Interactive Layout Design Interface with User Guidance

Jiahao Weng[1,a)] Haoran Xie[1,b)]

Abstract: It is challenging to create a visually appealing layout for common users, which is a time-consuming task even for skilled designers. In this work, we propose an interactive layout design system with user guidance and layout retrieval so that users can achieve satisfying design results. In particular, we focus on the design of academic presentation slides in this work. With the proposed design interface, the user can refer to the design guidance as a heat map, which is the layout distribution from our collected dataset. The proposed system is data-driven, which allows users to intuitively explore the design data. The user can then edit the layout to complete the layout design. In our user study, we verified the proposed interface by comparing it with conventional design interfaces. The results demonstrated that the proposed interface can achieve high retrieval accuracy while also providing a pleasant user experience.

1. Introduction

There is an increasing trend for online courses and academic conferences. In such cases, teachers and presenters may need to create well-designed slides to give a lecture or give a presentation. The audience ' s comprehension will be influenced by the quality of the slides [3], but it's quite a time-consuming for unseasoned users to create visually appealing slides. Although most of the software is very powerful, such as Microsoft PowerPoint, it might still be more difficult for inexperienced users to create well-designed slides than for experienced designers. Even though PowerPoint provides a toolkit to automatically generate layout suggestions for users to help them design, it has some limitations, which are that it cannot provide a unified style and also it cannot provide suggestions when there is too much content on one slide.

In this paper, we propose an interactive design interface by retrieving the slide layout to support common users. Since most academic presentations on the network are in video format, it is difficult for users to find references using the keyword-based retrieval system to create visually appealing slides, as shown in **Fig. 1**. Our system uses a slide detection technique to extract slides from the slide-based video and a cutting-edge image analysis approach to determine the layout of the slides. Then we apply CNN (Convolutional Neural Network) to extract the features of slides for further use when users retrieve slides. The interface of our system consists of three parts: the first is the canvas part used for user editing; the second is the heat-map canvas for prompting users, and the last one is the retrieval result part used for showing similar slides with the user editing. The design canvas can help users search and retrieve the reference slides conveniently. The heat-map canvas can inspire and prompt inexperienced users to think about how they can design the slide layout. The retrieval result page can provide some reference for users to let them pick the layout they prefer.

[1] Japan Advanced Institute of Science and Technology, Ishikawa, Japan
[a)] s2110022@jaist.ac.jp
[b)] xie@jaist.ac.jp

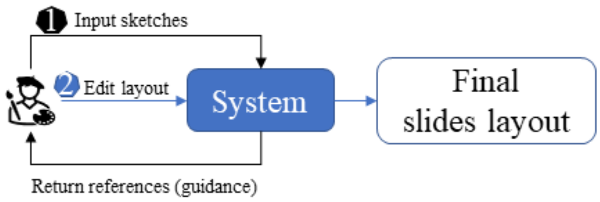

Fig. 1: The research concept illustration. First, users design the layout, and the system will provide guidance for them to help them improve their design.

2. Related Work

2.1 Interactive Design Interface

A design interface can use a freehand drawing that is abstract and lacks visual details. The sketch can let users express their intention intuitively. In such cases, some researches [4], [18] analyze the layout sketches to retrieve the web page, and other researches [11], [15] analyze the sketch stroke to edit the image. Previous researches [6], [8] also utilize user sketches to generate shadow guidance to improve user design skills, which can be used in motion retrieval [10] and calligraphy [5]. The design interfaces have been proposed for cartoon image generation [9] and facial images [11].

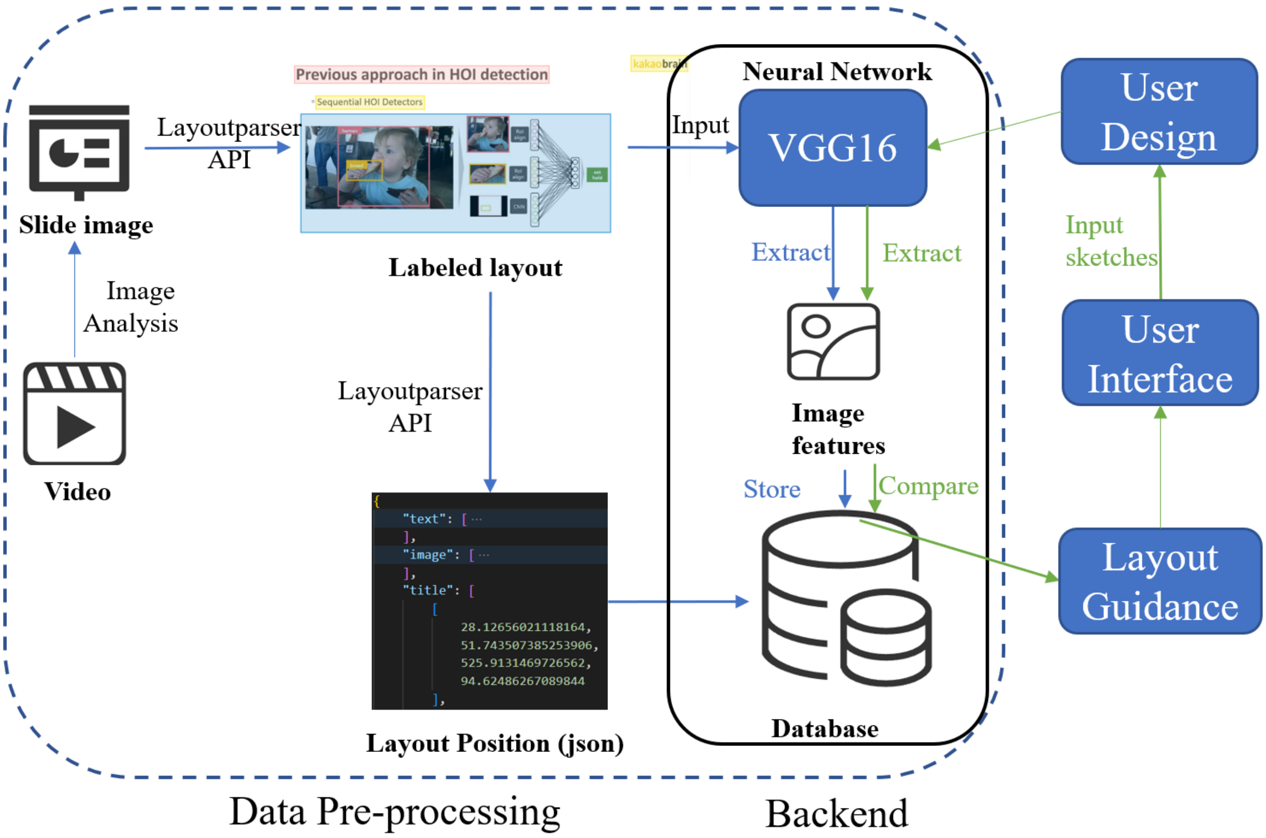

Fig. 2: The framework of our system The left part is the offline computation process for slide extraction; the middle part is the dataset collection; and the right part is the user interface.

In this work, we establish an interactive design interface to provide guidance to users to support them in designing the layout of slides.

2.2 Layout Design and Editing

The VINS [2] uses the UI (user interface) layout as input to retrieve some designs of mobile interfaces. Other research uses layout sketches to retrieve some example web pages that have similar layout design [4]. However, all of this research works only for the web design. In addition, some researchers use the neural network [7] and the transformer network [1] to generate the layout. Nevertheless, our work is focused on supporting users to design rather than replacing them to generate the layout.

2.3 Deep Learning for Layout

Nowadays, more and more researchers are applying the deep learning technique to analyze [14], [16], [17] and generate [1], [7] layout. The PubLayNet [17] also contribute a huge dataset for document layout analysis. In our work, we also utilize the deep learning method [14] to train a model for slides layout detection and analyze, and we apply the neural network [13] to extract the feature of the slides layout.

3. System Overview

In this work, we propose a sketch-based interface for retrieval and support users to design the layout of slides.

3.1 System Framework

The proposed architecture consists of three parts: the offline computation process for slide extraction (pre-processing); the construction of a slide layout feature database; and a user interface as shown in **Fig. 2**. First, we extract slides from academic conference presentation videos and store them in the database. Then, we use the layout-parser [12] to analyze and label the layout of each slide in our database. After that, we use the Convolutional Neural Network (CNN) to extract the features of the labeled slide and store them in the database. The proposed interface consists of three sections: the sketch canvas section, the heat map canvas section, and the result section.

When a user sketches a layout on the sketch canvas, the client will simultaneously submit an HTTP request. The back-end then processes the request and returns image data to the client in the HTTP response. The front-end will then render the data and show the slide image in the result section.

3.2 Slides Extraction

Slides extraction plays a critical role in this system as it is a basic stage in labeling layout, calculating the distribution, and extracting slide layout features as shown in the left part of Fig. 2. We first create an algorithm to compare the image hash of each frame in an academic conference presentation video. If the differential of two frames is larger than

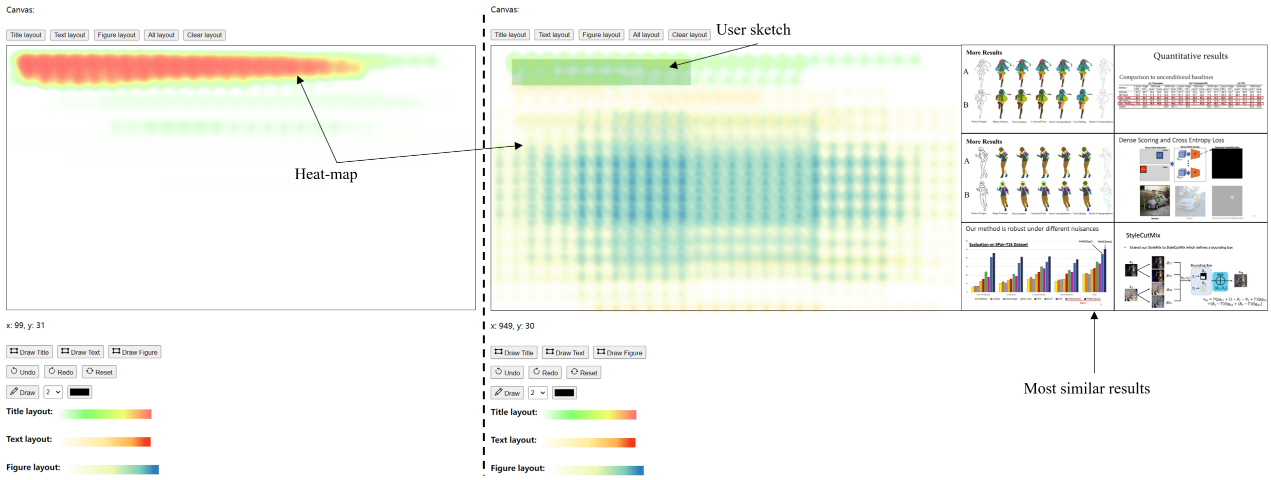

Fig. 3: The left part is the original sketch canvas, the right part is the retrieval page after user sketch the layout.

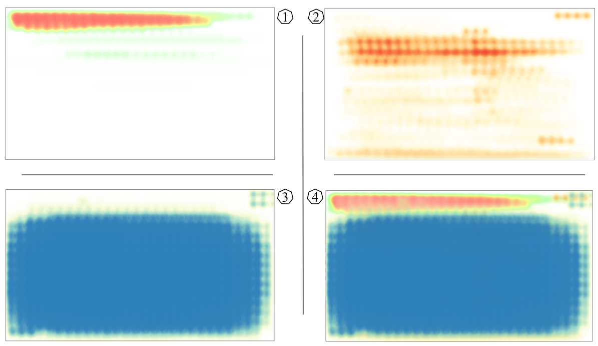

Fig. 4: Part 1 is the heat-map of title layout, part 2 is the heat-map of text layout, part 3 is the heat-map of figure layout, and part 4 is the heat-map of all the layout.

the threshold, it means that the slide has changed; the slide will then be extracted. After the extraction, we calculate the distribution of all the slides and then generate a heat map, which is used to provide reference and inspiration to users. Then we use Detectron 2 [14] to train a model to analyze the slides. After that, we apply layoutparser (a tool that comes with a set of layout data structures with carefully designed APIs that are optimized for document image analysis tasks: selecting layout elements in the document and visualizing the detected layouts) to automatically label the layout of each slide and use CNN to extract the feature. The layoutparser can apply the model trained by Detectron 2 and enables the extraction of complicated document structures using only several lines of code with the help of state-of-the-art deep learning models. Upon completion of pre-processing, all slide features are stored in the database and used for comparison with the user input sketch.

3.3 Design Interface

The design interface can help users retrieve some similar slides drawn by themselves conveniently and efficiently. In addition, it can help users design the slide layout and also inspire them. When users want to get some reference on how to design the slide layout, they can refer to the heat-map (it shows the distribution of each layout by different designers according to its shades of color), as shown in the left part of **Fig. 3**. Then the system will extract the features of the user input sketches and compare them with all the features stored in the database. The most similar slides will be rendered on the web page, as shown in the right part of Fig. 3. When users edit their sketches, the reference result will change simultaneously.

The heat-map shows the distribution of all the slide layouts in the database, as shown in **Fig. 4**. It was created to give users some ideas or inspiration on how to design the layout. We divided the heatmap into three parts: title, text, and figure. This can satisfy most design scenarios. The bottom of the heat-map canvas is the legend; the darker the color, the greater the distribution is. Users can see the distribution of each part or all the distribution by clicking the different buttons, as shown in fig. When users finish or edit the sketch, the distribution of the heat-map will also change simultaneously to continuously advise users how to design the layout.

4. User Study

First, we conducted an experiment to compare the conventional retrieval interfaces with our interface. We invited 12 participants (college students around 20-years-old; 7 males and 5 females) into 2 groups, and another 5 participants (college students around 20-years-old, 5 males) to do the evaluation. After a brief introduction of how to use our interface, we let the first group use the traditional interface, PowerPoint (PPT), and the second group use our interface. We first gave each group an academic paper (a poster about in the field of computer science), and let each participant design three slides (only design the layout): introduction, method, and result. After they finish this task, we ask another 5 participants to evaluate their design (only evaluating the layout). The participant will assess three factors: the content in the slides is reasonably organized, the layout is designed aesthetically, and the slides are designed in a consistent style.

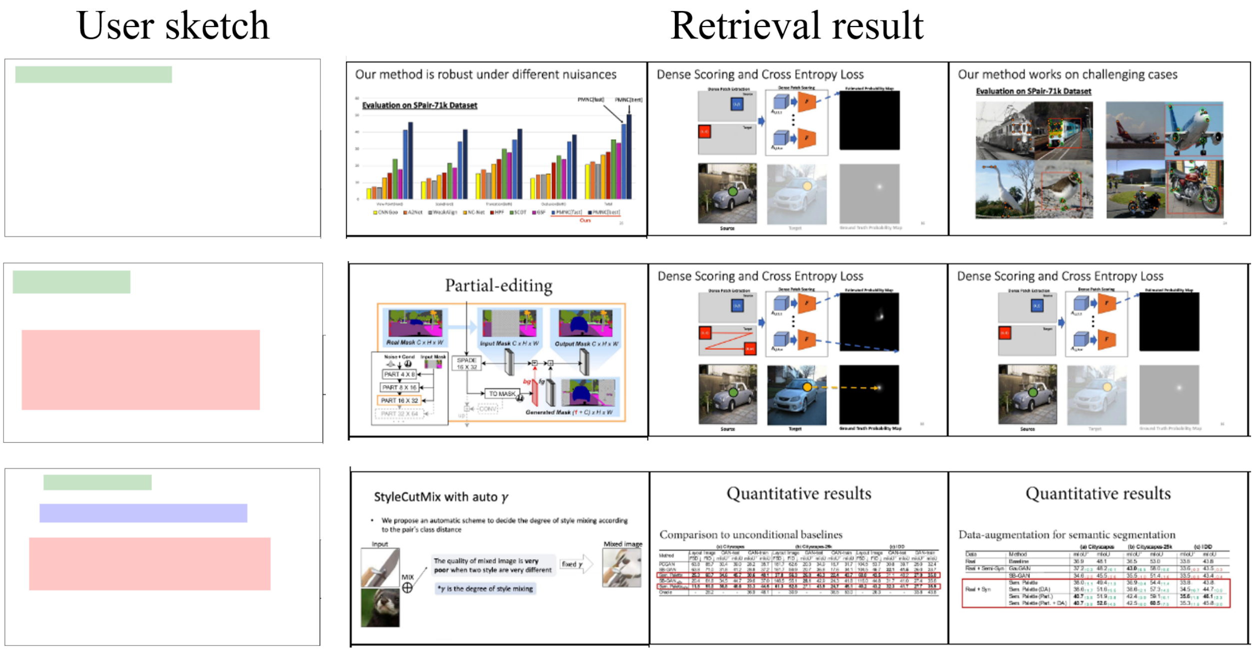

Fig. 5: The retrieval result of layout guidance. The green box is the title layout, the red box is the figure layout, and the blue box is the text layout.

4.1 User Experience

In the second experiment, we conducted a user study to verify user experiences by asking participants to complete the questionnaire. The questionnaire uses a 7-Point Likert Scale (1 for strongly disagreeing and 7 for strongly agreeing). We asked the six participants who used our interface in the first experiment to complete the questionnaire.

5. Results

We discuss the implementation details, the result of the layout design guidance, the result of the user study, and user feedback.

5.1 Implementation Details

In this work, our interface was programmed in Python 3.8 as a website application on Windows 11. A personal computer with AMD Ryzen5 5600X (6-Core), 3.70 GHz, NVIDIA RTX3060 GPU and 32GB RAM was used as the developing environment. In addition, we used Django 3.2 to establish the website, and Vue 3.0 to render the website. We use the Detectron2 to train our model on our own dataset (443 slide images). Our prototype requires 1.12s on average for retrieving the slides.

5.2 Layout Design Guidance

Fig. 5 shows some retrieval results. Different input sketch can retrieve different slides. Even if users didn't complete the whole layout, the system will also retrieve the most similar slides for them according to the sketches. The top part of Fig. 5 shows that the user only sketch the title layout, the system still work well.

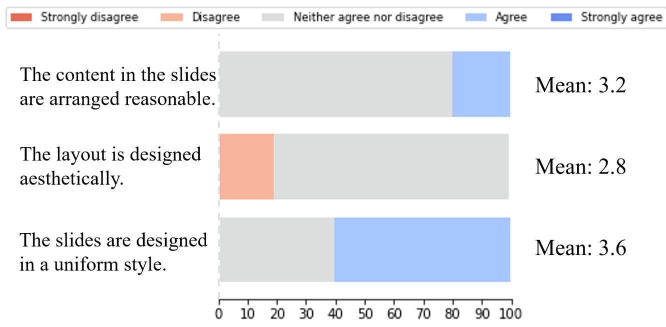

Fig. 6: The results of user study.

5.3 User Study

The questionnaire for comparing experiment uses a 5-Point Likert Scale (1 for strongly disagreeing and 5 for strongly agreeing). The result is shown in the **Fig. 6**. Our interface can provide users with satisfactory slides, whilst it can save their retrieval time. Furthermore, the design of a heat-map has been shown to inspire users on how to design layouts. **Fig. 7** shows some results designed by participants during the user study.

The result of user experience as shown in **Fig. 8**. Although our interface cannot help users design slides with aesthetic appeal, we will extend our research to have the ability to provide references for the content of slides and add more well-designed slides to our database in the future. The user study verified that our interface can help users design more uniform styles than the conventional interface.

6. Conclusion

We introduce an interface that can provide users with references and help them design slide layouts. Using our heatmap canvas can support and inspire users to edit the

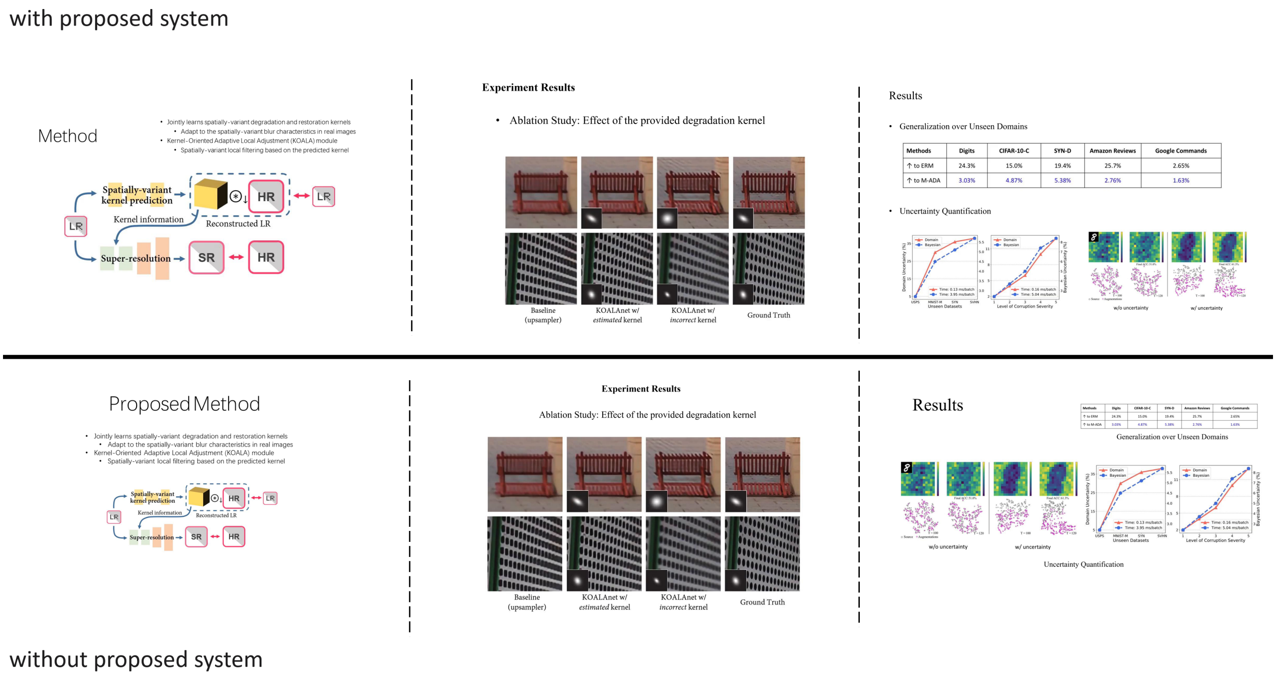

Fig. 7: Examples of the designed results by participants during the user study.

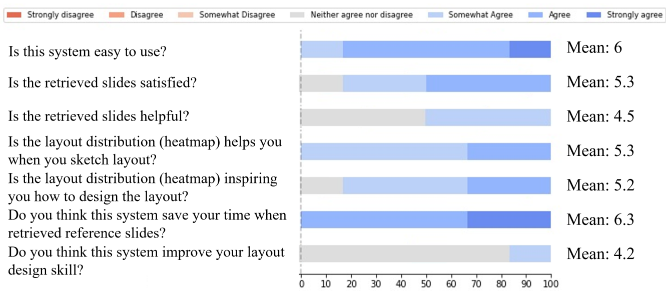

Fig. 8: The results of questionnaire.

layout. Then they can follow the recommended references to design, or they might be inspired by the references and design new layouts. This interface can help amateur users find references quickly and help them learn how to design the layout.

From our user study, we verified that our interface can help users design the slide in a uniform style and arrange the content more reasonably than the conventional interface. It also showed that users prefer to use interactive edits to retrieve the reference and that the reference can inspire them during the design of the slide layout.

However, there was also some negative feedback obtained in the user study. Some users complained that the retrieved slides cannot help them to design because the guidance does not meet their expectations. In this case, we will improve our database to add more quality and well-designed slides. There are 443 slides utilized in the user research. Sometimes, it might be insufficient for users to find the results they are satisfied with. Another complaint is that some users think our interface has difficulty in aesthetically design. This is because in this work we only focus on the layout design. However, it's only a part of slide design. In such a case, we will expand our research to ensure that our interface can provide references of slide content as well as inspire users on how to design detailed slide content.

Acknowledgements

We thank all participants in our user study. This work was partially supported by the JAIST Research Grant and the JSPS KAKENHI Grant JP20K19845, Japan.

References

[1] Arroyo, D. M., Postels, J. and Tombari, F.: Variational transformer networks for layout generation, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 13642–13652 (2021).

[2] Bunian, S., Li, K., Jemmali, C., Harteveld, C., Fu, Y. and Seif El-Nasr, M. S.: Vins: Visual search for mobile user interface design, *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–14 (2021).

[3] Garner, J. and Alley, M.: How the design of presentation slides affects audience comprehension: A case for the assertion-evidence approach, *International Journal of Engineering Education*, Vol. 29, No. 6, pp. 1564–1579 (2013).

[4] Hashimoto, Y. and Igarashi, T.: Retrieving Web Page Layouts using Sketches to Support Example-based Web Design., Citeseer (2005).

[5] He, Z., Xie, H. and Miyata, K.: Interactive Projection System for Calligraphy Practice, *2020 Nicograph International (NicoInt)*, pp. 55–61 (online), DOI: 10.1109/NicoInt50878.2020.00018 (2020).

[6] Huang, Z., Peng, Y., Hibino, T., Zhao, C., Xie, H., Fukusato, T. and Miyata, K.: dualface: Two-stage drawing guidance for freehand portrait sketching, *Computational Visual Media*, Vol. 8, No. 1, pp. 63–77 (2022).

[7] Lee, H.-Y., Jiang, L., Essa, I., Le, P. B., Gong, H., Yang, M.-H. and Yang, W.: Neural design network: Graphic layout generation with constraints, *European Conference on Computer Vision*, Springer, pp. 491–506 (2020).

[8] Lee, Y. J., Zitnick, C. L. and Cohen, M. F.: Shadowdraw: real-time user guidance for freehand drawing, *ACM Transactions on Graphics (TOG)*, Vol. 30, No. 4, pp. 1–10 (2011).

[9] Luo, S., Xie, H. and Miyata, K.: Sketch-based Anime Hairstyle Editing with Generative Inpainting, *2021 Nicograph International (NicoInt)*, pp. 7–14 (online), DOI:

10.1109/NICOINT52941.2021.00009 (2021).
[10] Peng, Y., Huang, Z., Zhao, C., Xie, H., Fukusato, T. and Miyata, K.: Sketch-based Human Motion Retrieval via Shadow Guidance, *2021 Nicograph International (NicoInt)*, pp. 42–45 (online), DOI: 10.1109/NICOINT52941.2021.00015 (2021).
[11] Portenier, T., Hu, Q., Szabo, A., Bigdeli, S. A., Favaro, P. and Zwicker, M.: Faceshop: Deep sketch-based face image editing, *arXiv preprint arXiv:1804.08972* (2018).
[12] Shen, Z., Zhang, R., Dell, M., Lee, B. C. G., Carlson, J. and Li, W.: LayoutParser: A Unified Toolkit for Deep Learning Based Document Image Analysis, *arXiv preprint arXiv:2103.15348* (2021).
[13] Simonyan, K. and Zisserman, A.: Very deep convolutional networks for large-scale image recognition, *arXiv preprint arXiv:1409.1556* (2014).
[14] Wu, Y., Kirillov, A., Massa, F., Lo, W.-Y. and Girshick, R.: Detectron2, `https://github.com/facebookresearch/detectron2` (2019).
[15] Xiao, C., Yu, D., Han, X., Zheng, Y. and Fu, H.: SketchHairSalon: Deep Sketch-based Hair Image Synthesis, *ACM Transactions on Graphics (Proceedings of ACM SIGGRAPH Asia 2021)*, Vol. 40, No. 6, pp. 1–16 (2021).
[16] Xu, Y., Li, M., Cui, L., Huang, S., Wei, F. and Zhou, M.: Layoutlm: Pre-training of text and layout for document image understanding, *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 1192–1200 (2020).
[17] Zhong, X., Tang, J. and Yepes, A. J.: Publaynet: largest dataset ever for document layout analysis, *2019 International Conference on Document Analysis and Recognition (ICDAR)*, IEEE, pp. 1015–1022 (2019).
[18] 中洲俊信, 山地雄土, 柴田智行 and 井本和範: 手書きスケッチによる直感的な資料検索システム, **ヒューマンインタフェース学会論文誌**, Vol. 18, No. 3, pp. 141–152 (2016).