Architectural Sketch Recognition and Generation

JAECHANG KO, Department of Architecture, Texas A&M University
SAMRAAT GUPTA, Computer Science & Engineering, Texas A&M University
JACOB STRICKLAND, Computer Science & Engineering, Texas A&M University
SETH POLSLEY, Computer Science & Engineering, Texas A&M University
TRACY HAMMOND, Computer Science & Engineering, Texas A&M University

Sketching is an essential part of the design process for architects, designers, and students alike. It allows them to quickly convey ideas, visualize spatial relationships, and explore new design possibilities. However, not everyone is naturally skilled at sketching, and even experienced designers can struggle with the time and effort required to create professional-looking sketches. In this study, we present a machine learning framework that can transform user sketches into polished, professional-looking drawings. Our framework is based on the Mask RCNN algorithm, which is trained to recognize the objects and details in a user sketch and replace them with high-quality, stylized versions. This allows users to create professional-looking sketches in a fraction of the time it would take to do so by hand. The results of this study will demonstrate the potential of our tool to improve sketching skills and enable designers and students to explore design options more creatively, without the stress and pressure of traditional sketching methods.

ACM Reference Format:
Jaechang Ko, Samraat Gupta, Jacob Strickland, Seth Polsley, and Tracy Hammond. 2022. Architectural Sketch Recognition and Generation . In *CSCE 624: Sketch Recognition Final Report, Fall 2022, Texas A&M University*. TAMU, College Station, TX, USA, 14 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

The phenomenon of drawing lies prominently between the imagination of the architect and the design of a building. Sketching and architecture go hand and hand. Like many other fields, architects do not just go headfirst into building anything. Ideation can begin with nothing more than a simple sketch that is later developed into something more. With the advancement of technology, architecture has moved into the digital world, and sketching will always remain an essential part of design. Sketching is much more than a scribble on a piece of paper in the architectural field. It helps to convey ideas, demonstrates functionality, visualizes blueprints, and illustrates anything that requires interaction. The importance of sketches, however, has declined in stature compared to the past. Advances in Computer-aided design (CAD) tool 3D modeling have facilitated many design processes for architects and designers but have made the various artistic interpretations and actions they possess increasingly forgotten. Consequently, students majoring in architecture and architects/designers in practice have adopted 3D modeling tools rather than sketching behavior.

This phenomenon is based on the psychological pressure with the uncertainty of sketches' quality [18]. The sketch is an expression of the designer's mind, but at the same time, it plays a role as a business medium that requires architects to persuade his/her clients to make it in the real world. From the owner's point of view, they want their building to be beautiful and functional since they invest a lot of money in constructing buildings. The designer and the owner want attractive and compelling sketches to envision the final output. However, it inevitably requires much effort and

CSCE 624: Sketch Recognition. Fall 2022.

Final Report for CSCE 624

time for high-quality sketches [6] . In addition, there is no guarantee that the sketch is as beautiful as the time they invested. Therefore, architects reluctantly rely more on CAD and rendering tools from the initial design stage instead of practicing the unfamiliar act of sketching. Another reason for this decline is the remarkable development of rendering tools. Graphical tools such as Photoshop, Encape, V-ray, Keyshot, or Lumion have enabled architects to produce results with guaranteed quality even though they modeled simple geometries [5]. This repetitive behavior is a process that requires more skill in handling tools than the talent required by artists, so it is a more persuasive working process for architects who need to use their given time efficiently. In addition, the versatility and practicality of hand drawing sketches in the workflow is challenging. While the architect sketched with great effort and time, the sketch is selected or discarded by the answer of the client's "Yes" or "No". Even though the idea addressed in the sketch is selected, those efforts are only an early-stage idea [7]. Commonly, their form and aesthetic are only maintained at the end of the project. In this case, it is almost impossible for the result of a long-time invested sketch to have persistence. On the other hand, 3D modeling is more efficient and continuous because architects can easily modify it based on customer feedback.
However, these CAD tools do not promote their creativity. Sketches using other GUIs require much time and effort compared to hand drawings. In addition, it does not respond resiliently to modifications and design changes. For this reason, frequent modifications and design changes have been seen as a significant barrier to transforming architectural design into a cliche that has lost its personality. As a curator to collect data and control their creative works, architects should define the parameters, design the workflow, and allocate roles to humans and computers.
Therefore, our goal is to expand the applicability of sketching to a more resilient and practical methodology to implement into their routine for the architect, designer, and students. Our approach is to use machine learning to seamlessly merge procedural modeling and interactive sketching, thus enabling an interactive design process leveraging both the intuitiveness, freedom, and flexibility of sketching and the precision, exactness, and detail amplification of procedural modeling. The user does not need to specify tedious procedural rules or rule parameters; instead, they are recognized from the sketch automatically thus enabling untrained users to quickly create complex procedural models. The final output is a sketch tool easily accessible to students, designers, and architects from psychological pressure and stress. Thus, they can explore various design options more creatively, bridging the gap in their sketch abilities. Sketching is a crucial skill for them and it allows them to create beautiful representations, seek variations of existing buildings, and explore new uncharted designs.

2 RELATED WORK

Many researchers have conducted research on sketch beautification and object detection using machine learning approaches. Our liturature review goes into depth on these topics and how they relate to this work.

2.1 Sketch Beautification

2.1.1 Geometrical features and strokes. Beautification of freehand sketches is integral for building robust sketch understanding systems and sketch-based interfaces for CAD. Murugappan et al [12]. proposed a framework that recognizes geometric constraints between primitives and suggested an interface for constraint-driven beautification through user-interaction. The initial steps in this paper include segmentation and recognition, which splits the strokes into its constituent primitives. As a result, freehand sketch is converted to a more simplified representation where the different strokes are closely approximated by a set of parameterized geometric primitives. Langbein et al. [10] used a constraint-based approach to beautify boundary representation models reconstructed from 3D range data. They find geometric regularities approximately present in the model and impose a consistent subset of them to refine the model. Zou and Lee [20] used a similar approach to beautify 3D polyhedral models reconstructed from 2D sketches composed of line segments. Both these methods use priorities to select a subset of constraints in case of inconsistencies. Igarashi et al [9] adopted a strategy to beautify a single stroke (exactly one line segment) one after another to prevent accumulation of recognition errors. The system generates multiple candidates of possible intended geometry which are displayed next to the strokes. This often cluttered the scene making it complicated. Though a high efficiency was achieved compared to the traditional design systems, there were restrictions on the users drawing style as it supported only line segments. Tracy et al. [4] proposed a concept of entropy, which distinguishes between shapes and text strokes. In this paper, the concept of entropy was adopted to classify two different categories. The entropy measurement of a text stroke was higher than that of a shape stroke, which means the difference in information density. This Entropy serves as a single aur that implements the structural properties of shape and text strokes, and can be used to generate a classification between them.

2.1.2 Machine learning approach. Nearly all of these approaches rely on heuristics and empirical parameters, which limits their extensibility. In many cases, there is no automated procedure for selecting optimal parameter values. By contrast, a general-purpose machine learning approach naturally extends to incorporate any number of features. Peterson et al. [14] use a classifier-based method for grouping strokes. That work aims to categorize pen strokes from a freehand drawing into clusters representing distinct elements. To do this, the authors first compute features that capture the spatial and temporal characteristics of the pen strokes as well as the spatial connections between them. A statistical classifier then uses these characteristics to categorize the strokes into different groups. A second classifier then examines pairs of strokes of the same type to determine if they should be clustered together into an object. Similarly, Blagojevic et al. [17] have applied data mining techniques to separate text and shape pen strokes in digital ink diagrams. In particular, they compute a variety of features characterizing the spatial and temporal properties of the pen strokes and use classifiers to distinguish text from shapes. They evaluated the performance of several classifiers, such as LADTree and LogitBoost, for this task. Both of these research efforts demonstrate that these techniques can be more effective than prior approaches.

2.2 Mask R-CNN

The main machine learning method we apply in this work is the Mask R-CNN method proposed in [8], which is a simple, flexible, and general frame work for object instance detection. The first stage of the model is the Region Proposal Network (RPN), which proposes candidate object bounding boxes and predicts the class of the object. The same as the Faster R-CNN method[15], the building blocks for Mask R-CNN. The second stage of Mask R-CNN runs in parallel to the first, and outputs a binary mask for each region of interest. This is a change to most R-CNN systems, which usually classifies objects based on mask predictions. For the application in this work the bounding box output of the model was used and the mask was not.

The Mask R-CNN model used is pre-trained on the Common Objects in Context (COCO) data set by Microsoft [11]. The data set contains 91 common object categories in a total 2,500,000 instances across 328,000 images. Images were collected containing common objects in context of that object, and the images were then annotated using crowdsourcing through Amazon Mechanical Turk. The annotations followed three steps: category labeling where each object was labeled in the broad categories given, instance spotting where each instance of the category was marked in the image, and instance segmentation where the images were manually segmented and a tight fitting bounding box was generated for each object based on the segmentation.

From our reading we could not find any sources that applied Mask R-CNN on sketch data, and specifically architecture-based sketch data.

2.3 Transfer Learning

Applying Mask R-CNN in this work is an example of transfer learning due to the lack of existing training data in the domain and resources required to train a model from scratch. By creating a high-performance learner trained with more easily obtained data from different domains, the learner can train a new learner on a different domain with different data with high level results [19]. One real world example is consider two people who want to learn to play the piano. One person has no previous example with music and one is an experience guitar player. The person with the music background will be able to transfer their previous music knowledge on a different instrument to learn a new instrument more easily [13]. Using the same example applied to this work, a learner trained on recognizing real-world objects will be able to learn to recognize sketch objects faster than a new learner trained only on sketches by applying past training.

3 METHODOLOGY

Our system pipeline can be seen as in Fig 1. We took our custom rough architectural sketch data set and trained a Mask R-CNN model. The model then recognizes the objects, classifies them, and generates a bounding box representing their location in the image. We then beautify the image by replacing the rough sketches with a beautified drawing. For the use case, the user would submit a image containing a sketch and the output would be the beautified sketch following the prior steps. The details of the pipeline sections will be discussed below.

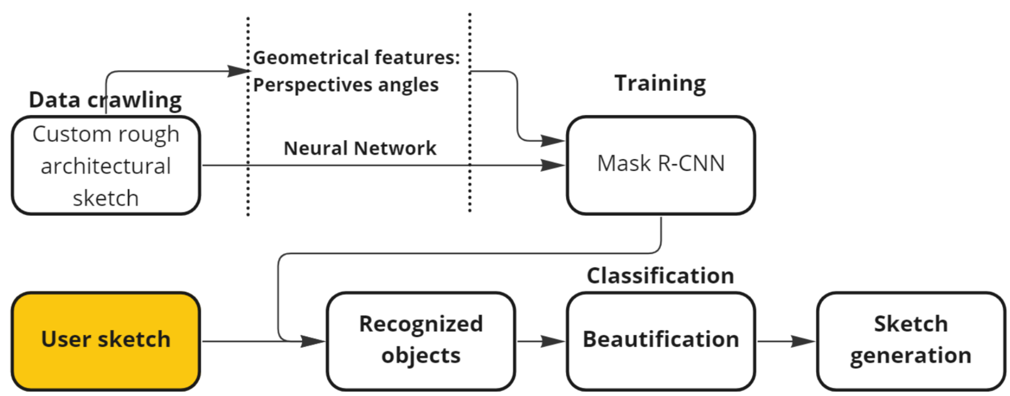

Fig. 1. Model pipeline (placeholder)

3.1 Data set

Our data set contains 50 hand-drawn architectural sketches using Microsoft paint, all of size 500 pixels by 500 pixels. There are four classes of objects inside the images that have been labeled through their color: building (black), person (red), car (blue), and tree (green). Each image in the data set contains at least one of each of these classes. All of the images were drawn from various two point perspectives.

Our data set was annotated using VGG Image Annotator (VIA). This allowed bounding boxes to be places around the

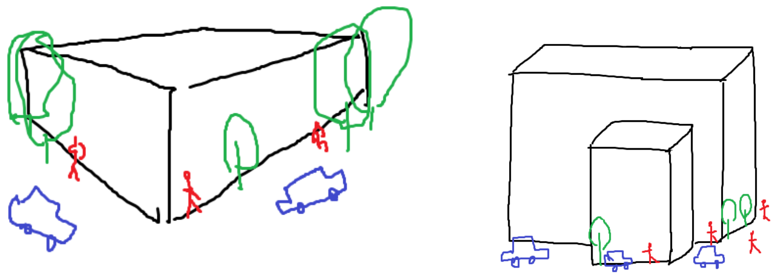

Fig. 2. Example images from custom data set

images, labeled by the class of object inside the bounding box. A JSON file was generated containing all images in the data set, with each image having multiple bounding boxes associated with it.

For the purpose of our training, the data was split into 85% training data and 15% validation data, with separate images generated in the user study for testing.

3.2 Mask R-CNN

For this experiment we used the Mask R-CNN pre-trained model, the source code of which can be found at [8]. Rather than using the original code from 2018, we used a version that has been updated to TensorFlow 2.0. This allowed us to take advantage of newer versions of Python and supporting libraries, along with making set up of the virtual environment easier.

To set up the virtual environment Anaconda was used, with a base of Python 3.7.11. The dependencies from the requirements.txt in the GitHub repository installed the correct libraries, and the setup.py file installed the custom mcrnn library used in all versions of Mask R-CNN. Due to training and evaluating the model on GPU, CUDA 10.1 was needed along with cuDNN 7.6 as the environment was ran on a Windows PC with TensorFlow version 2.2[3]. The last item needed was the pre-trained COCO weights. One online tutorial was very useful in assisting with virtual environment related to this specific repository found [2].

The COCO weights for the model were used as the initial weights in training. For each item in the training data set, the image and JSON file were passed into the model. Using CPU to train the system took 60 minutes per epoch and by training on GPU this time was able to be reduced to 3 minutes per epoch. The default parameters of the model were selected due to lack of time and computational resources required for further testing.

3.3 Beautification

After detecting the bounding boxes for all the elements in the scene, we need to replace them with appropriate images with higher level of detail. But it is not enough to simply replace the bounding boxes with images since we need to take into account the background, size and orientation of the element to be replaced. Thus, the steps performed in order to generate the final image are as follows-

- Remove all sketch elements except the building
- Reorder bounding boxes so that the elements in the background are drawn first
- Randomly select an image for each category of element to be replaced

- Resize image to fit the bounding box

Initially, we were replacing the building with a more detailed building but we felt that this was not in line with the vision of the artist since that is the most important part of an architectural sketch. So instead, we remove the foreground elements from the input sketch and keep the building as it is. We instead focus on making sure that the other elements(trees,cars etc.) fit the scene and create a coherent final image.

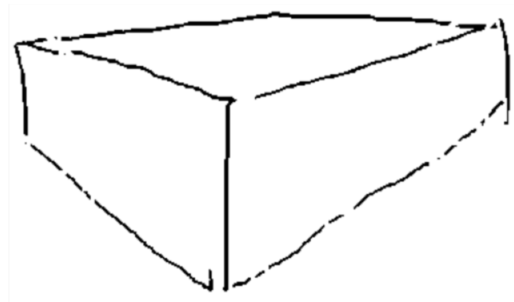

Fig. 3. Isolated image of the building

4 EVALUATION PLAN

A pilot study was conducted to evaluate the proposed application's effectiveness. A total of 10 participants will be selected for this experiment. All participants will have some experience in drawing by hand, but will have no prior experience with the proposed drawing tool. Experiment process is as follows [fig 4]:

1. The participants will be randomly assigned to one of two groups: experimental group (group1: designers) and evaluation group (group2: clients).

2. The experimental group will be asked by group2 (clients) to generate a hand drawing sketch in a given 30 minutes session. Any activities, such as text, call, conversation, or scribbling on papers, will be strongly prohibited while group1 draws a sketch. Group2 will be monitored in isolated spaces and the author will provide them invoices while they hire the architects (two-hundred dollar per hour).

3. Both group will be asked to evaluate their sketch and satisfaction (Questionnaires survey).

4. In the second session, group1 will be asked to generate a quick sketch in a 1-minute session, based on the same design inquiries and same building perspectives.

5. After the second drawing session, both groups will be surveyed again to evaluate their sketch qualities and satisfaction level.

In this pilot study, the Independent variable is the use of our application, and the dependent variable is fee for the design, calculated by the time for the sketches. Participants in group2 will ask to draw architectural objects in sketches, such as buildings, trees, cars, entourages, people, and landscape. In this process, authors maintain the observation stance, in which researchers only talk with participants once they ask for help to minimize the intervention. The conversations are strongly limited to 'instruction' to use the tool or experiment procedure. In this study, the authors aim to determine whether using a tool to quickly generate a sketch improves client satisfaction and their willingness to use the tool compared to using a traditional method of drawing a sketch in 30 minutes. To ensure fair judgement, the authors will provide the judging group with two images of the same design. Additionally, the authors will limit

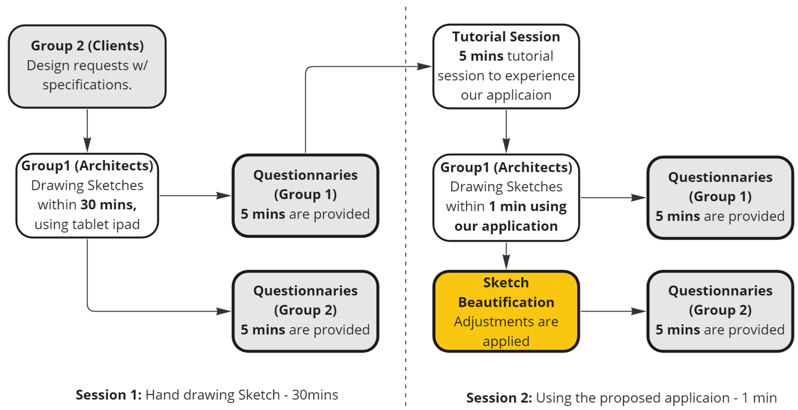

Fig. 4. Evaluation plan

the drawing time to one minute to compare the performance of the two groups under time pressure. This will help to demonstrate the benefits of using this tool to create high-quality sketches within a limited time.

5 RESULTS

5.1 Mask RCNN

Fig 5 shows an example of the generated bounding boxes and masks generated overlaid on testing images from the trained Mask R-CNN model. Using the eye test on our testing images it would appear that the model produces near-perfect results in object detection. The majority of the objects that the model in unable to recognize are trees that are close to other trees. In terms of generated bounding boxes for recognized objects, the model fairly closely matches the expected bounding box. There are some misses, but overall the model is close enough in most cases. With more time and computational resources, the parameters of the model could have been tuned better which may have led to better performance. More data used in training would also improve the performance.

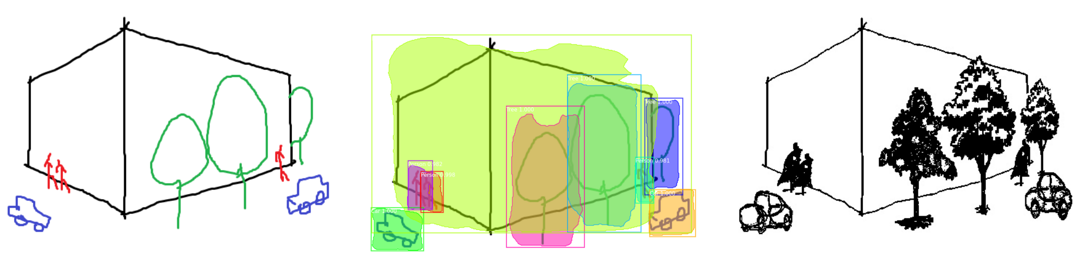

Fig. 5. Example test image through the stages of the model

While developing our system we were unable to determine any quantitative statistics to evaluate the effectiveness of our model. Our analysis involves seeing the number of correctly recognized and then classified objects in the image. We also observe the bounding box compared to its expected result given the object. Our preference is to error on the side of slightly larger than slightly smaller to ensure we are capturing all of the sketch when it is passed through the beautification process. Beyond that the magnitude of difference between the actual bounding box and predicted bounding box was evaluated using the subjective eye test on the intermediary result, along with the final beautified image. The focus of the project was on the user study and providing a useful tool to architects rather than developing the model.

5.2 Statistical Analysis

A paired T-test is used to compare the means of two related or dependent samples. This type of test is appropriate since authors want to see if there is a significant difference between the means of the same group. After the questionnaire survey, authors retrieved multiple related dependent samples to compare, and authors believe that the relationship between the samples may affect the results of the test. The authors of the study used the observation method to record the actions and behaviors of Group 1 during the experiment session. The participants were not given any input or guidance from the moderator and were given 30 minutes to draw a sketch as instructed by Group 2. After the observation, the authors administered questionnaires to both groups and provided invoices to Group 2 for the cost of the sketches. Throughout the experiment, the authors continued to observe the subjects' speech, emotional expression, and reactions.

-Sketch comforts and satisfaction: The questionnaire to group 1 is about the comforts and satisfaction regarding their sketch. Based on the data provided, it appears that the paired t-test shows a statistically significant difference between the two sets of data: Only hand drawing(pre), using our sketch support tool (post). In question 1-3, while

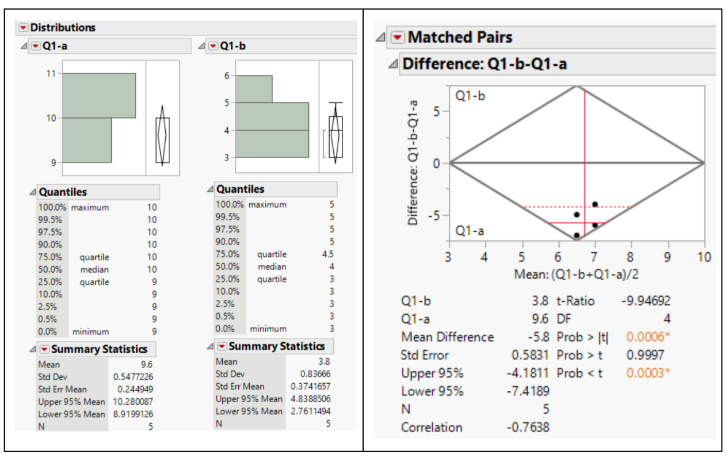

Fig. 6. Q1 statistics summary and matched pairs

hand drawing (pre) satisfied with their sketch quality and expected positive comments from the group2, post group's satisfaction and expectation score was statistically lower (p-value were 0.0006, 0.0004, 0.0012) lower than hand drawing trial.

-Amendments: Question 4 and 5 are measures to see how many features participants want to add or fix after the sketching session based on the sketch methodology: pre vs post. The paired t-test showed that both design methodologies

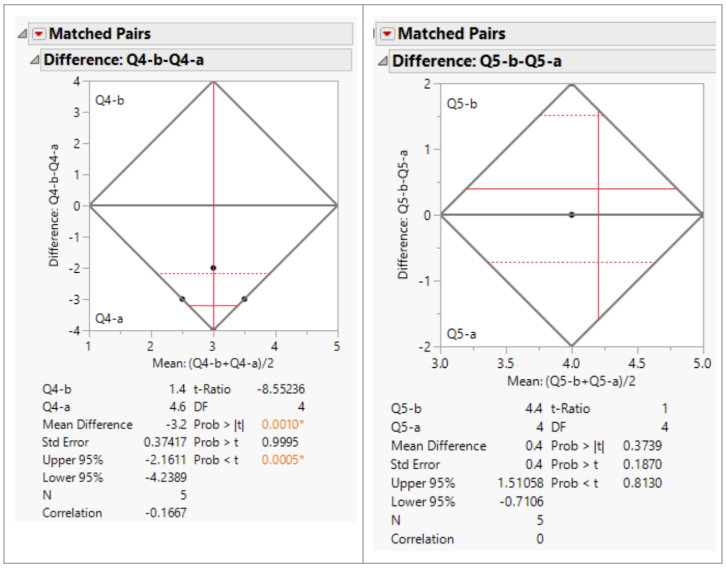

Fig. 7. Q4 and 5 matched pairs

were conducive to the design project, but participants preferred to use the hand sketching method for making changes and additions to their designs. In the interview session, the authors asked which parts of the design the participants wanted to fix or add, and found that all of the participants in the post-test group wanted to focus on the building itself, while the participants in the pre-test group were more interested in improving the entourages, such as people, cars, and trees.

-Expectation: Group 2 represented a client group. In question 9, the participants evaluated both pre and post-test group sketches and scale the score. It has been used as a comparison with question 6, which is a score by group 1 to score their work (scale 1-100) [fig 8]: The paired t-test showed that the post-test group scored themselves lower on the sketch satisfaction survey when using the tool, compared to when they drew the sketch by hand. This difference was statistically significant ($p < 0.0061$). In contrast, group 2's response in question 9 shows that both pre and post were similar across both trials, and the difference was not statistically significant ($p>0.9378$).

-Payment willingness: In questions 11 and 12, the authors measure the clients' willingness to use the tool based on the time to generate the sketches versus the invoice fee. The questionnaire results show that the clients' group is willing to pay for the additional sketches in both methodologies ($p>0.0890$). However, in the case of generating more sketches (in the questionnaire, 10 sketches) all participants in group 2 responded that they are unlikely to ask for the sketches while they are willing to pay money in the case of post-trial, using the sketch support tool [fig 9]. The paired t-test showed that, in general, clients were willing to pay for additional sketches, regardless of which design methodology was used ($p > 0.0890$). However, when asked about their willingness to generate more sketches (10 sketches, in this case), all participants in the pre-test group responded that they were unlikely to ask for the additional sketches while they are willing to pay money in the case of post-trial, using the sketch support tool.

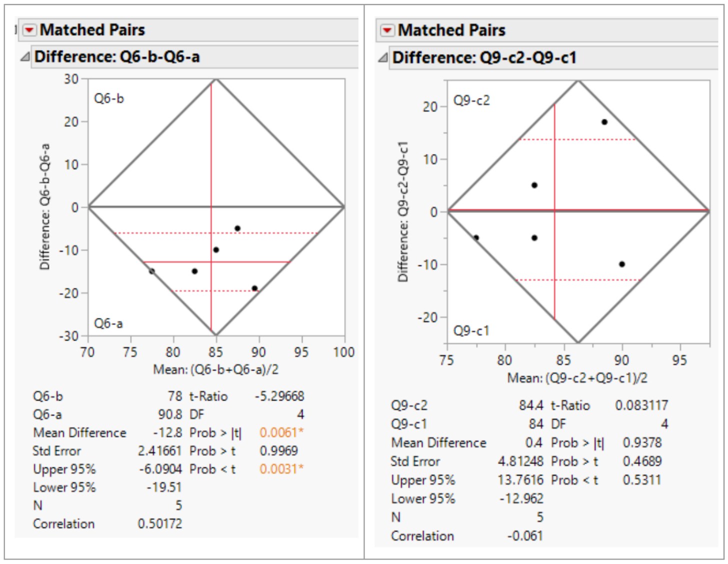

Fig. 8. Q6 and 9 matched pairs

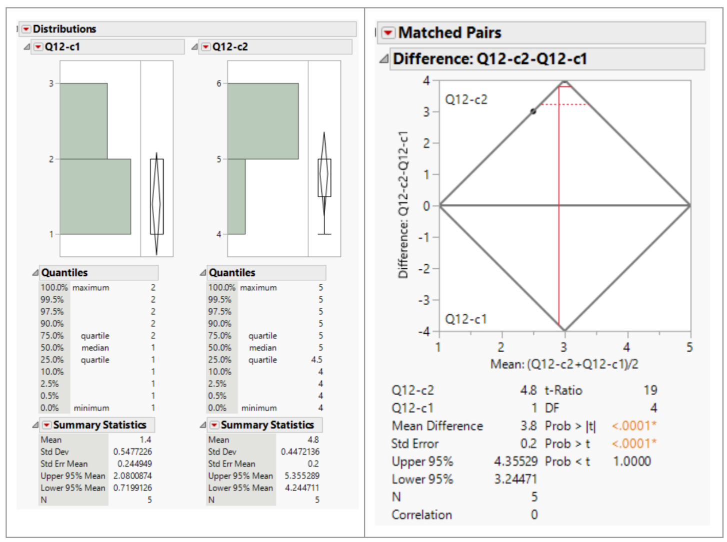

Fig. 9. Q12 matched pairs

6 DISCUSSION

6.1 U-Net

Our first attempt at a solution was to use a modified U-Net [16], where the code was based on the TensorFlow tutorial for image segmentation on the Oxford-IIIT Pets data set[1]. Instead of trying to predict if pixels are in the foreground, background, or border our plan was to predict what color a pixel should be from the black and white image. The U-Net used the pre-trained MobileNetV2 ans the encoder and the pix2pix up-sampling as the decoder. As seen from a sample result in Figure X, results from this method were not successful. This is believed to be due to the large amount of white pixels present in the image, the model could predict an entirely white image and result in 95%+ accuracy. We attempted to remedy this by weighting the classes, such as each class proportional to the number of pixels present in the training set, but with various attempts this was able to produce marginally better results. A larger training set may have helped produce better results. While completing this process we realized that if we were successfully able to produce a near-correct color image based on the black and white image, we did not know where to place the beautified version, which led us to annotate the bounding boxes on our data set and attempt the Mask R-CNN approach. We believe there is potential in the idea of producing the correct colored image based on the black and white image that could see use in future development of this work.

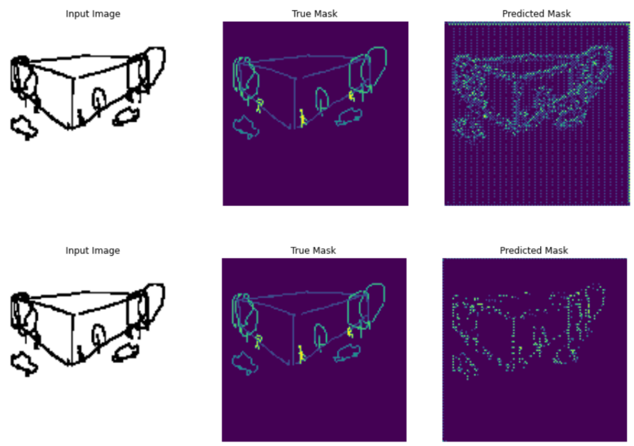

Fig. 10. Results from pixel-by-pixel color prediction using U-Net

6.2 Paired t-test

The hand-drawing group reported higher levels of satisfaction with their results and confidence in dealing with customers compared to the group using the sketching tool. The client also rated the quality of the hand-drawn sketches higher. However, there was no difference in the results and satisfaction with the process of creating the sketches. (Even the 3rd quartile of the post-group was higher.) In the interviews, group2 said that sketching support tool was satisfied in terms of obtaining a decent quality sketch in a short amount of time. When authors asked about if group2 participants are willing to request multiple sketches, they said they would not use hand-drawing, while all respondents said they would request it if they used the sketching support tool. Overall, our findings suggest that the sketching tool does not necessarily produce higher quality work than hand-drawing, but users positively evaluate it for its ability to quickly create good-quality sketches. The difference between the two sets of data is most likely due to the fact that the tool is having a negative impact on the quality of the sketches. This could be because the tool still needs to be well-suited to the task of hand sketching, or because the participants were not familiar with how to use the tool effectively. It's important to note that this analysis is based solely on the data provided, and other factors (such as the participants' experience with drawing, their familiarity with the tool, and any potential biases in the study design) may also be contributing to the observed difference. To gain a more complete understanding of the impact of the tool on sketch quality, it would be helpful to conduct further research and analysis.

7 CONCLUSION

In conclusion, this study has presented a machine learning framework that can transform user sketches into professional and aesthetically pleasing sketches. By using the Mask RCNN to recognize the sketches and replace them with professionally drawn versions, the tool allows users with little sketching experience to easily create professional looking sketches. In a paired t-test, the tool was shown to be effective in allowing users to create professional looking sketches in a fraction of the time it would take to create them by hand. This tool is valuable for students, designers, and architects as it allows them to explore different design options more creatively and without the psychological pressure and stress of having to create perfect sketches by hand. Overall, the tool has the potential to greatly enhance the sketching abilities of its users and improve their design work.

8 FUTURE WORK

There are improvements that could be made in the pipeline. Currently colored images have to be used in training and for evaluating the model due to the constraints on replacing the sketched objects with their beautified objects. With a successful implementation of a black and white to color image (similar to the U-Net mentioned in Methods) black and white images could be used for training the model which would more closely follow a use case of an architecture sketch. The colored image could be then generated after predictions to generate the beautified sketch. An alternative to this would be to replace the objects in the black and white images, but that is well outside the time scope provided in the project.

There are also improvements to the beautified sketch that could be made such as adding perspective to the image. When this was attempted for this work, an issue occurred where the sketch was too sparse to gather the transformation matrix. A more complex input sketch would solve this issue, such as windows added to the building, but that defeats the purpose of the project to an extent. Other solutions could be arrived to with more time.

There are also simple additions that could be made to the work with more data, such as recognizing more objects. The building sketch is also not modified, so applying a form of line beautification or replacing the image based on the perspective is another add-on that would make our tool more robust.

REFERENCES

[1] [n.d.]. https://www.tensorflow.org/tutorials/images/segmentation

[2] 2022. *Object detection using Mask RCNN: end-to-end from annotation to prediction.* https://www.youtube.com/watch?v=QntADriNHuk

[3] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. 2015. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. https://www.tensorflow.org/ Software available from tensorflow.org.

[4] Akshay Bhat and Tracy Hammond. 2009. Using Entropy to Distinguish Shape Versus Text in Hand-Drawn Diagrams. *IJCAI International Joint Conference on Artificial Intelligence*, 1395–1400.

[5] Daniel Cachapa, Armando Colombo, Martin Feike, and Axel Bepperling. 2007. An approach for integrating real and virtual production automation devices applying the service-oriented architecture paradigm. In *2007 IEEE Conference on Emerging Technologies and Factory Automation (EFTA 2007)*. 309–314. https://doi.org/10.1109/EFTA.2007.4416783

[6] Peter Carruthers. 2006. *The Architecture of the Mind.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199207077.001.0001

[7] Iain L. Fraser and Rod Henmi. 1993. Envisioning Architecture: An Analysis of Drawing.

[8] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. 2017. Mask R-CNN. https://doi.org/10.48550/ARXIV.1703.06870

[9] Takeo Igarashi, Satoshi Matsuoka, Sachiko Kawachiya, and Hidehiko Tanaka. 1997. Interactive Beautification: A Technique for Rapid Geometric Design. In *Proceedings of the 10th Annual ACM Symposium on User Interface Software and Technology* (Banff, Alberta, Canada) *(UIST '97)*. Association for Computing Machinery, New York, NY, USA, 105–114. https://doi.org/10.1145/263407.263525

[10] Frank Langbein, David Marshall, and Ralph Martin. 2004. Choosing consistent constraints for beautification of reverse engineered geometric models. *Computer-Aided Design* 36 (03 2004), 261–278. https://doi.org/10.1016/S0010-4485(03)00108-8

[11] Tsung-Yi Lin, Michael Maire, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, and Piotr Dollár. 2014. Microsoft COCO: Common Objects in Context. https://doi.org/10.48550/ARXIV.1405.0312

[12] S. Murugappan, S. Sellamani, and K. Ramani. 2009. Towards Beautification of Freehand Sketches Using Suggestions *(SBIM '09)*. Association for Computing Machinery, New York, NY, USA, 69–76. https://doi.org/10.1145/1572741.1572754

[13] Sinno Jialin Pan and Qiang Yang. 2010. A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering* 22, 10 (2010), 1345–1359. https://doi.org/10.1109/TKDE.2009.191

[14] Eric Peterson, Thomas Stahovich, Eric Doi, and Christine Alvarado. 2010. Grouping Strokes into Shapes in Hand-Drawn Diagrams. *Proceedings of the National Conference on Artificial Intelligence* 2.

[15] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. 2015. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. https://doi.org/10.48550/ARXIV.1506.01497

[16] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. https://doi.org/10.48550/ARXIV.1505.04597

[17] Thomas F. Stahovich and Hanlung Lin. 2016. Enabling Data Mining of Handwritten Coursework. 57, C (jun 2016), 31–45. https://doi.org/10.1016/j.cag.2016.01.002

[18] Christian Swann, Lee Crust, Patricia Jackman, Stewart Vella, Mark Allen, and Richard Keegan. 2017. Psychological States Underlying Excellent Performance in Sport: Toward an Integrated Model of Flow and Clutch States. *Journal of Applied Sport Psychology* 29 (01 2017), 00–00. https://doi.org/10.1080/10413200.2016.1272650

[19] Karl Weiss, Taghi Khoshgoftaar, and DingDing Wang. 2016. A survey of transfer learning. *Journal of Big Data* 3 (05 2016). https://doi.org/10.1186/s40537-016-0043-6

[20] H. Zou and Yong Tsui Lee. 2007. Constraint-based beautification and dimensioning of 3D polyhedral models reconstructed from 2D sketches. *Computer-Aided Design* 39 (11 2007), 1025–1036. https://doi.org/10.1016/j.cad.2007.08.002

A APPENDIX

Pilot Test Questionnaire Survey

Name: ___________________ Date:__________ Testing Group: _______

Please do not write any identifying marks on the survey as participants are meant to be anonymous. All information will be kept confidential. Any concerns can be communicated to (Person, contact info). Thank you for your time and cooperation.

Answer the following questions by circling the most appropriate answer. (First Session)

1. **Do you think the time was enough to draw the sketch? (Scale 1-10) __________**
2. **Do you think your sketch reflects your idea and design intention? (Scale 1-10) __________**
3. **Do you think your client will be satisfied with your sketch? (Scale 1-10) __________**
4. **How many objects do you want to improve if you have more time?**

   None *one or two* *Two or three* *Four or five* *More than five*
5. **Do you think this sketch methodology(1st: hand, 2nd: tool) is conducive to the project?**

   Strongly Disagree *Disagree* *Neutral* *Agree* *Strongly Agree*
6. **Scale your work score (1-100) _________________**
7. **Are you satisfied with the result of processing the client's request about design modification?**

   Strongly Disagree *Disagree* *Neutral* *Agree* *Strongly Agree*
8. **If another chance is given, do you expect that you can conduct better quality sketches?**

   Strongly Disagree *Disagree* *Neutral* *Agree* *Strongly Agree*

Answer the following questions by circling the most appropriate answer. (group C only)

9. **Was the sketch quality satisfactory, considering the time for the sketch? Scale the score. _____**
10. **Was the sketch process satisfactory?_____**
11. **Are you willing to spend more money and time to get the sketch? $120/hour usually.**

    Strongly Disagree *Disagree* *Neutral* *Agree* *Strongly Agree*
12. **Are you willing to spend more money and time to get the 10 sketches? $120/hour usually.**

    Strongly Disagree *Disagree* *Neutral* *Agree* *Strongly Agree*

Fig. 11. Pilot test questionnaire