2026 IEEE/SICE International Symposium on System Integration (SII)
January 11-14, 2026. Cancún, México

Automatic Sewing Pattern Generation from Garment Images Using Segmentation and Conditional GANs

Hikaru Suzuki[1], Alessandro Moro[2], Sarthak Pathak[3] and Kazunori Umeda[4]

Abstract— **An automatic method for generating sewing patterns corresponding to dress images is proposed in this study. In garment production, the creation of sewing patterns, the blueprints for garment construction, from design sketches is a highly complex process that demands substantial expertise and experience. Most existing studies focus on learning from entire garments; however, they face the challenge of reduced shape reproduction accuracy for small parts with diverse shapes, such as collars and sleeves. The proposed method segments a garment image into three main parts—bodice, sleeve, and collar—and inputs each part into a specialized sewing pattern generation model, enabling faithful reproduction of even small and complex garment parts. A custom training dataset consisting of garment images and their corresponding sewing pattern images used in actual garment production is constructed. In addition, a part segmentation model and part-specific GAN-based sewing pattern generation models are developed. The proposed method is capable of adapting to diverse garment shapes and variations across parts, thereby enhancing both the accuracy and efficiency of sewing pattern creation in garment production workflows.**

I. INTRODUCTION

In recent years, the fashion industry has placed increasing emphasis on individuality, with many designers presenting works that fully reflect their creativity. However, fashion trends change almost every season, requiring the release of new collections within limited time frames. Under such tight schedules, it is often challenging for designers to fully express their unique styles while maintaining high-quality garment production.

In the following, we outline the standard procedure involved in garment production. First, a fashion illustration is created, which is then used as the basis for drafting sewing patterns. Fabric is subsequently cut according to the patterns and sewn to complete the garment. Sewing patterns, often referred to as the "blueprints" of a garment, are indispensable in apparel production. Creating these patterns is a highly skilled task that requires considerable experience and knowledge in dressmaking. In some cases, limitations in a maker's technical ability can hinder the creation of the desired garment design.

Kato et al. [1] proposed a method that utilizes Generative Adversarial Network(GAN) [2] to generate garment designs that retain the characteristics of a brand. They then evaluated the difficulty of drafting sewing patterns from these generated designs by professional pattern makers. Their findings revealed that the difficulty of sewing pattern creation depended more on the pattern maker's knowledge and experience than on the quality of the generated design itself. Given this context, if sewing patterns—the most technically demanding stage of garment production—could be automatically generated from garment images or fashion illustrations, it could simplify the apparel manufacturing process and contribute to improving working conditions in the fashion industry. Lijuan et al. [3] proposed a method to predict sewing patterns from garment images by combining 3D garment models with a Transformer network [4]. However, the generated sewing patterns often differed in format from those used in actual dressmaking, and the shapes of the input garments were not accurately reproduced. These limitations were attributed to the lack of real sewing pattern data in the dataset and insufficient consideration of garment shape diversity. In our prior work [5], we explored sewing pattern generation from garment images using deep learning models. The approach involved merging all garment parts into a single sewing pattern image and training the model to generate the entire pattern at once. While this method performed reasonably well for large, clearly shaped parts such as skirts, it failed to accurately generate patterns for smaller and more geometrically complex parts such as collars and sleeves. In the case of collars, pattern generation was often unsuccessful altogether. This is due to the fact that treating the sewing pattern as a whole resulted in an abundance of features for the model to learn, causing it to prioritize the more prominent features of large, easily recognizable parts. Consequently, the accuracy for small, complex parts decreased significantly. This study constructs a custom dataset containing real sewing patterns and segments garment images into three parts: bodice, sleeve, and collar. A dedicated deep learning model is then built for each part. In the proposed method, the input garment image is first divided into these three parts, and each segmented part is fed into its corresponding sewing pattern generation model to produce the sewing pattern. This approach enables the generation of accurate and practical sewing patterns that faithfully reflect fine details of the garment's shape, thereby contributing to the efficiency of the garment production process and supporting design work.

[1]Precision Engineering Course, Graduate School of Science and Engineering, Chuo University, 1-13-27 Kasuga, Bunkyo-ku, Tokyo, Japan. h.suzuki@sensor.mech.chuo-u.ac.jp

[2]RITECS Inc., 3-5-11 Shibasaki, Tachikawa-shi, Tokyo, Japan.

[3]College of Engineering Shibaura Institute of Technology. spthak@shibaura-it.ac.jp

[4]Department of Precision Mechanics, Faculty of Science and Engineering, Chuo University, 1-13-27 Kasuga, Bunkyo-ku, Tokyo, Japan. umeda@mech.chuo-u.ac.jp

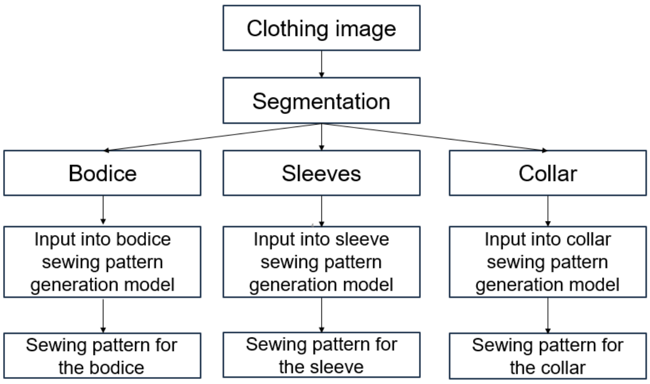


Fig. 1: Flow of the proposed method

II. METHOD

A. Concept

This study aims to generate sewing patterns, which serve as the blueprints for garment construction, from a single garment image. To achieve this, the garment is segmented into three major parts—bodice, sleeves, and collar—and a dedicated sewing pattern generation model is built for each part.

When sewing patterns for all garment parts are generated in a single process, smaller and more geometrically complex parts, such as sleeves and collars, often suffer from poor shape reproduction. To address this issue, the proposed method trains separate models for each part, enabling the capture of part-specific geometric characteristics and allowing the generated patterns to more faithfully reflect fine structural details.

B. Overview

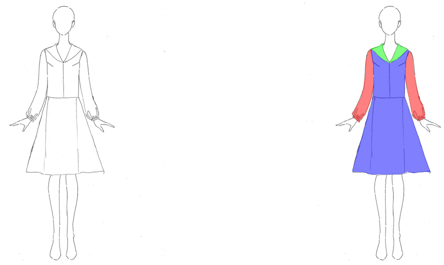
The processing flow of the proposed method is illustrated in Fig.1. In this study, experiments are limited to sewing pattern generation for one-piece dresses.

First, the input garment image is divided into three parts—bodice, sleeves, and collar—using the Segment Anything Model 2 (SAM2) [6]. Segmentation is performed based on predefined seed points, and the corresponding regions are extracted. The extracted part images are then placed at the center of a black background image, converting them into a format suitable for the input of the sewing pattern generation model.

For the generation model, Pix2Pix [7] is adopted, and independent models are constructed for each garment part. Each model is trained using paired data consisting of the input part image and its corresponding sewing pattern image.

During test, each part image is fed into its corresponding sewing pattern generation model to produce the sewing pattern for that part.

The following sections describe each of these steps in detail.



(a) Clothing image (b) Part mask image

Fig. 2: Image used for segmenting clothing

C. Segmentation and Preprocessing

We first segment the garment image into three parts—bodice, sleeves, and collar—using SAM2[6] with part masks that indicate each region. An example of the part segmentation mask used in this study is shown in Fig.2. This mask allows stable region specification in SAM2[6] and enables consistent extraction of the bodice, sleeves, and collar. To highlight shape information relevant to pattern generation, a Sobel filter is applied to extract contour features. After segmentation, each part is cropped from the original image and placed at the center of a 256×256 pixel black canvas, ensuring a consistent input format for the Pix2Pix[7]-based generation model.

D. Pattern Generation Models

In this study, we adopt Pix2Pix[7], a conditional generative adversarial network (Conditional GAN, CGAN) [8], to generate sewing patterns corresponding to each garment part image.

Pix2Pix [7] is specialized for image-to-image transformation tasks and is capable of learning from paired input and output images. This makes it well suited to our task of directly generating sewing pattern images (output) from garment part images (input), as it can effectively leverage edge information and shape features. Moreover, Pix2Pix [7] preserves local correspondences between the input and output images during transformation, which contributes to improve reproduction of fine shape details.

The preprocessed part images were used as the input, and the corresponding sewing pattern images were used as the output. Separate Pix2Pix [7] models were constructed and trained for each of the three garment parts: bodice, sleeve, and collar. This approach enables each model to learn the shape-specific characteristics of its respective part, thereby improving the fidelity of shape reproduction.

The loss function was defined as the sum of the adversarial loss, which encourages realistic outputs, and the L1 loss, which promotes pixel-level consistency with the ground-truth patterns.

Garment parts generally exhibit a high degree of shape variation, and differences in shape often lead to significant changes in the structure of the sewing patterns. When multiple shape variations are trained within a single model, the model tends to prioritize frequently occurring or distinctive shapes in the training data, while the generation accuracy for less common shapes decreases. This effect was especially pronounced for collars, where inter-class shape differences are substantial; training all collar types together resulted in outputs biased toward certain shapes, making it difficult to accurately reproduce the unique characteristics of each type.

To address this issue and simplify the task, we first classify the input images by shape type and then perform training with dedicated generation models for each category. Specifically:

- Collar: Classify into one of five collar shape types and input to the corresponding collar-specific generation model.
- Sleeve: Classify as either short or long sleeve and input to the respective generation model.
- Bodice: Classify based on the presence or absence of a waistline seam and input to the corresponding generation model.

This strategy allows each model to focus on a restricted set of shape features (e.g., corner curvature, length, width), enabling the generation of sewing patterns that more faithfully reflect the shape characteristics of the input garments.

E. Dataset Construction

In this study, we constructed a custom dataset consisting of paired input images (garment images) and output images (corresponding sewing pattern images) for training the sewing pattern generation models. First, garment images containing diverse shapes for each of the three target parts—bodice, sleeve, and collar—were collected from the internet. Since the shape of a garment can vary significantly depending on the wearer's body shape and pose, the dataset intentionally included images of garments with the same design worn by different individuals, as well as images where the wearer adopts various orientations and poses. This design choice was intended to enhance the model's ability to handle shape diversity.

Next, as described in Subsection II-C, a Sobel filter is applied to the collected images to emphasize shape boundaries, seams, wrinkles, and other details useful for sewing pattern generation.

We then annotated the part regions using LabelMe [9], enclosing each part with a polygon. Based on the annotated coordinates, each part was cropped and placed at the center of a black background image with a resolution of 256×256 pixel to create the training input images. Furthermore, all cropped part images were horizontally flipped for data augmentation, effectively doubling the dataset size and improving model generalization.

For the output images, sewing patterns corresponding to the input garment images were designed. In the case of collars, multiple variations (two to four) were prepared even for the same collar type, ensuring greater shape diversity. The produced sewing patterns were scanned, edited, and placed at the center of a 256×256 pixel square image with a gray background. The gray background was chosen to enhance the visibility of differences between the input and output images in the image-to-image translation framework of Pix2Pix [7], thereby potentially improving learning efficiency.

TABLE I: TYPE OF COLLARS, NUMBER OF SHAPE VARIATIONS, AND NUMBER OF IMAGES

Collar Type	Shape Variations	Total Images
Shirt collar	4	100
Stand collar	3	80
Sailor collar	4	93
Flat collar	4	91
Shirt collar with band	2	98
Total Images		**462**

TABLE II: TYPE OF SLEEVES AND NUMBER OF IMAGES

Sleeve Type	Total Images
Long sleeve – straight	60
Long sleeve – tight	60
Long sleeve – puff	60
Long sleeve – flare	60
Short sleeve – straight	60
Short sleeve – puff	60
Short sleeve – flare	60
Total Images	**420**

The final dataset consist of paired input and output images for each part. Table I, II, and III summarize the shape variations and number of images for each dataset category.

III. EXPERIMENT

A. Experimental Environment

The experiments were conducted on a PC equipped with a single NVIDIA GeForce RTX 4080 GPU (24 GB VRAM). Python 3.12 was used for implementation, with PyTorch 2.2 as the deep learning framework.

The dataset used in the experiments was custom-built for this study, targeting dresses and dividing them into three main parts: bodice, sleeves, and collar. For each part, paired data consisting of an input image (garment image) and a corresponding output image (sewing pattern image) was prepared. The dataset was split into training, validation, and test sets with a ratio of 8:1:1, and training was performed independently for each part.

The training conditions for all models were set to a batch size of 1, a learning rate of 0.0001, and a total of 1000 epochs, with models saved every 100 epochs. During training, the transition of the loss function was recorded, and among the saved models, the one with a comparatively low loss value and judged by two experienced garment makers to most faithfully reflect the shape of the input garment in the generated sewing pattern was adopted as the final model.

B. Results of Garment Image Segmentation

This section presents the results of garment part segmentation using SAM2[6] as a preprocessing step for sewing pattern generation.

Fig.3 shows the input garment image and each part segmented using the proposed method.

As shown in Fig3, the bodice, sleeves, and collar parts were successfully and appropriately extracted.

TABLE III: TYPE OF BODICES AND NUMBER OF IMAGES

Bodice Type	Total Images
A-line	66
Panel line	62
Box line	68
Waist seam (Gathered skirt)	70
Waist seam (Straight skirt)	65
Waist seam (Trapezoid skirt)	61
Total Images	**392**

(a) Input (b) Part mask image

(c) Collar (d) Sleeve (e) Bodice

Fig. 3: Segmentation results

C. Result of Sewing Pattern Generation

This section presents the output results of the sewing pattern generation models for each garment part (collar, sleeve, and bodice) and conduct a qualitative evaluation of each model based on these results. Fig.4 and Fig.5 show the generation results for collars, Fig.6 shows the results for sleeves, and Fig.7 shows the results for bodices.

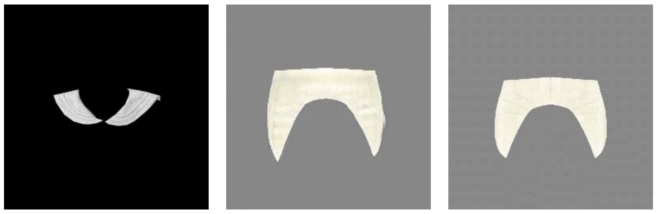
From Fig.4 to Fig.7, it can be observed that the models for collars and sleeves generated sewing patterns that relatively well reflect the shapes of the input images. In contrast, for bodices, parts of the generated shapes were distorted, indicating that accurately reproducing the input shapes was more challenging compared to the other part-specific models.

D. Quantitative Evaluation

We evaluate shape accuracy using Intersection over Union (IoU) [10], which measures the overlap between the generated pattern and the ground-truth pattern. IoU is defined as:

$$\mathrm{IoU} = \frac{|A \cap B|}{|A \cup B|} \tag{1}$$

For each garment part, the IoU [10] values were calculated over the entire test dataset, and their mean was obtained. The results are presented in Table IV, V, and VI. The number of test samples for each part was 51 for collars, 42 for sleeves, and 42 for bodices.



(a) Input (b) Output (c) Groundtruth

Fig. 4: Collar pattern generation results 1 (sailor collar)

(a) Input (b) Output (c) Groundtruth

Fig. 5: Collar pattern generation results 2 (flat collar)

Among the parts, collars achieved the highest mean IoU [10] (0.760), followed by sleeves (0.678) and bodices (0.552). Collars and sleeves generally have more constrained shapes and regions, making it easier for the model to generate sewing patterns that reflect the input garment shapes. In contrast, bodices exhibit greater shape diversity and are more affected by decorative elements, leading to a reduction in shape reconstruction accuracy. These tendencies and their underlying causes will be discussed in detail in the following section.

E. Qualitative Evaluation

In this section, we compare the conventional Single-generator Model with the proposed Part-specific Multi-generator Model to verify the effectiveness of the proposed approach.

The qualitative evaluation was conducted on three garment parts: bodice, sleeves, and collar, using the following two criteria:

1) Recognizability: This metric evaluates whether the generated pattern possesses an appropriate shape as the intended garment part, such that an experienced garment maker can immediately identify which part it represents. If the shape is ambiguous or does not form a valid pattern piece, recognizability is judged to be low.

2) Shape Accuracy: For parts judged recognizable, the generated pattern was compared with the ground-truth sewing pattern corresponding to the input garment. The evaluation focused on whether the contour shape, curvature, and proportional balance were appropriately reproduced. An experienced garment maker assessed how well the characteristic features of the input garment were preserved.

The evaluation results are summarized in Table VII and Table VIII.

From Table VII, the Single-generator Model shows moderate recognizability for the bodice and sleeves, but fails to

(a) Input (b) Output (c) Groundtruth

Fig. 6: Sleeve pattern generation results

(a) Input (b) Output (c) Groundtruth

Fig. 7: Bodice pattern generation results

generate a valid collar shape. In contrast, the Part-specific Multi-generator Model achieves 100% recognizability for all parts, indicating that separating the garment into components improves the stability of part-wise shape generation, particularly for parts with distinctive geometric structures such as collars.

In Table VIII, the Single-generator Model could not be evaluated for collar shape accuracy because the generated collar was not recognizable as a valid pattern. The Part-specific Multi-generator Model, however, successfully reproduces the collar shape with a shape accuracy of 72.5%. Shape accuracy for sleeves is also improved, suggesting that local contour and length relationships are better preserved when each part is modeled independently.

From Table VII and Table VIII, it can be observed that the single-generator model fails to sufficiently learn the local geometric characteristics of each part, and consequently tends to collapse when generating components with strong shape constraints, such as collars and sleeves.In contrast, the proposed method employs independent generators for each garment part, allowing the model to effectively learn local shape characteristics. As a result, it can produce stable and sewing-ready pattern shapes.

Next, we visually evaluate and compare the pattern shapes generated by each method. Fig.8 shows the input garment image, the patterns generated by each model, and the ground-truth sewing patterns.(a) is the garment image used as the input.(b) shows the ground-truth sewing patterns corresponding to (a).(c) is the pattern generated by the conventional single-generator model.(d) shows the patterns generated by the proposed part-specific model, in which the independently generated parts are combined into a single image.

As shown in Fig.8, the proposed method is visually confirmed to be effective in reproducing the shapes of individual garment parts. First, the single-generator model (c) failed to

TABLE IV: IoU RESULTS FOR COLLAR TYPES

Collar Type	Mean IoU
Sailor collar	0.818
Shirt collar	0.738
Shirt collar with collar back	0.824
Stand-up collar	0.630
Flat collar	0.789
Overall Average	**0.760**

TABLE V: IoU RESULTS FOR SLEEVE TYPES

Sleeve Type	Mean IoU
Short sleeve	0.660
Long sleeve	0.697
Overall Average	**0.678**

generate a valid collar shape, whereas the proposed model (d) successfully reconstructed a recognizable collar pattern. In addition, the sleeves generated by the single-generator model exhibited distortion and lost bilateral symmetry, while the proposed model generated two sleeves with stable and symmetric shapes.

On the other hand, for the bodice, the skirt shape produced by the proposed model (d) differs from the ground-truth pattern (b), and the single-generator model (c) generated a shape closer to the correct pattern.

F. Discussion

1) Garment Segmentation: In the proposed method, the garment images were successfully segmented into three parts—bodice, sleeve, and collar—with sufficient accuracy for use as inputs to the sewing pattern generation models. However, the segmentation process required additional steps such as creating auxiliary images and manually setting seed points, which introduced extra preprocessing effort. Future work should focus on developing a more automated and efficient part extraction method that maintains segmentation accuracy while reducing manual intervention.

2) Sewing Pattern Generation Models: According to the IoU [10] evaluation results, the collar and sleeve achieved relatively high shape consistency, whereas the bodice exhibited lower values compared to the other two parts.

The high performance of the collar and sleeve models can be attributed to the relatively limited spatial extent of these parts and their lower susceptibility to shape variation caused by the wearer's posture. These characteristics make it easier for the generation models to learn and reproduce edge and contour features accurately. In particular, for the collar, prior fine-grained classification into five distinct shape types, followed by training dedicated models for each type, is considered to have contributed to improved shape fidelity.

In contrast, the lower IoU [10] for the bodice may be explained by the following factors:

- High shape variability : The bodice is heavily influenced by the wearer's body shape and posture, resulting in greater variability compared to other parts. This diversity may have made it more difficult for the model to learn consistent shape features.

TABLE VI: IoU RESULTS FOR BODICE TYPES

Bodice Type	Mean IoU
Bodice skirt	0.649
Bodice dress	0.455
Overall Average	**0.552**

TABLE VII: RECOGNIZABILITY COMPARISON

Method	Bodice [%]	Sleeves [%]	Collar [%]
Single-generator model	81.8	80.0	0.0
Proposed model	**71.4**	**100**	**100**

- Large area and high information density : The bodice occupies the largest area of a garment and often contains numerous non-essential visual elements such as decorative patterns, making it more challenging to extract accurate shape features and thus reducing generation accuracy.

As potential improvements, preprocessing techniques that remove internal patterns or decorations and reconstruct occluded regions could be introduced. Furthermore, considering that garments inherently possess a three-dimensional structure, reconstructing the 3D garment shape from the input image prior to pattern generation is a promising direction. Such approaches could enhance the clarity of structural features, thereby enabling more accurate and practically applicable sewing pattern generation.

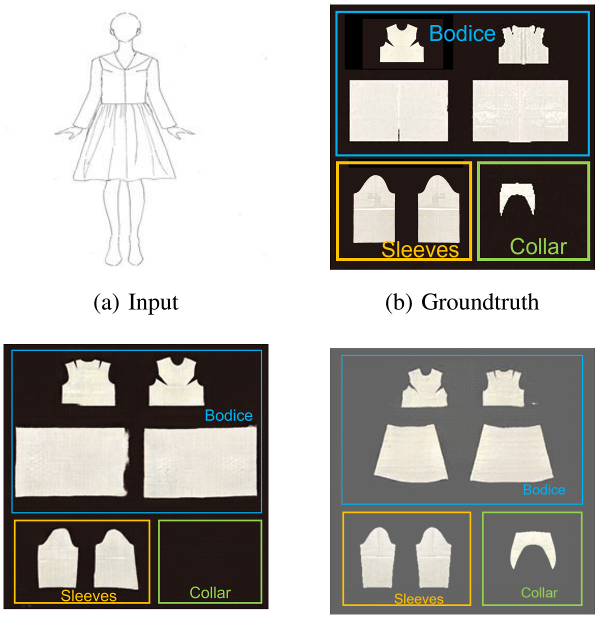
IV. CONCLUSION

In this study, we proposed a method for automatically generating sewing patterns corresponding to the bodice, sleeves, and collar from garment images. First, by employing SAM2 [6] in combination with part segmentation mask images, we performed part segmentation and preprocessing to create input data suitable for sewing pattern generation. Subsequently, by constructing a dedicated generation model for each part, we demonstrated that the proposed approach can produce sewing patterns that more faithfully reflect the shapes of smaller parts—something that conventional methods, which learn all garment parts jointly, find it difficult to achieve. Experimental evaluation showed that the proposed method achieved high IoU [10] values for the collar and sleeves, whose shapes tend to be relatively consistent. In contrast, for the bodice—whose shape is more affected by the wearer's posture, body type, and occlusion—a decrease in generation accuracy was observed.

For future work, to improve shape reproducibility across all parts, we plan to incorporate image processing techniques that remove patterns or decorations within the part regions and compensate for occlusions. Moreover, we will investigate methods that reconstruct the garment's three-dimensional structure prior to predicting its sewing patterns. While existing studies have attempted to reconstruct garment 3D structures for sewing pattern generation, they have not yet addressed the challenges of handling diverse shapes or producing sewing patterns suitable for practical use. Therefore, our future goal is to develop a method that overcomes these limitations, enabling the generation of sewing patterns that can adapt to a wide variety of shapes and be directly applied in real garment production.

TABLE VIII: SHAPE ACCURACY COMPARISON

Method	Bodice [%]	Sleeves [%]	Collar [%]
Single-generator model	53/0	80.0	–
Proposed model	**42.9**	**54.8**	**72.5**



(a) Input (b) Groundtruth (c) Single-generator model (d) Proposed model

Fig. 8: Comparison of the single-generator model, Proposed model, and the ground-truth sewing patterns.

REFERENCES

[1] N. Kato, H. Osone, K. Oomori, C. W. Ooi, and Y. Ochiai, "GANs-based Clothes Design: Pattern Maker Is All You Need to Design Clothing," in Proceedings of the 10th Augmented Human International Conference 2019, Reims France: ACM, Mar. 2019.

[2] I. J. Goodfellow et al., "Generative Adversarial Networks," June 10, 2014, arXiv: arXiv:1406.2661.

[3] L. Liu, H. Xu, X. Pan, X. Wang, S. Bi, Z. Deng, and Z. Xu, "Towards garment sewing pattern reconstruction from a single image," *ACM Transactions on Graphics*, vol. 42, no. 4, pp. 1–13, 2023.

[4] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 30, 2017.

[5] H. Suzuki, K. Miura, A. Moro, S. Pathak, and K. Umeda, "Realistic sewing pattern generation and dataset construction for GAN-based fashion synthesis," in Proc. 2025 SICE Festival with Annual Conference (SICE FES 2025), 2025.

[6] N. Ravi et al., "SAM 2: Segment Anything in Images and Videos," Oct. 28, 2024, arXiv: arXiv:2408.00714.

[7] P. Isola, J. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2017, pp. 5967–5976.

[8] M. Mirza and S. Osindero, "Conditional Generative Adversarial Nets," Nov. 06, 2014, arXiv: arXiv:1411.1784.

[9] K. Wada, "labelme: Image Polygonal Annotation with Python," Zenodo, 2021.
Available: https://github.com/wkentaro/labelme

[10] H. Rezatofighi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid, and S. Savarese, "Generalized Intersection over Union: A Metric and A Loss for Bounding Box Regression," Apr. 15, 2019, arXiv: arXiv:1902.09630.