Studies on anime-illustration to body-outline translation

Rintaro Matsumura and Hiroyuki Arai. Nippon Institute of Technology

***Abstract*—In this study, we introduce a body-outline translation method using U-GAT-IT to assist in anime-illustration production. Our results suggest that focusing on instance correspondence during training effectively enhances translation performance.**

***Index Terms*—GAN,U-GAT-IT,Image-To-Image translation**

I. Introduction

There has been a surge in the number of amateur anime-illustrators. Additionally, given the challenges associated with creating anime-illustrations, various kinds of support services have emerged. anime-illustrators sometimes refer to character poses available online. However, as depicted in Fig. 1 on the left, when a character is adorned with intricate decorations or has a background, discerning the pose might become challenging. Although there are related methods that estimate character joints as seen in the center of Fig. 1, relying solely on joints may not suffice for reference. To address this issue, we propose a method to extract the body-outline, as demonstrated on the right in Fig. 1.

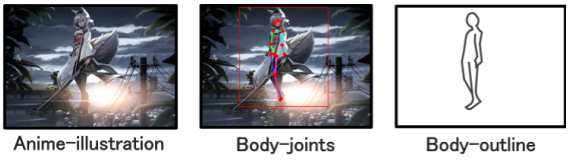

Fig. 1: Examples of body-outline translation

II. Proposed Method(s)

There is an application of GANs in the area of deep learning methods to translation one image into another. Among these, the related model U-GAT-IT [1] demonstrates superior translation capabilities, even for images that undergo significant shape changes, when compared to traditional image translation methods. The model excels at tasks that factor in shape, like translation images of cucumbers into bitter gourds. Therefore, it is anticipated to be effective in body-outline translation as well.

U-GAT-IT incorporates the ”ADALIN” method, which allows for a balance between texture and shape translation. However, for accurate balanced translation, we hypothesize that instances across two domains must be explicitly matched. For the translation task in this study, especially when limbs adhere closely to the front or side of the torso, recognizing instances becomes difficult, potentially leading to insufficient translation. Given these characteristics, we believe that a dataset explicitly designed with instance-level correspondence can facilitate more accurate translation. To this end, we prepared two types of body-outline images for our experiments, as illustrated in Fig. 2. Limbs-colored images are ”model A”, while uncolored ones are ”model B”. For the sake of foundational examination in this study, the dataset was created using 3DCG.

Fig. 2: Datasets

III. Findings and Discussion

To evaluate the translation results, a three-tiered qualitative evaluation criterion as shown in Table. 1 was established and assessed visually: (1) significant pose deviations or body fragmentation, (2) unnatural details, (3) No issues.

From Table. 1, model A shows a 13% reduction in critical defects compared to model B. In fact, Fig. 3 upper confirms model A's limb translation improvements, Demonstrating the Effectiveness of Instance Awareness. However, Fig. 3 lower shows model A struggles with limb overlap in intricate decorations. This is likely due to the failure to recognize limbs, resulting in a lack of instance correspondence. Two countermeasures were considered and tested. First, enlarging images improved model A's critical defects by 7%, as seen in Table. 1. Improvement is likely due to more pixels in limb area. Second, adding a Sub Class Detector (SCD) to model A increased satisfactory results by 4% (Table. 1). SCD's retention of decoration data likely aided instance segmentation.

TABLE 1: Qialitative assessment results

Evaluation Categories	Model A	Model B	Model A (Wide image)	Model B (with SCD)
(1) Critical Defects	23.8%	36.3%	16.5%	23.6%
(2) Room for Improvement	44.8%	41.6%	50.9%	41.2%
(3) Satisfactory	31.4%	22.1%	32.6%	35.4%

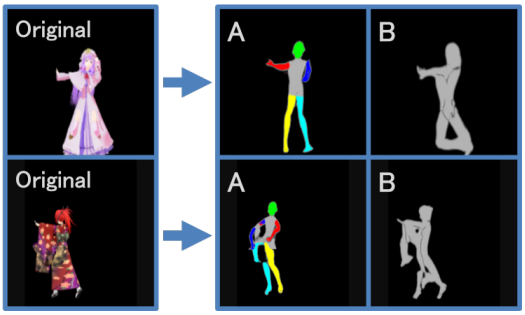

Fig. 3: Experimental results

IV. Conclusion

In this study, we aimed to explore body-outline translation using U-GAT-IT. The experiments with an instance-aware dataset demonstrated both effectiveness.

References

[1] Junhob Kim et.al. ”U-GAT-IT:Unsupervised Generative Attentional Networks with Adaptive Layer-Instance Normalization for Image-to-Image Translation”. arXiv:1907.10830, 07.25 2019.