Late-resizing: A Simple but Effective Sketch Extraction Strategy for Improving Generalization of Line-art Colorization

Appendix & Supplementary Material

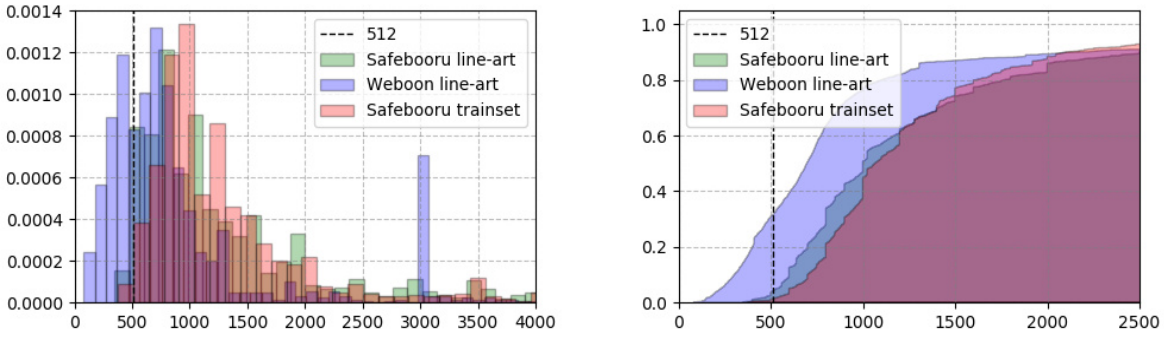
1. Appendix

1.1. Appendix-A: Distribution of resolutions

To support the assumption presented in Section 3.1 of the main text that we consider the scale normalization only for reducing the size, we analyze the distribution of the maximum axis size for the source color images of the training dataset. As shown by the red area in Fig. 1 (b), more than 90% of the images in the training dataset had an axis size of 512 or larger. The Safebooru line-art dataset, which is one of the representative target sketch dataset, showed the similar distribution, as shown in the green area in Fig. 1. Therefore, the target sketch also requires scale normalization to reduce size of image and this normalization causes inevitable downsizing artifacts, as described in Section 3.2.2 of our main text. The area marked in blue in Fig. 1 shows the distribution for Webtoon line-art dataset, where more than 70% of sketches had a larger axis size than 512, even though this dataset contains partial contents such as hands and faces.

1.2. Appendix-B: Comparison of mixing different types of sketch extraction methods

To show that the proposed late-resizing strategy is effective even in the case of mixing different types of several sketch extraction methods, we supplement an ablation study based on the hint-based coloring method very recently proposed by Yuan and Simo-Sera [2]. From a model structural point of view, Yuan and Simmo-sera [2] proposed a concatenated spatial attention block (CSA) inspired by SECat module proposed in Tag2pix [1]. In terms of the sketch extraction method, sketchKeras and XDoG line detector were used in combination similar to Tag2Pix [1]. For comparison, we applied a late resizing strategy instead of an early resizing strategy to XDoG. Note that we preserved strategy for sketchKeras in the same way due to the limitation of its input resolution (512×512). Furthermore, we compared these results with the case of using only late-resized XDoG without the other method, i.e., sketchKeras. Table 1 summarizes the settings used for comparison and Figs. 2, 3, 4 and 5 show the qualitative results of the comparison. As shown in (a) and (b) of Figs. 2, 3, 4 and 5, the case of applying late-resizing showed better color expression and less color bleeding for all kinds of sketches, despite the small percentage of



(a) Histogram of distribution (b) Histogram of cumulative distribution function

Figure 1. Histogram of the distribution of resolutions. The horizontal axis represents the maximum axis size of images, and the vertical axis represents the density.

Table 1. **Experimental settings.**

Method	Model	Sketch extraction pipeline			
		Pre-resize	Line detection	Binarization	Post-resize
(a) Yuan [2]	UNeT + CSA	512×512	sketchKeras + XDoG(9:1)	simplification & tanh	-
(b) Yuan [2]-ours	UNeT + CSA	512×512 -	sketchKeras (90%) XDoG (10%)	simplification tanh	- 512×512
(c) Yuan [2]-ours only XDoG	UNeT + CSA	-	XDoG (100%)	tanh	512×512

XDoG. Interestingly, the case of using late-resized XDoG alone showed the best results. This results indicates that the late-resizing strategy can effectively improve the generalization of the colorization model and training data to which downsizing augmentation is not applied has a potential risk.

2. Additional Result

We present additional results for the visual comparison performed with AlacGAN baseline and that applied proposed late-resizing strategy, as shown in Figs. 6, 7, 8 and 9. It is noteworthy that the line segment remained sharp when using late-resized XDoG. We found that a gap of intensity between the line segment of binarized sketch and that of ground truth makes the colorization model drive the line segment brighter or darker. Here, we analyzed that the downsizing augmentation, which is one of the benefit of proposed late-resizing strategy, resolves the gap of intensity, as described in Section 3.2.2 of the main text.

References

[1] Hyunsu Kim, Ho Young Jhoo, Eunhyeok Park, and Sungjoo Yoo. Tag2pix: Line art colorization using text tag with secat and changing loss. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 9056–9065, 2019.

[2] Mingcheng Yuan and Edgar Simo-Serra. Line art colorization with concatenated spatial attention. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3946–3950, 2021.

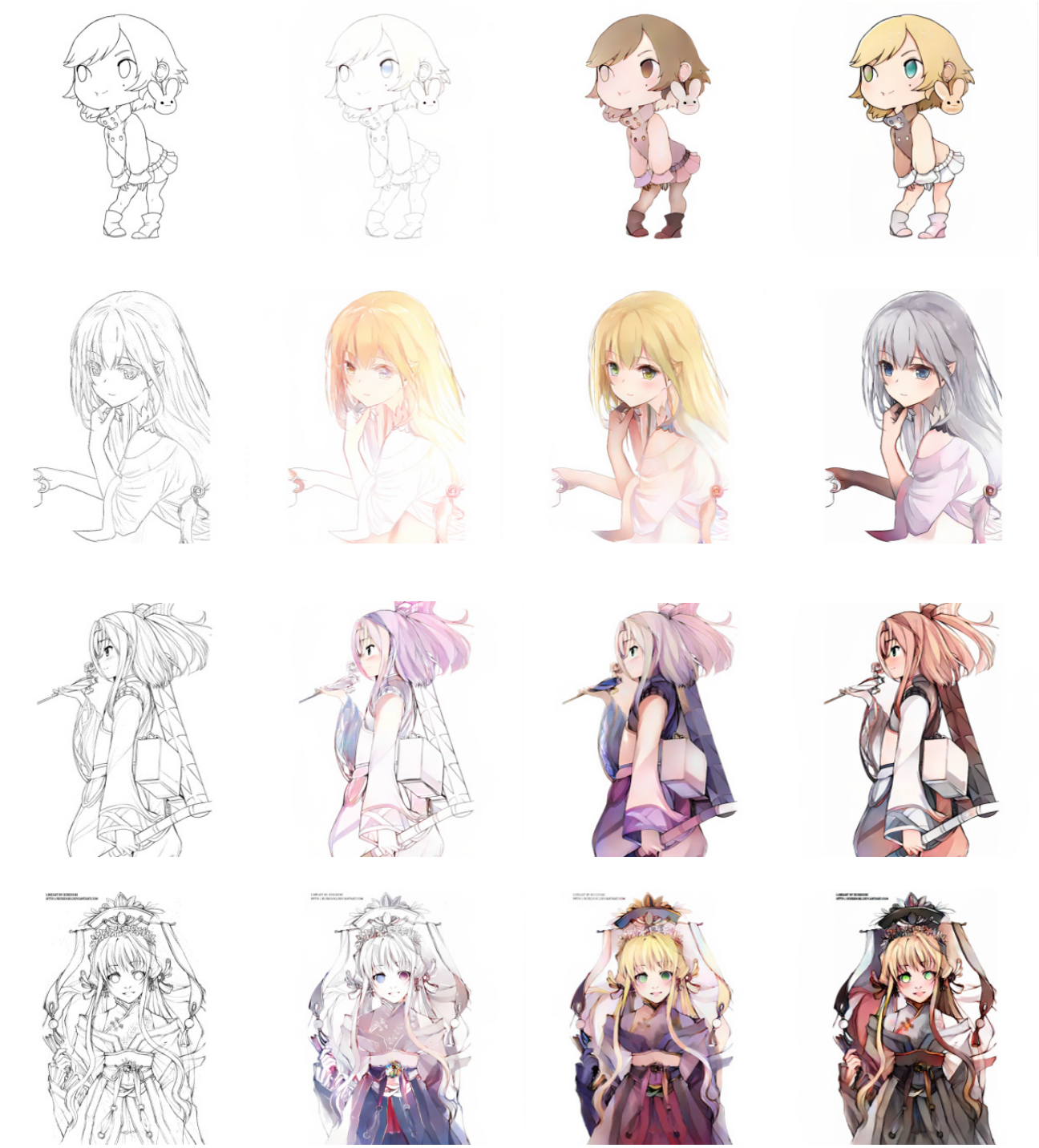
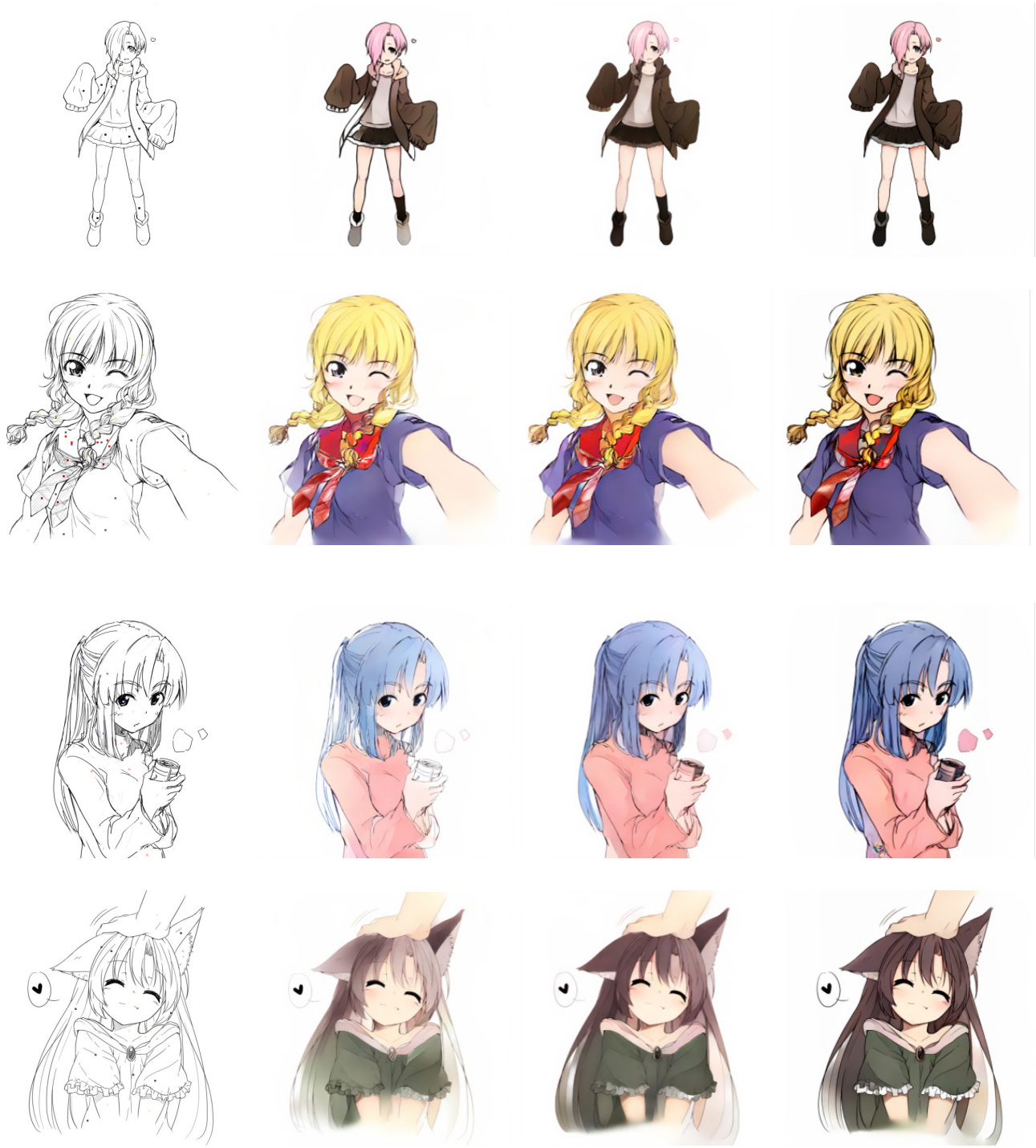


Figure 2. Visual comparison according to the application of the proposed method for Safebooru line-arts without color hints. (a) shows line-arts and (b)-(d) show colorized results using original setting, late-resized XDoG with other line detection methods in combination and late-resized XDoG solely, respectively.

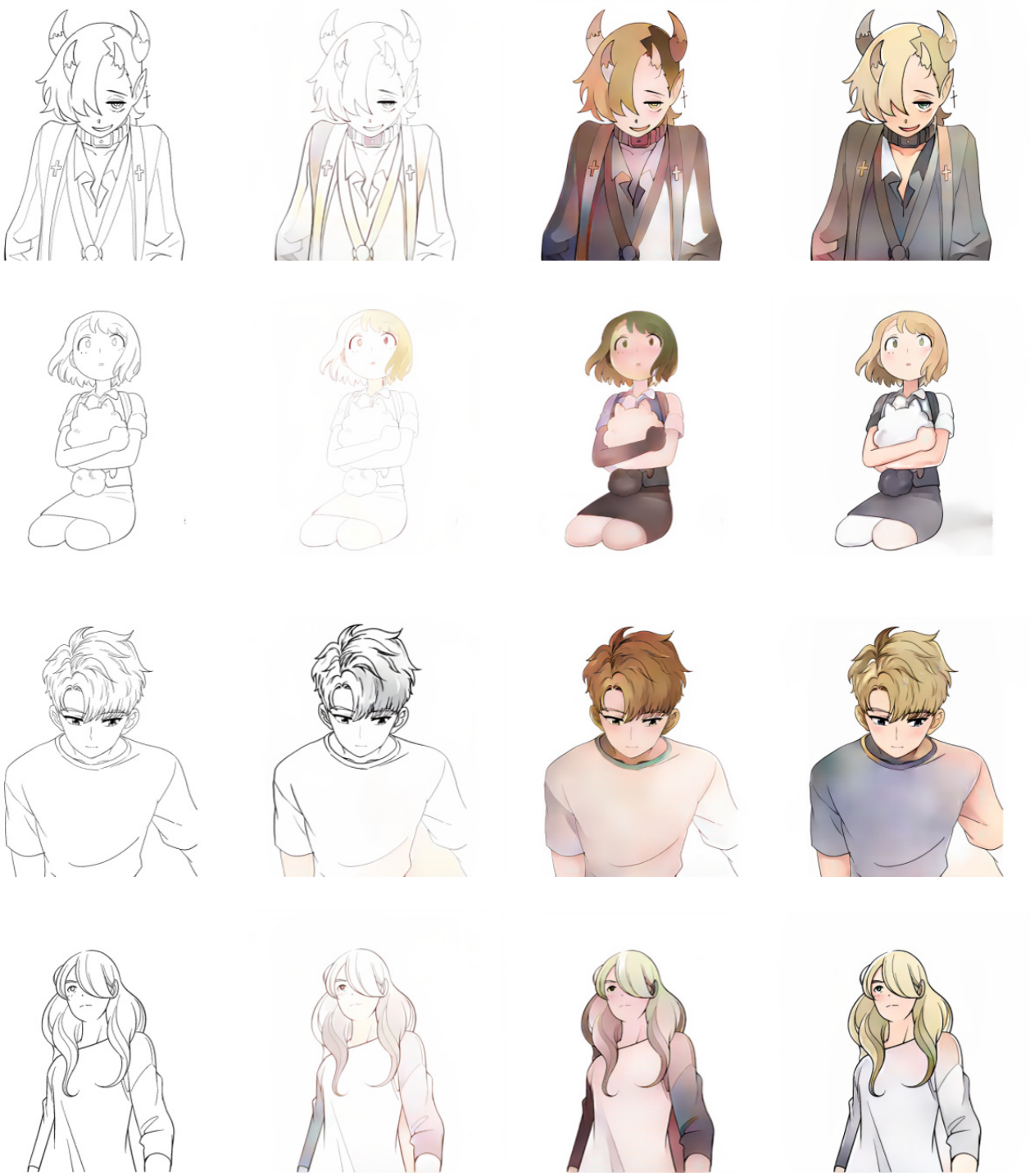
Colorized results for **Safebooru** line-arts **with** color hints **(Yuan)**



(a) line-art (b) Yuan (c) Yuan-ours (d) Yuan-ours only XDoG

Figure 3. Visual comparison according to the application of the proposed method for Safebooru line-arts with color hints. (a) shows line-arts and (b)-(d) show colorized results using original setting, late-resized XDoG with other line detection methods in combination and late-resized XDoG solely, respectively.

Colorized results for **Webtoon** line-arts **without** color hints **(Yuan)**



(a) line-art (b) Yuan (c) Yuan-ours (d) Yuan-ours only XDoG

Figure 4. Visual comparison according to the application of the proposed method for Webtoon line-arts without color hints. (a) shows line-arts and (b)-(d) show colorized results using original setting, late-resized XDoG with other line detection methods in combination and late-resized XDoG solely, respectively.

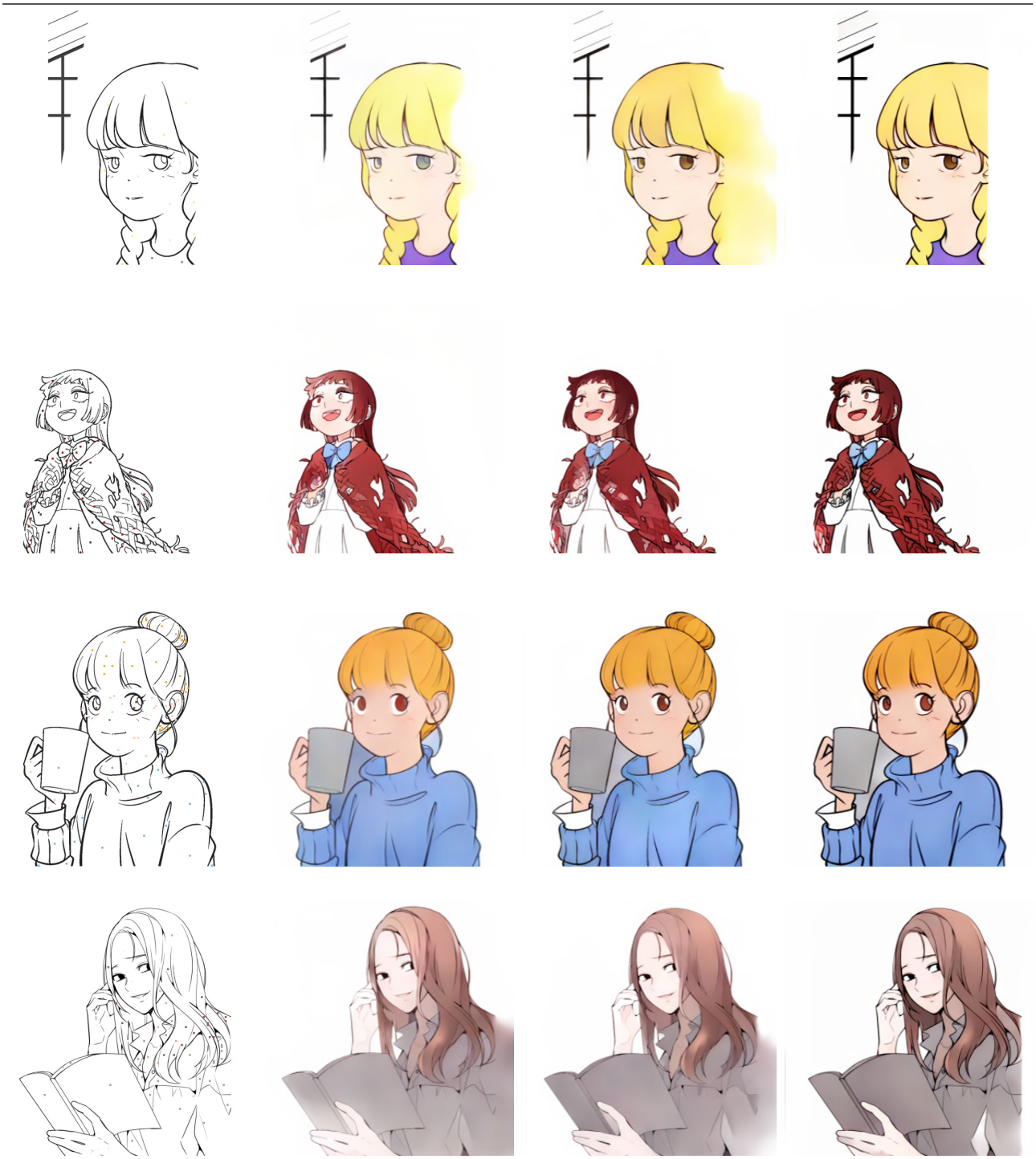
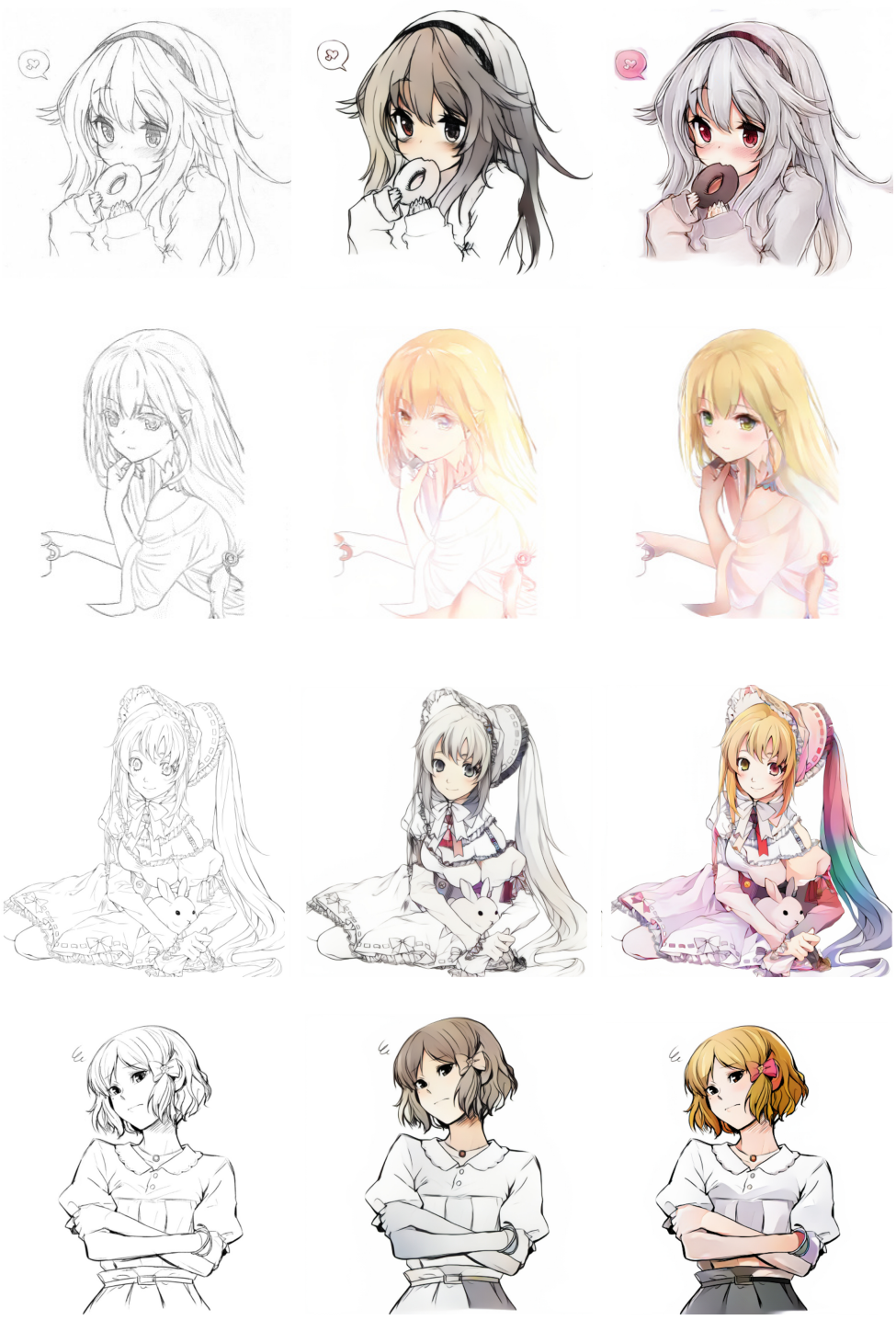


Figure 5. Visual comparison according to the application of the proposed method for Webtoon line-arts with color hints. (a) shows line-arts and (b)-(d) show colorized results using original setting, late-resized XDoG with other line detection methods in combination and late-resized XDoG solely, respectively.

Colorized results for **Safebooru** line-arts **without** color hints **(AlacGAN)**



(a) line-art (b) Alac (c) Alac-ours

Figure 6. Visual comparison according to the application of the proposed method for Safebooru line-arts without color hints. (a) shows line-arts and (b)-(c) show colorized results using original setting and late-resized XDoG with other line detection methods in combination, respectively.

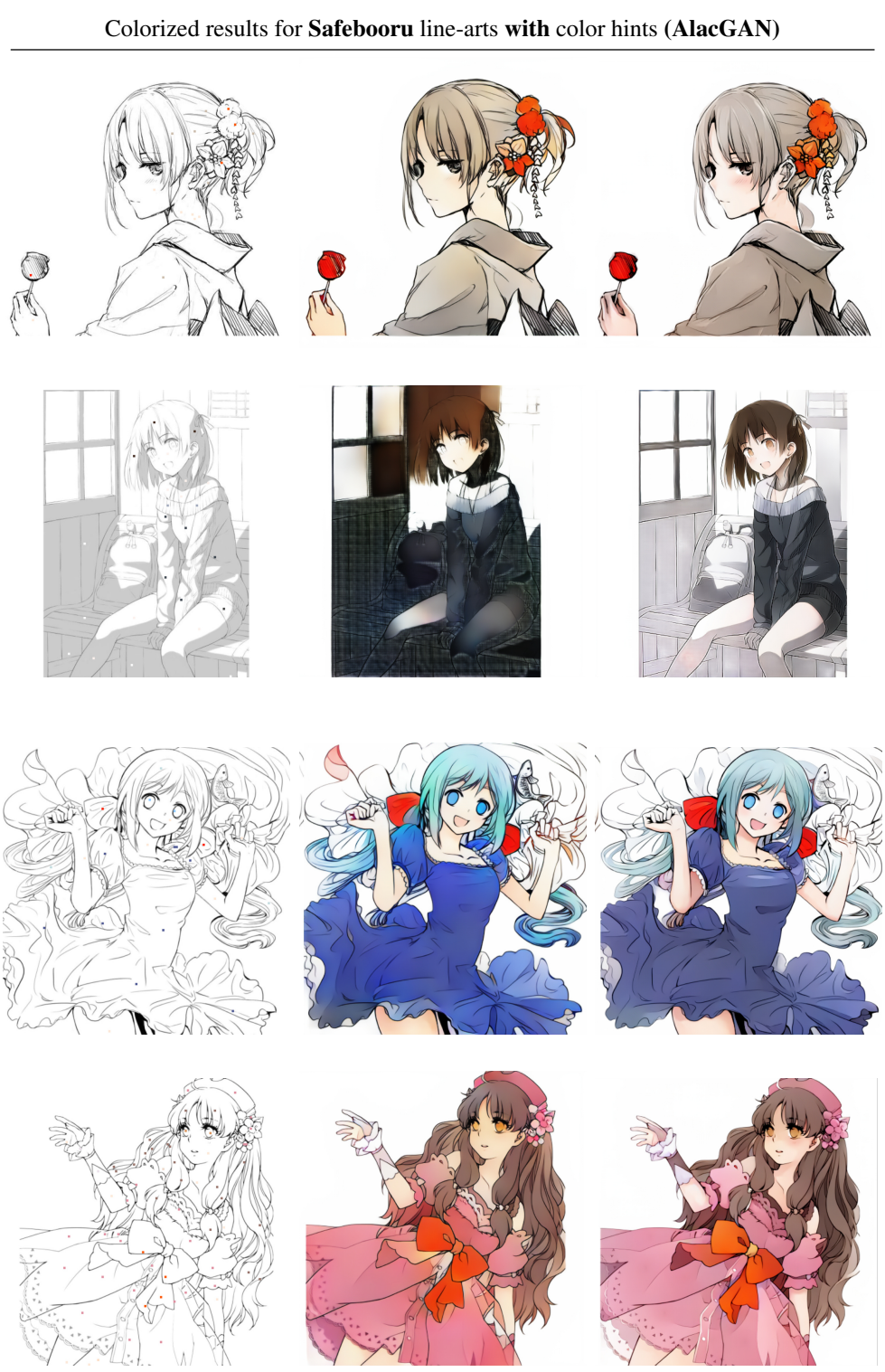
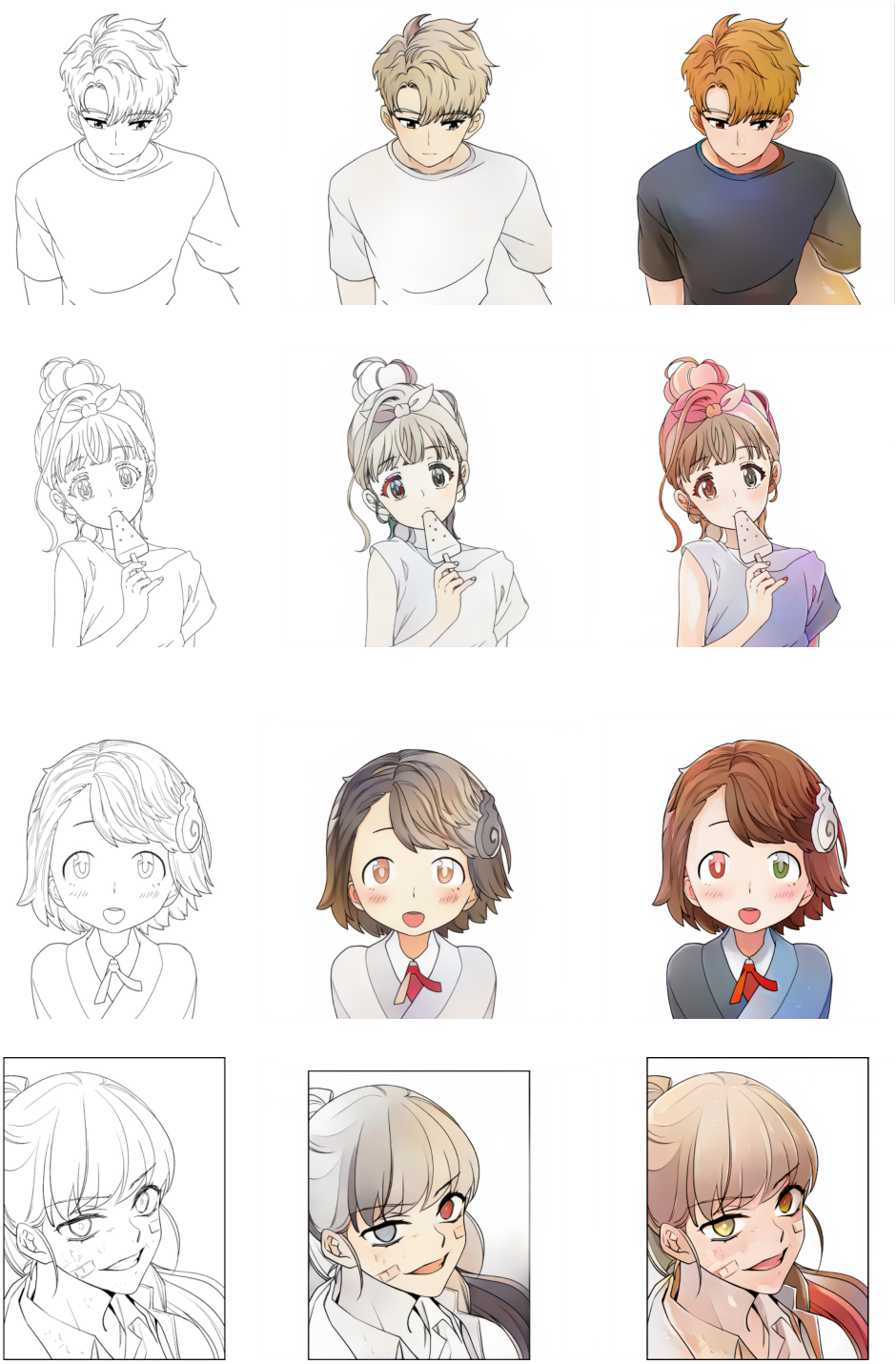


Figure 7. Visual comparison according to the application of the proposed method for Safebooru line-arts with color hints. (a) shows line-arts and (b)-(c) show colorized results using original setting and late-resized XDoG with other line detection methods in combination, respectively.

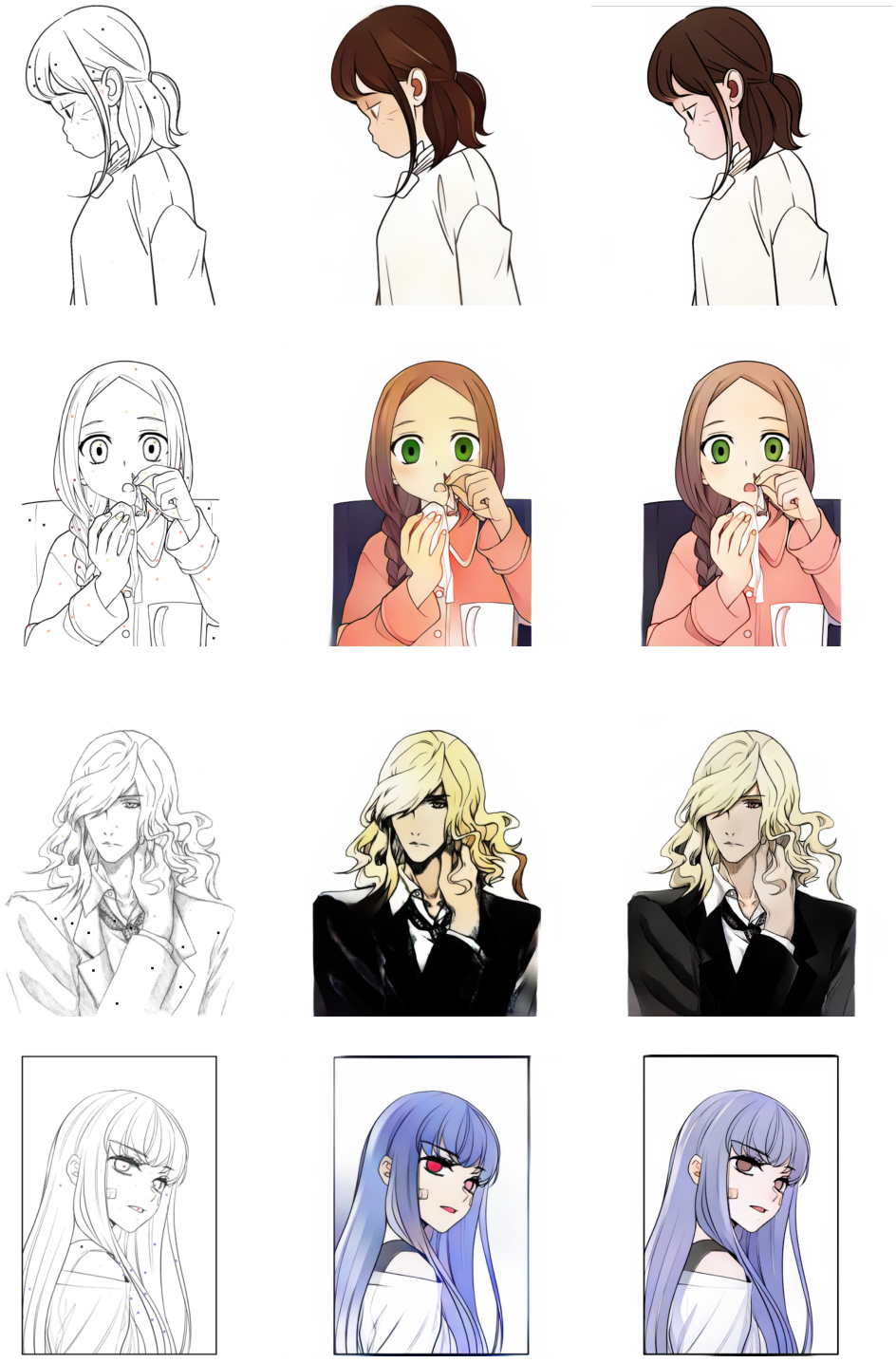
Colorized results for **Webtoon** line-arts **without** color hints **(AlacGAN)**



(a) line-art (b) Alac (c) Alac-ours

Figure 8. Visual comparison according to the application of the proposed method for Safebooru line-arts without color hints. (a) shows line-arts and (b)-(c) show colorized results using original setting and late-resized XDoG with other line detection methods in combination, respectively.

Colorized results for **Webtoon** line-arts **with** color hints **(AlacGAN)**



(a) line-art (b) Alac (c) Alac-ours

Figure 9. Visual comparison according to the application of the proposed method for Safebooru line-arts with color hints. (a) shows line-arts and (b)-(c) show colorized results using original setting and late-resized XDoG with other line detection methods in combination, respectively.