learning framework extract features implicitly. This implicit feature extraction at each layer does not give the user the ability to give priorities to certain aspects or features while retrieval.

Therefore, in Chapter 5, high level semantic analysis of floor plans and their subsequent retrieval is proposed. Additionally, a provision to adjust the weights of each feature while retrieval is given in the proposed technique. Such a provision can prove to be more practical to users while looking for similar floor plans sufficing their specific preferences. We observed a MAP value of 0.3047 on the ROBIN dataset using this technique. Although, the MAP value obtained here is less than what was obtained using the deep learning framework for image retrieval, the advantages however are that, weighted feature fusion helps in emphasizing on certain features when a user is querying and has specific feature preferences while retrieval.

In Chapter 6, the querying mode changes and sketch based querying is proposed. Such a query makes the system more convenient for searching from the user's end. The conjunction of

CNN, Cyclic GAN and autoencoders was used in this particular framework for domain adaptation. The MAP value obtained in this case was impressive at a value of 0.642 for ROBIN dataset.

Figure 7.1 presents a bar graph comparing the MAP values obtained by each technique. Im-Im represents, image to image retrieval and Sk-Im represents, sketch to image retrieval as proposed in Chapter 6.

7.3 SCOPE OF FUTURE WORK

The work presented in this thesis, definitely opens the scope of a chain of tasks in many diverse applications, for researchers to solve in future. Some of them are given below.

- In the future, content based retrieval taking a hybrid scheme of querying can be researched. Here, the idea can be that a user can draw/scribble over a floor plan image and in real-time the query modifies according to the addition or removal of the components done by the user in the floor plan image. Such a system can be more interactive to use. Such multimodal inputs would make the process of query representation very convenient to the end-user, and would provide a rich set of specifications to further accurately retrieve similar floor plan samples.

- Future research should concentrate on capturing of better and usable content from floor plans to help in real world scenarios. Researchers may explore on deducing high-level features, which for example can be in the form of accessibility of each room, in terms of shortest path of a room to other rooms or to the exit in a floor plan. This scenario can find application in maintaining safety situations in big architectural buildings and retrieving floor plans abiding by the safety standards set by a regulatory body.

- The main issue with floor plan analysis research is the dearth of publicly available datasets as architects do not prefer to share their designs in the open. Therefore, a good future scope can be to develop a larger dedicated repository of real-world floor plans to foster research in the area of analysis and retrieval in floor plans.

- Acquiring the floor plan images may not always be in digital form. The alternate mode could be through high-resolution scanning. However, scanned documents at times are corrupted with transformations like shear, rotation, etc. Also, during digitization some noise might get added or existing noise in the document might get enhanced. Retrieval of floor plan images under such an uncontrolled situation forms a challenging research problem.

- In Chapter 6, Fig. 6.7, we already propose, a unified multimodal system which can classify query into sketch or image and accordingly perform image to image or sketch to image retrieval. Future research can also explore the combination of feature extraction techniques such as, hand-crafted techniques and deep learning techniques for retrieval. For example, in chapter 4, as we observe that deep learning fails to explicitly give preference to a feature while retrieval which was possible through the technique mentioned in chapter 5. Therefore, a suggested modification can be exploring an approach of incorporating feature importance into neural network learning. The weights connected to an important feature can be biased to have an absolute value larger than the weights connected to the other features.

Annexure A

Floor plan datasets ROBIN and S-ROBIN

Dataset plays a pivotal role in any image processing problem. Availability of a public dataset sometimes makes the life of a researcher easy. However, this is true for only generic problems, and not so for problems which are of some specific nature. Floor plan image analysis is such a problem where a large collection of samples is hardly available. For this reason two public datasets have already been proposed by researchers. They are: (i) SESYD [Delalandre *et al.*, 2010b] and (ii) CVC-FP [de las Heras *et al.*, 2015]. The former has 10 classes of floor plans, where there are 100 samples/class. The latter has 122 scanned floor plan documents divided into 4 categories based on the origin and style. The SESYD dataset is inappropriate for application to a retrieval task due to very less difference between a pair in a class of floor plan images. Also, in the CVC-FP dataset, the samples are insufficient in number for the task of floor plan retrieval.

To train and evaluate the contributions presented in this thesis, two floor plan databases, named Repository Of BuildIng floor plaNs (ROBIN) and Sketched-Repository Of BuildIng floor plaNs (S-ROBIN) are created. S-ROBIN are the sketch replicas of the ROBIN dataset containing floor plan images. These datasets have significantly aided in the content analysis and retrieval of similar floor plans as proposed in this thesis. The primary motivation while creating the ROBIN dataset was to have: i) sufficient number of sample to execute the experiments, and ii) variation (inter and intra class) among the samples. While searching for a dream house, the customer generally has the follwing sequence of requirements in his/her mind; i) the number of rooms inside the house, ii) the global shape of the house, iii) what is the pattern in which the rooms are connected to each other, and iv) if a user is looking for a well furnished home. then which furniture should be placed in the house and in which arrangement. In the following subsections, each of the points mentioned above are elaborated in detail.

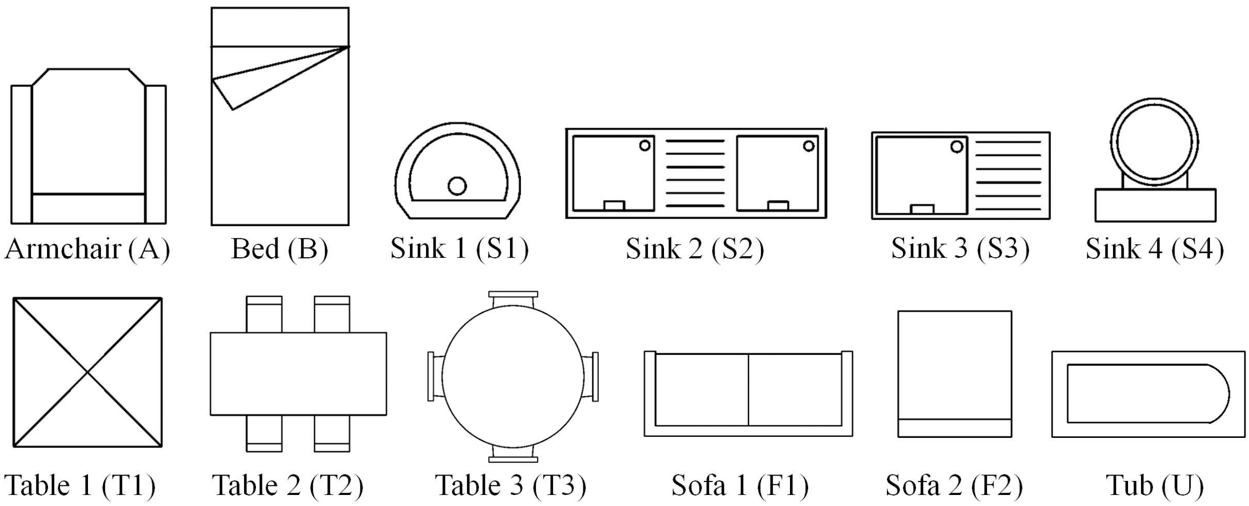

Figure A.1. : Symbol set used as furniture in ROBIN and S-ROBIN dataset.

A.1 NUMBER OF ROOMS

In ROBIN there are three categories of floor plans, depicting the three broad classes of houses. There are houses with 3, 4 and 5 rooms. The purpose and the subsequent label of each room varies depending on the various decors present in them. To be consistent with the existing floor plan datasets, the furniture symbols used in ROBIN dataset are similar to the ones used in SESYD dataset [Delalandre *et al.*, 2010a] as shown in Fig. A.1.

A.2 GLOBAL SHAPE OF THE HOUSE

The second important criteria kept in mind while designing the dataset was the global or overall shape of the layout. To design the floor plans the rooms were arranged either in the horizontal or the vertical direction with respect to each other. While designing, due consideration was given to the fact that, there is no arbitrary shaped floor plan image and also that curvilinear structures in the layout walls are avoided. Moreover, to increase the variability among the samples, difference in arrangement of rooms across 3, 4 and 5 room layout categories was also introduced. Figure A.2 (a), (b) and (c) depict the example of variation in global shapes across a layout category.

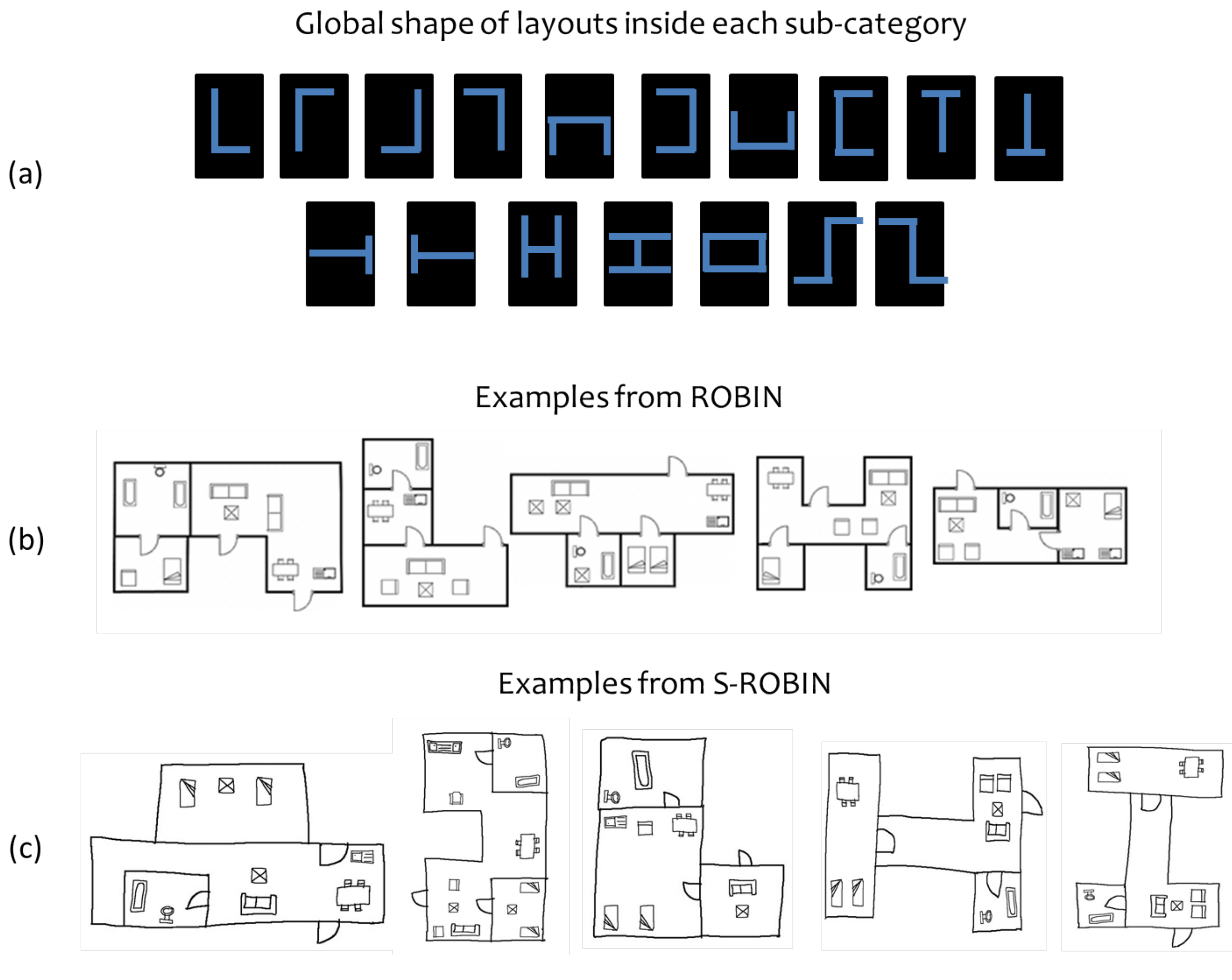

Figure A.2. : 3-Room layouts in the ROBIN and S-ROBIN dataset.

A.3 CONNECTIVITY WITHIN A HOUSE

Doors are placed at the desired places in the floor plans to ensure connectivity between the rooms. Orientation of each door is given special consideration to ensure proper direction of movement within the house. For every floor plan, there is atleast one door lying on the outer wall of the floor plan that opens inside the house and provides entry/exit for the house. Windows are appropriately placed on the walls of the floor plan to ensure proper ventilation and aesthetics. This makes the floor plan more realistic. Figure A.3 (b) can be observed for the arc-type door symbols satisfying entry-exit criteria in the floor plans.

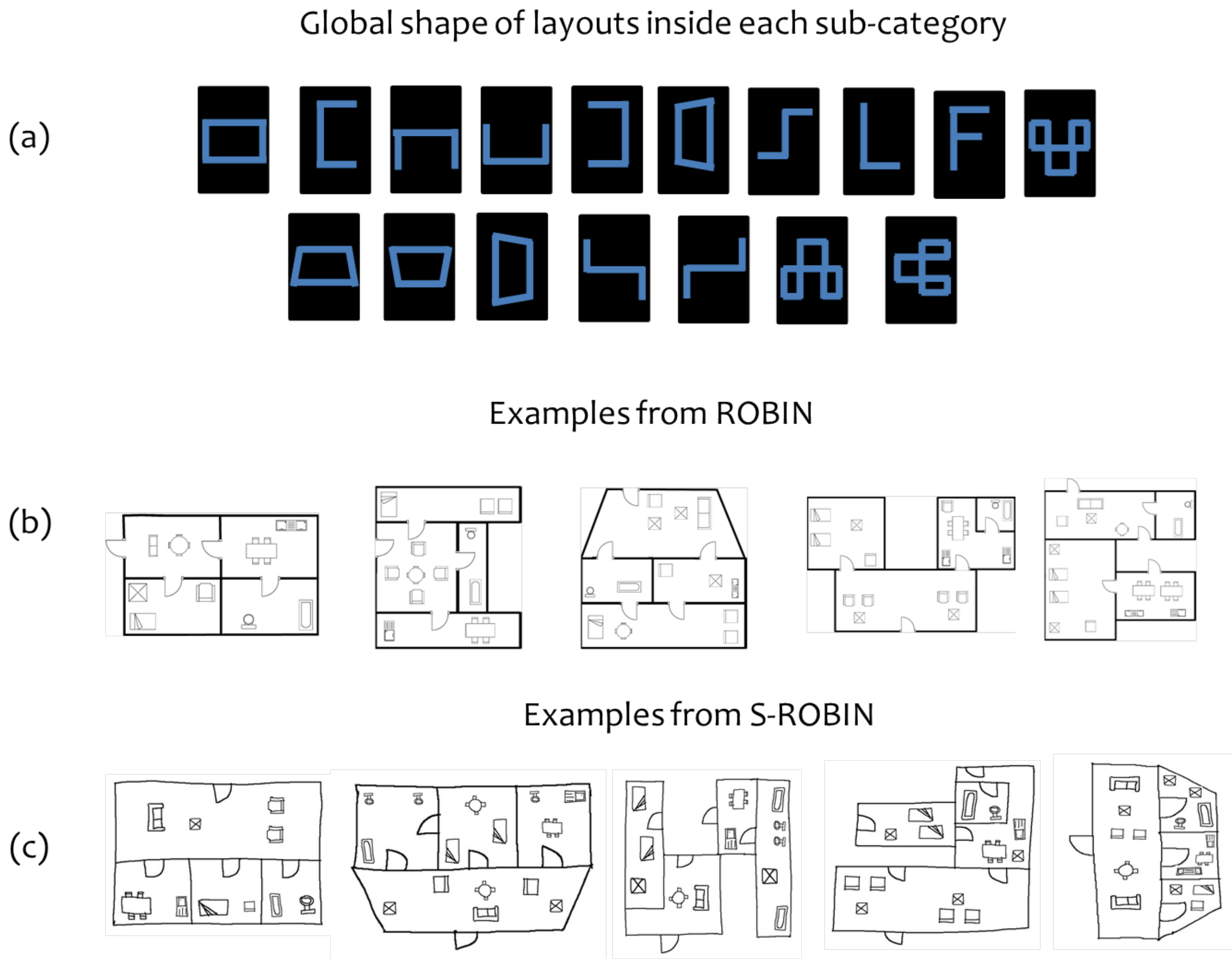

Figure A.3. : 4-Room layouts in the ROBIN and S-ROBIN dataset.

A.4 PLACEMENT OF FURNITURE

Furnitures or decors are inseparable part of a floor plan. To make the floor plans of ROBIN realistic, decor symbols as proposed by Delalandre *et al.* [2010a] were utilized. As shown in Fig. A.1, there are 12 different floor plan symbols. In each floor plan a subset of the entire symbol set was chosen and placed inside every room to give a definition to the room. As the furnitures are indicators of the purpose of the rooms, therefore, while designing, sufficient care has been taken to ensure that a house contains a i) kitchen, ii) bedroom, iv) washroom and a v) lobby. Moreover,

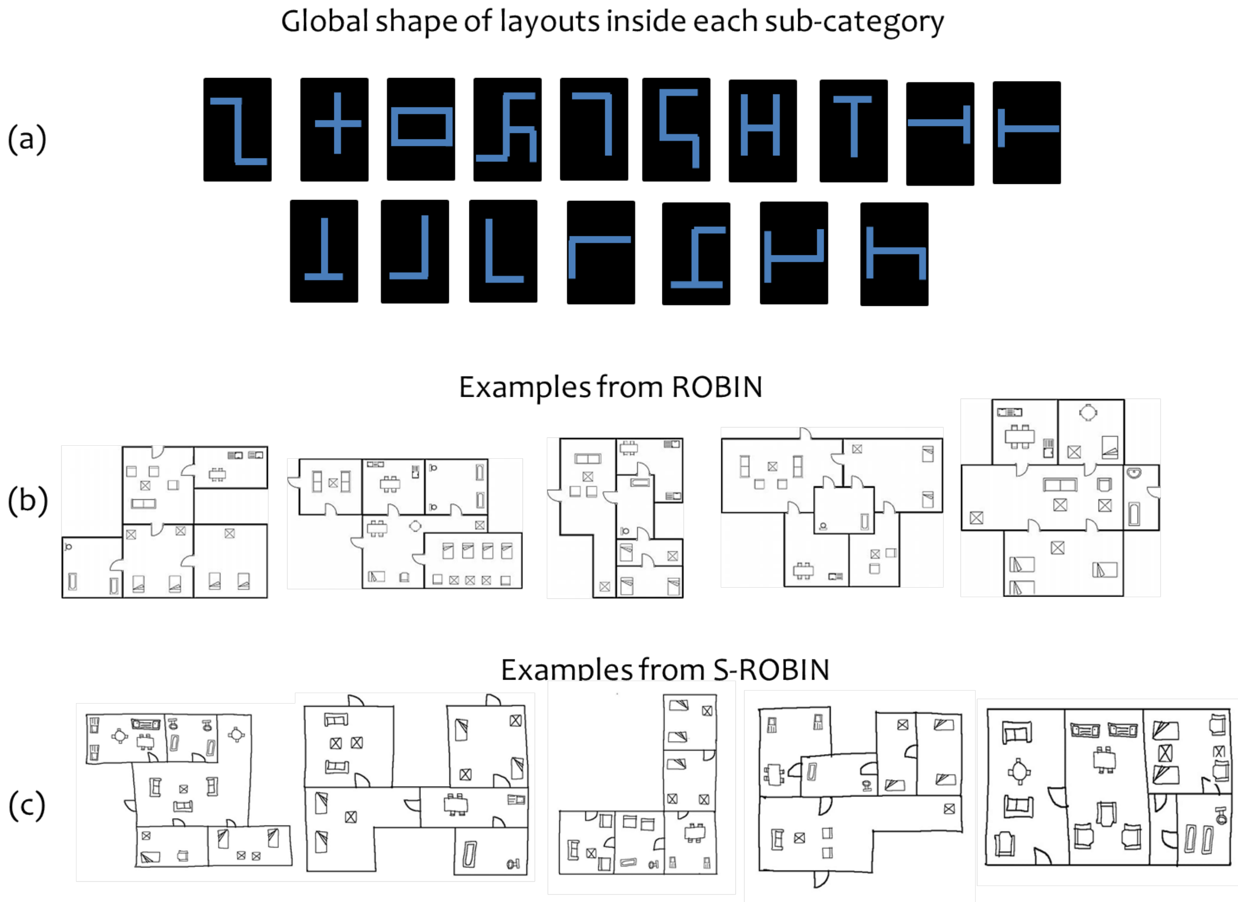

Figure A.4. : 5-Room layouts in the ROBIN and S-ROBIN dataset.

special consideration was given to the fact that the furniture is not placed in an arbitrary manner. For example, furniture should not be placed in a way that it blocks the door. Also, sufficient area should be left vacant after the furniture placement for navigation purposes.

A.5 DATASET CREATION PROCESS

ROBIN dataset was created on similar lines as the SESYD dataset [Delalandre *et al.*, 2010a]. Basic room structure of the floor plan was built using rectangular shape in Microsoft powerpoint. Thickness of each wall was kept uniform to be around 40 pixels. The decor symbols were inserted inside the rooms from the symbol library and door symbols were carefully placed on the room walls. On the other hand, to capture hand drawn sketched floor plans in S-ROBIN, volunteers were asked to draw the floor plans on a digital platform using a Wacom tablet. Wacom tablet is an interactive tool, where the user can draw on a board like tablet using a pen stylus conveniently and the resultant floor plan sketch is captured on the Microsoft paint software by adjusting the thickness of the strokes of the pen stylus. Since, offline mode is being used, the hand drawn floor plans are stored as an image and used for experimentation.

A.6 STATISTICS OF THE DATASET

Both the datasets contain 510 floor plans divided into 3 categories of layouts each with 17 sub categories. The unique characteristic of the datasets is that they are designed keeping in mind the need of a potential buyer. Every prospective buyer wants to have certain amenities and functionalities in their house. Thus, in the dataset, there are three broad categories, which are different from each other in terms of the number and type of rooms present in a floor plan. The broad categories are layouts with 3-rooms, 4-rooms and 5-rooms each which differ in the global shape of layout across their sub-categories. Each broad category is further classified into 17 sub-categories depending upon the global layout shape of the floor plan. Standard notations as mentioned by Delalandre *et al.* [2007], are followed for all the floor plans so that the datasets can be used for other floor plan analysis tasks as well, and not only for retrieval. The dataset has 12 different categories of furnitures inside each layout. These datasets help in better visualization of the floor plans and aid in efficient capturing of various high-level features while fine-grained retrieval. Illustrations of the various sub-categories in the floor plans and how their shapes are modeled are shown in Fig. A.2, Fig. A.3 and Fig. A.4. The size of the floor plans in ROBIN vary from 825×708 to 2916×1845 whereas the resolution in S-ROBIN dataset ranges from 591×517 pixels to 1483×884.

A.7 CONCLUSION

The rationale behind the creation of ROBIN and S-ROBIN dataset was to create a floor plan repository that can be used for research purposes by the Document Analysis and Retrieval (DAR) community. Both ROBIN and S-ROBIN dataset are sufficient for carrying out tasks like, retrieval, semantic segmentation, door to door navigation, etc. The entire dataset along with the annotations can be downloaded from [S-ROBIN, 2018].