Pacific Graphics 2021
E. Eisemann, K. Singh, and F.-L Zhang
(Guest Editors)

Volume 40 (2021), Number 7

Line Art Colorization Based on Explicit Region Segmentation –Supplemental Material–

Ruizhi Cao, Haoran Mo, and Chengying Gao†

Sun Yat-sen University

1. Translation between Region Map and Skeleton Map

We propose an explicit segmentation fusion mechanism, which utilizes regional segmentation information stored in skeleton maps to alleviate the color bleeding effects. The skeleton map firstly introduced in DanbooRegion dataset [ZJL20] is not a conventional (1D) skeleton, but a translation of a region map essentially. The translation is performed because skeleton maps are learnable and can be directly predicted by a neural network given a line art. On the contrary, the region maps are unlearnable and cannot be predicted because regions are unordered and represented by random colors, and thus L1 loss is meaningless.

Region to Skeleton. Given a region map of a line art image, the skeleton map is generated by: First, detect the edge of the annotated regions of the line art. Then, extract the skeleton of the regions. Afterwards, combine the region edge (as background) and region skeleton (as foreground). Finally, the skeleton map is produced by applying a smoothing operation to the combined image, which is similar to a distance transformation.

Skeleton to Region. Given a skeleton map, binarization operation is first employed to produce a watershed marker. We use a threshold of 0.549 for the binarization because it works well with our datasets. Next, we use the watershed algorithm [NP14] with the generated marker to obtain the region map.

2. Post-processing

Although segmentation information allows the colorization networks to generate better colorized results by reducing color bleeding artifacts, other kinds of artifacts may still be produced, such as unnatural motley or checkerboard artifacts. We propose an optional post-processing method to further reduce these kinds of artifacts in small areas.

The post-processing method is mainly based on the region maps. Our approach generates a colorized image $\hat{y}$ as well as a skeleton map $\hat{s}$. With the conversion introduced in Section 1, we obtain the corresponding region map $\hat{r}$ from $\hat{s}$. With the segmented regions and the output result $\hat{y}$, we fill each region with the median color sampled from all pixels inside that region. This step erases all the edges, and the result is a flat painting style image $\hat{f}$ without edges. To address this problem, we merge the input line art x and the flat image $\hat{f}$ by Eq.(1) to obtain the final result $\hat{o}$:

$$\hat{o}_{i,j} = \begin{cases} \hat{f}_{i,j} & x_{i,j} > 180 \\ (1-\alpha)\hat{f}_{i,j} + \alpha x_{i,j} & x_{i,j} \leq 180 \end{cases}, \tag{1}$$

where $\hat{o}_{i,j}$, $\hat{f}_{i,j}$ and $x_{i,j}$ are the RGB values in the (i,j) pixel in $\hat{o}$, $\hat{f}$, and x, respectively. α is a scalar and set to 0.4.

Results of applying post-processing are shown in Figure 1.

3. More Results

3.1. Effectiveness of Avoiding Color Bleeding

Figure 2 shows the results of avoiding color bleeding and improvement on color contrast.

3.2. Performance of Fusion Mode

Figure 3 shows the results of the two fusion modes of segmentation information.

In Figure 4, we show some examples where the “Direct Concatenation” mode works better than the “Dual-branch” mode. For example, in the top row example, the color of the neck from “Dual-branch” mode is weird. In the example from the second row, there exists inconsistent color in the tiptoe with “Dual-branch” fusion mode.

3.3. Reference-based Colorization

Figure 5 and Figure 6 show the results of reference-based colorization compared with MUNIT [HLBK18].

References

[HLBK18] HUANG X., LIU M.-Y., BELONGIE S., KAUTZ J.: Multimodal unsupervised image-to-image translation. In *Proceedings of the European conference on computer vision (ECCV)* (2018), pp. 172–189.

[KJPY19] KIM H., JHOO H. Y., PARK E., YOO S.: Tag2pix: Line art colorization using text tag with secat and changing loss. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (2019), pp. 9056–9065.

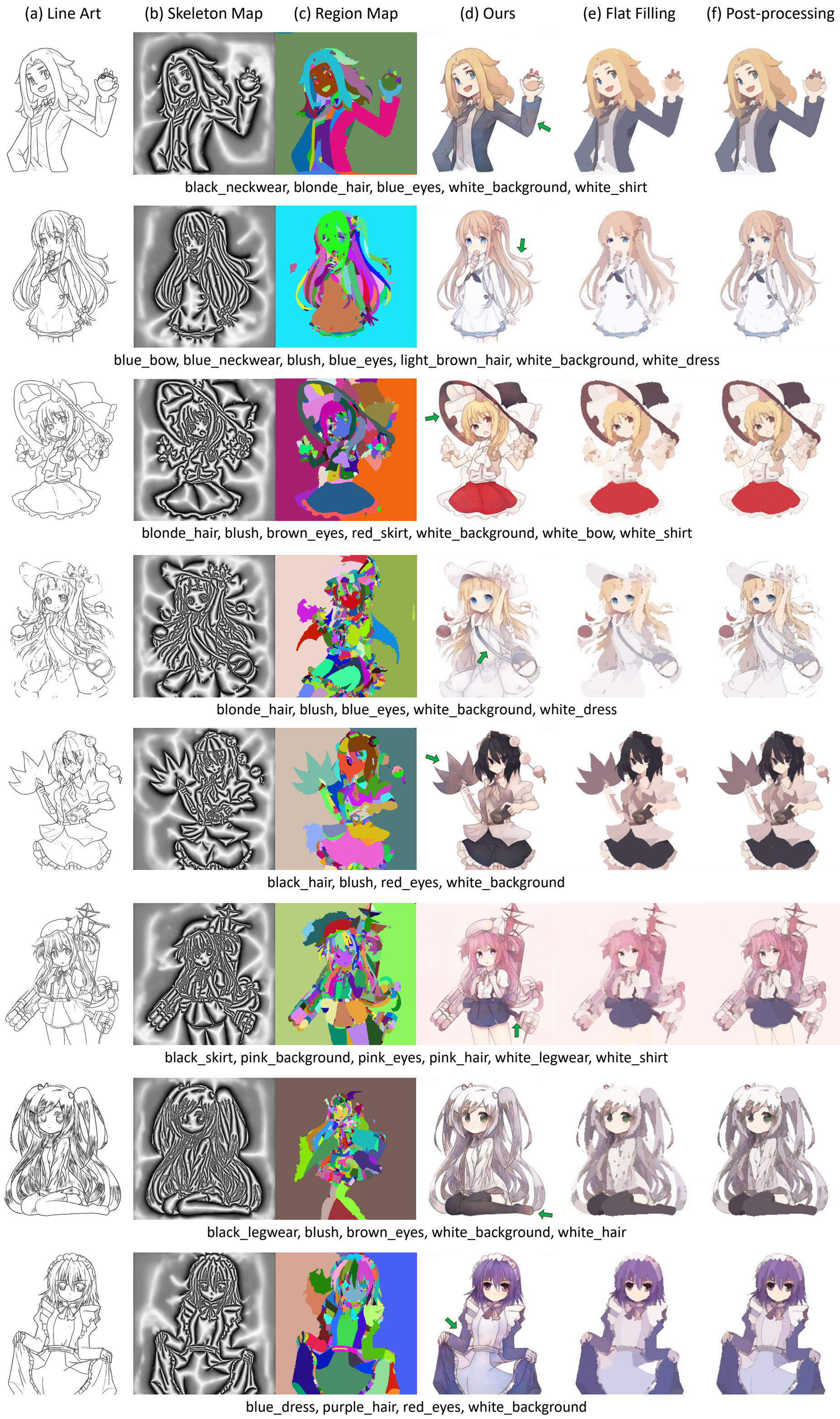

Figure 1: *Post-processing of alleviating artifacts based on region maps. Flat painting style (e) and line art (a) added result (f).*

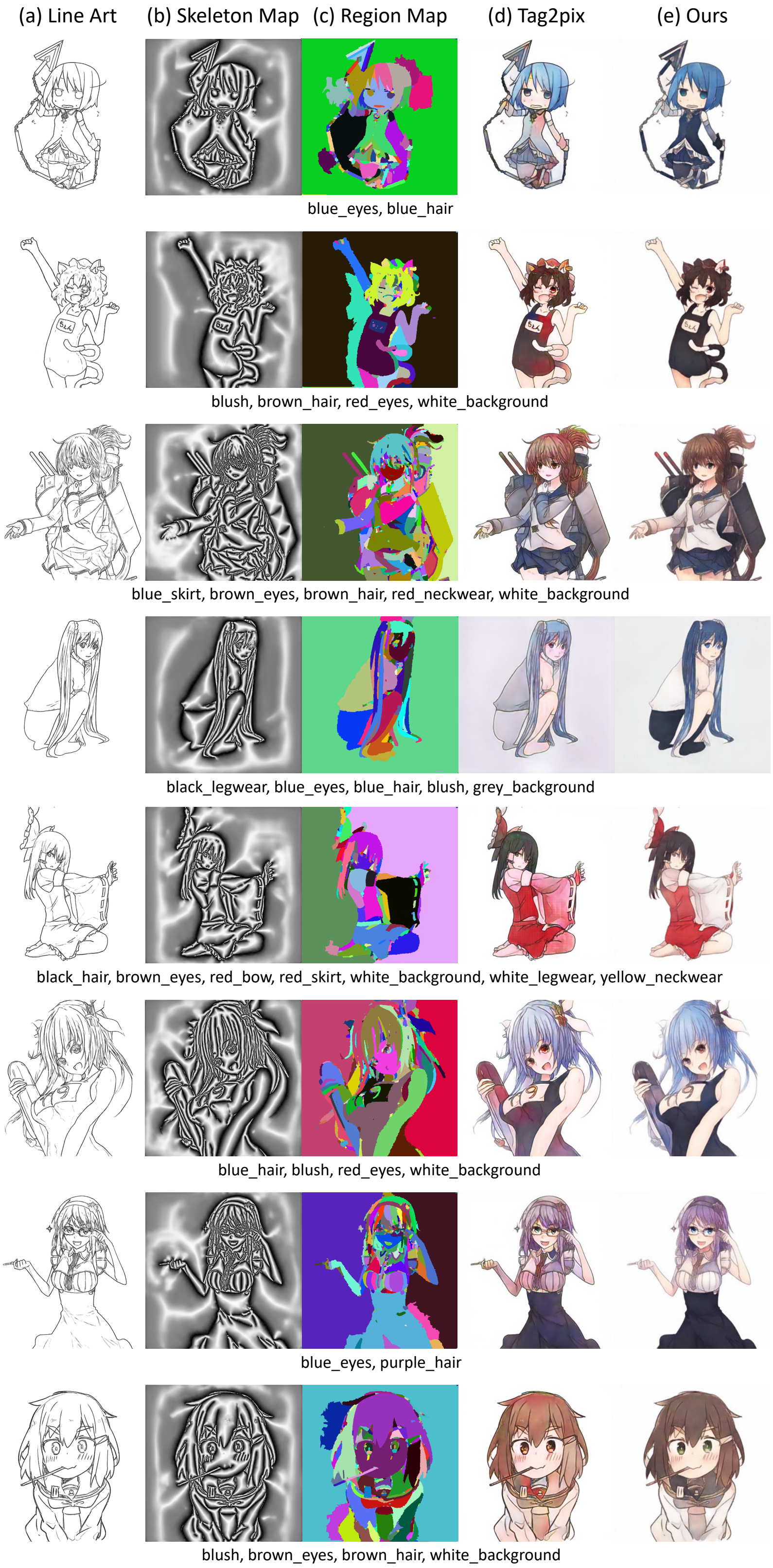

Figure 2: *Effectiveness of alleviating color bleeding and improving color contrast compared with Tag2pix [KJPY19]. Our results are from the framework with a dual branch.*

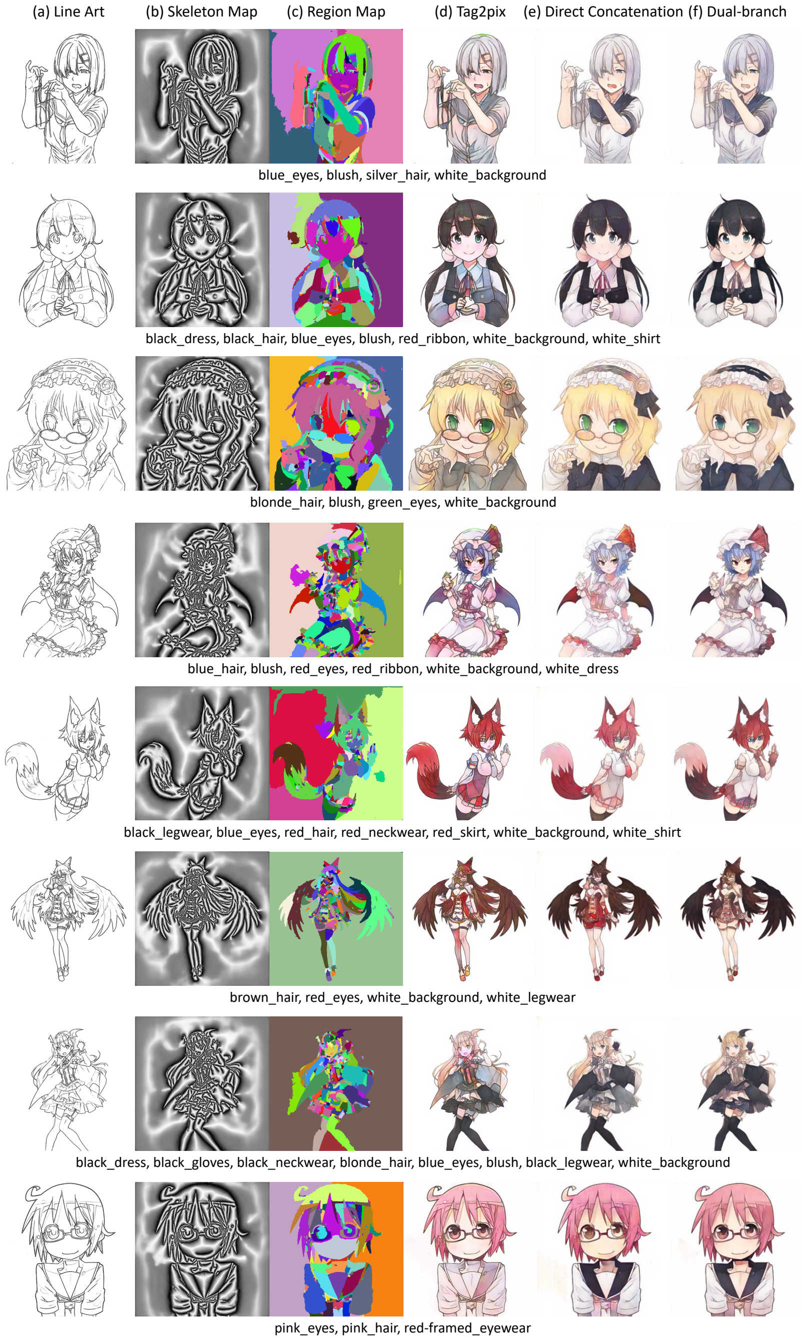

Figure 3: *Effectiveness of the two fusion modes of the segmentation information.*

Figure 4: *Examples where "Direct Concatenation" mode ("Concat") works better than "Dual-branch".*

Figure 5: *Results on reference-based line art colorization. Our results are from model by incorporating MUNIT [HLBK18] with explicit segmentation information in a direct concatenation mode.*

Figure 6: *Results on reference-based line art colorization. Our results are from model by incorporating MUNIT [HLBK18] with explicit segmentation information in a direct concatenation mode.*

[NP14] NEUBERT P., PROTZEL P.: Compact watershed and preemptive slic: On improving trade-offs of superpixel segmentation algorithms. In *2014 22nd international conference on pattern recognition* (2014), IEEE, pp. 996–1001.

[ZJL20] ZHANG L., JI Y., LIU C.: Danbooregion: An illustration region dataset. In *European Conference on Computer Vision (ECCV)* (2020).