Confidence-aware Practical Anime-style Colorization

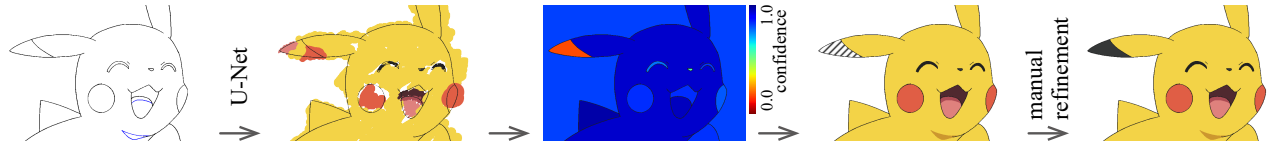
Daichi Ishii
ishii.daichi.ia8@is.naist.jp
NAIST
Ikoma, Nara, Japan

Hiroyuki Kubo
hkubo@tsc.u-tokai.ac.jp
Tokai University
Minato-ku, Tokyo, Japan

Seitaro Shinagawa
shinagawa.seitaro.si8@is.naist.jp
NAIST
Ikoma, Nara, Japan

Akinobu Maejima
akinobu.maejima@olm.co.jp
OLM Digital, Inc.
IMAGICA GROUP Inc.
Setagaya-ku, Tokyo, Japan

Takuya Funatomi
Satoshi Nakamura
Yasuhiro Mukaigawa
NAIST
Ikoma, Nara, Japan



(a) line-drawing (input) (b) per-pixel semantic label (c) per-region confidence (d) strategic withdrawal (e) colorized image (output)

Figure 1: Illustration of the proposed pipeline. Given an input line-drawing (a), U-net[Ronneberger et al. 2015] predicts the per-pixel semantic labels (b). Because the prediction of the right ear edge is inconsistent, the confidence of the region is low (c). Low confident regions are filled in using a noticeable slash line pattern (d) to be colorized manually. After a tiny manual refinement, the complete image can be obtained (e). Images used with permission. ©Nintendo, Creatures, GAME FREAK, TV Tokyo, ShoPro, JR Kikaku, ©Pokémon

ABSTRACT

We propose practical anime-style colorization of an input line-drawing. The key idea is the strategic withdrawal which reflects the prediction confidence that indicates the expected accuracy of the predicted color labels. Furthermore, we investigate the relation between the proposed confidence, prediction accuracy, and number of automatically colorized regions to maximize the efficiency of the colorization process including both automatic prediction and manual correction for practical use in production.

CCS CONCEPTS

• **Computing methodologies → Image manipulation**.

KEYWORDS

anime-style, automatic colorization, production

ACM Reference Format:
Daichi Ishii, Hiroyuki Kubo, Seitaro Shinagawa, Akinobu Maejima, Takuya Funatomi, Satoshi Nakamura, and Yasuhiro Mukaigawa. 2020. Confidence-aware Practical Anime-style Colorization. In *Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY*. ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3388767.3407331

SIGGRAPH '20 Talks, August 17, 2020, Virtual Event, USA

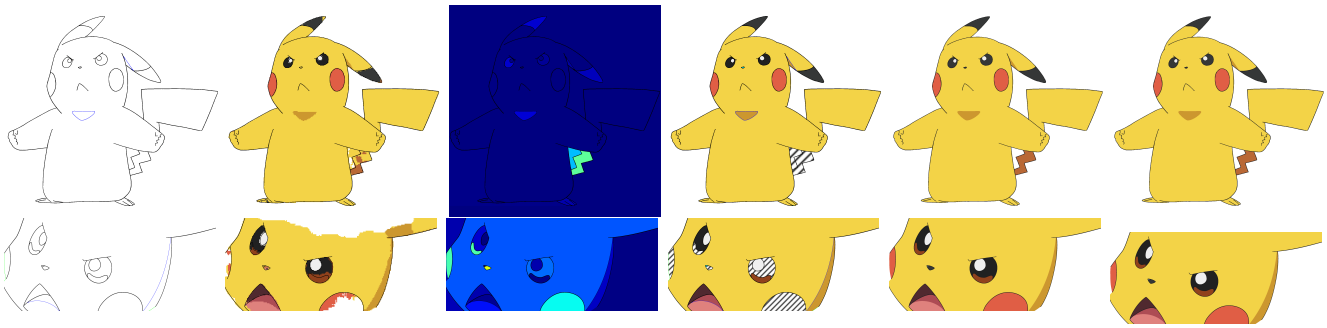
ACM ISBN 978-1-4503-7971-7/20/08.
https://doi.org/10.1145/3388767.3407331

1 INTRODUCTION

Colorizing a draft line drawing of anime characters is a time-consuming process in a production pipeline, because artists need to colorize numerous images manually. To automate this process, we developed fully convolutional network (FCN)-based semantic segmentation to predict the color labels of the corresponding parts of a character [Ramassamy et al. 2018]. The accuracy of the above prediction is relatively good; however, there remain several incorrect colorized regions. To correct the colors, it is necessary to address two steps: *1. distinguish incorrect colorized regions from the "mostly" correct colorized image*, and then *2. fill in the correct color*. According to interviews with professional artists in production, the first step is more troublesome. More precisely, they prefer the region to remain uncolorized rather than colorized in the wrong color. To avoid colorizing in incorrect colors as much as possible, we propose confidence-aware colorization that withdraws automatic colorization according to the prediction confidence. This system only colorizes the region of higher confident, and also fills the region using a noticeable pattern to be colorized manually in lower confident region.

2 CONFIDENCE-AWARE COLORIZATION

Ramassamy et al. mentioned [Ramassamy et al. 2018] that the anime-style colorization problem can be regarded as a semantic segmentation task, which predicts the per-region label. We basically follow the same strategy in their work, but replace the FCN with U-net [Ronneberger et al. 2015] to improve the prediction accuracy and introduce a withdrawal strategy based on the prediction confidence to reduce the cost for fixing incomplete regions. Figure 1 illustrates the pipeline for our method.



(a) input line-drawing (b) estimated per-pixel labels (c) per-region confidence (d) strategic withdrawal (e) manual refinement result (f) ground truth

Figure 2: Step-wise images in our pipeline. Applying the U-net to the input line-drawing (a), per-pixel labels are obtained (b). Determined by the proposed confidence (c), strategic withdrawal regions which confidences lower than $t_h = 0.8$ are filled in slash line pattern (d) to be colorized manually in (e). Note, higher-lower confidence shows in blue-red, respectively in (c). (f) shows the ground truth. Images used with permission. ©Nintendo, Creatures, GAME FREAK, TV Tokyo, ShoPro, JR Kikaku, ©Pokémon

2.1 Per-region Semantic Label Estimation

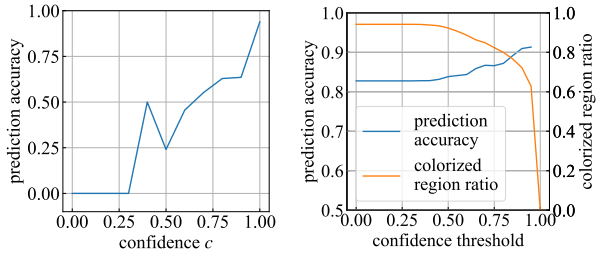
We collected approximately 1,000 image pairs (line-drawings and corresponding manually colorized images) of the yellow character shown in Fig. 2 from a TV series for the dataset. U-net is trained on the dataset with some data augmentation such as affine transformation and random cropping. From the input line-drawing (Fig. 2(a)), U-net predicts the color label for each pixel (b).

2.2 Strategic Withdrawal

To decrease the number of falsely predicted regions, we strategically withdraw the automatic colorizing process in the region that is difficult to predict using U-net. Thus, we introduce a novel parameter to indicate the confidence of the predicted color label for each region to be colorized. Let the number of pixels assigned to the most frequently predicted label denoted by index p, be $R_p(\Omega)$ on the closed region Ω of the character. Confidence $c(\Omega)$ is defined as $c(\Omega) = R_p(\Omega)/S(\Omega)$ shown in Fig. 2 (c). Note, $S(\Omega)$ represents the number of pixels in Ω. If the confidence $c(\Omega)$ of the region is higher than the threshold $t_h = 0.8$ in these examples, then our system colorizes the image, and if the confidence is lower, then the system strategically withdraws the automatic colorization and fills the region using a noticeable pattern, as shown in Fig. 2 (d) to be colorized manually (e), later. By automatically colorizing only higher confident regions and manually colorizing the remaining regions, we significantly increase the efficiency of the entire colorization process.

3 DISCUSSION AND CONCLUSION

For validation, Fig. 3(a) shows the relationship between confidence and the prediction accuracy evaluated by 228 image pairs not used in the training process. Because the graph shows an obvious positive correlation, the proposed confidence c works for estimating the prediction accuracy. Therefore, higher confidence c is preferred if more accuracy is needed; however, more withdrawing results more regions to be colorized manually. Figure 3(b) shows this trade-off between the prediction accuracy and the ratio of the automatically



(a) confidence and prediction accuracy (b) trade-off between accuracy and colorized region ratio

Figure 3: (a) demonstrates higher confidence region tends to be colorized correctly. (b) shows the trade-off between the expected prediction accuracy and the area of higher confident regions.

colorized regions along with the confidence threshold. For example, if we choose the confidence threshold $t_h = 0.8$, then over 80% of the regions are colorized by the system and 90% accuracy is expected. Note that, we ignore very small regions in the evaluation because it is difficult to manage these regions correctly using the U-net training process.

To conclude, we proposed a practical method that takes account of prediction confidence for anime-style colorization of an input line-drawing. It will contribute reducing the cost of manual correction.

REFERENCES

Sophie Ramassamy, Hiroyuki Kubo, Takuya Funatomi, Daichi Ishii, Akinobu Maejima, Satoshi Nakamura, and Yasuhiro Mukaigawa. 2018. Pre- and Post-Processes for Automatic Colorization Using a Fully Convolutional Network. In *SIGGRAPH Asia 2018 Posters (SA '18)*. Article Article 70, 2 pages.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. *LNCS* 9351, 234–241. https://doi.org/10.1007/978-3-319-24574-4_28