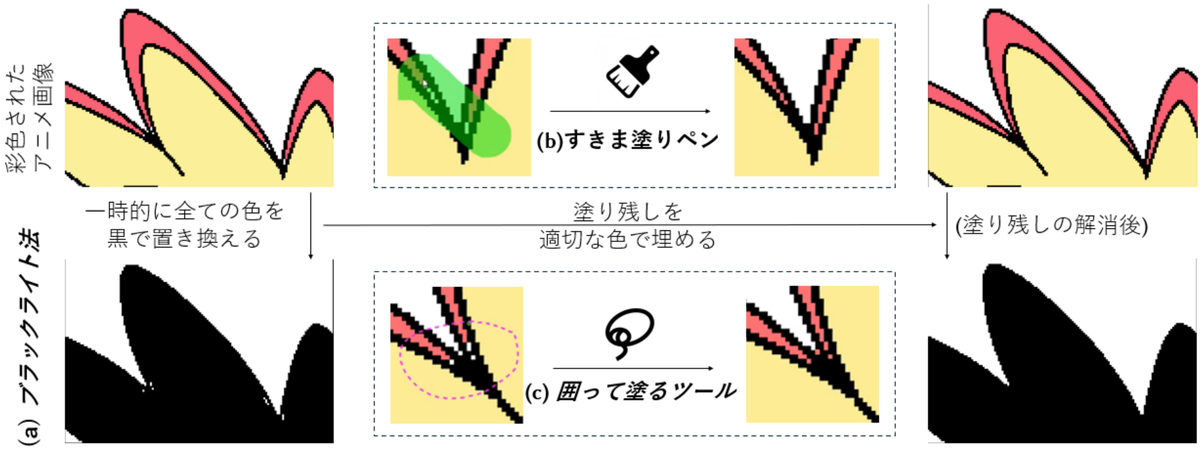
Accepted: CHI' 26 (Full Paper) / SIGGRAPH' 25 (Poster) / WISS '25 (Demo)

GapFill: アニメ調彩色における塗り残しの解消を支援するツール

河野 将大* 前島 謙宣† 小山 裕己* SECHAYK Yotam* 五十嵐 健夫*

*東京大学 †OLM Digital Inc./IMAGICA GROUP Inc.

背景



【背景】
- アニメの「**仕上げ（彩色）**」工程は、手作業による**バケツ塗り**が主流
- 毛先などの鋭角部で「**塗り残し**」が頻発・**検出や解消に余分な手間**
- 既存ツールや自動彩色手法 [1] では、小さい塗り残しへの対応が不十分

【既存ツール】（事前調査により判明した、CSP [2]のツール）

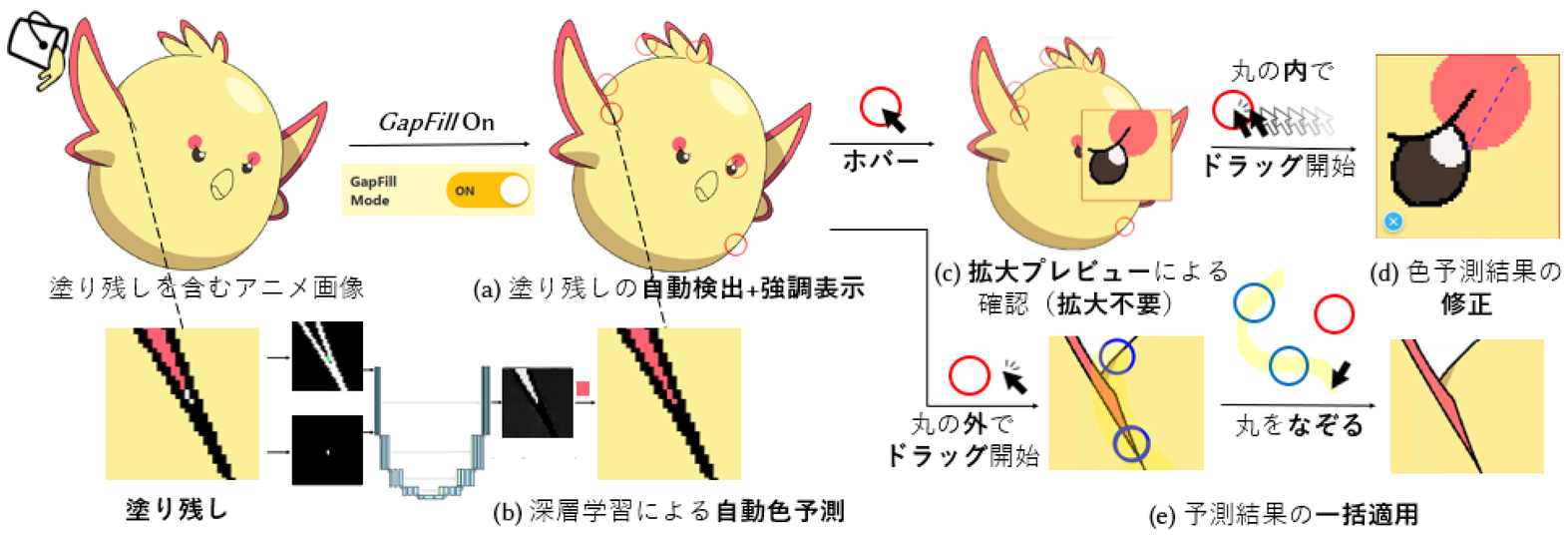
(a)「**ブラックライト法**」による検出
一時的に全ての塗りを黒へ置換・未着色部が白く浮かび上がるように

(b)「**すきま塗りペン**」による解消
ブラシでなぞった軌跡中の閉領域を、指定色でまとめて塗りつぶす

(c)「**囲って塗るツール**」による解消
投げ縄で囲んだ範囲内にある閉領域を、指定色でまとめて塗りつぶす

【GapFillの設計思想】
- 必然的な「**塗り残し検出**」「**拡大操作**」「**色選択**」の反復操作を省略
- 主に周囲の色情報などの**局所的な文脈**を参照した色の提案
- **現場の多様な操作体系を踏襲**した、直感的な操作性・**AIの手動制御**の余地

手法・実装



【GapFillの機能】

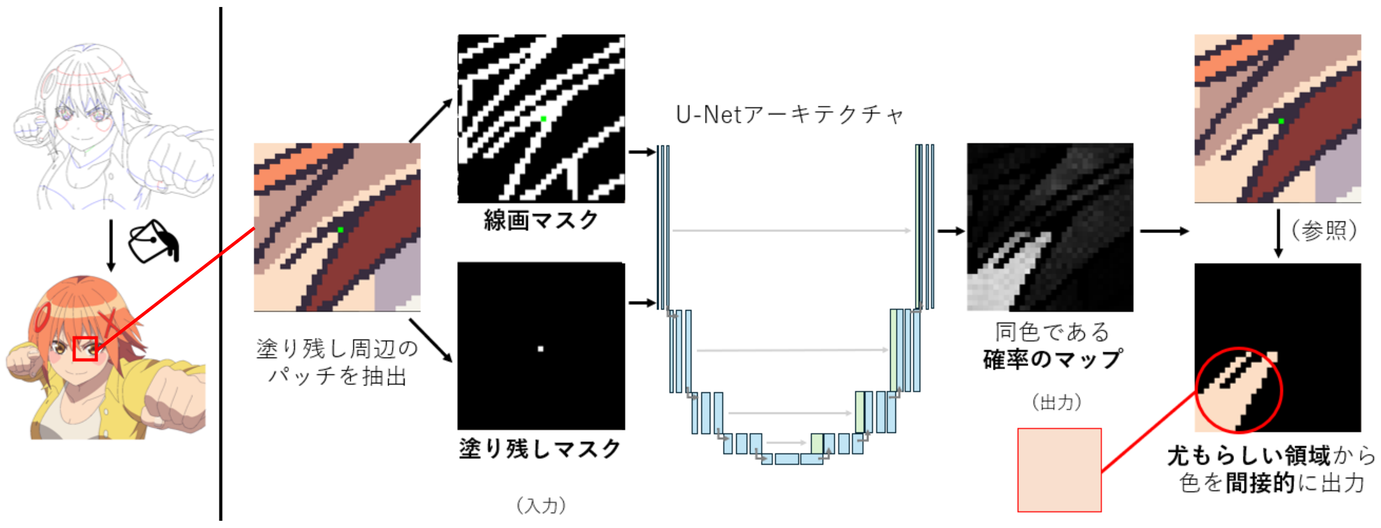
(a) 塗り残しの自動検出＋強調表示
BFSにより、未着色の閉小領域を自動で検出
円形のマーカーで可視化（「**検出**」の省略）

(b) 深層学習による自動色予測
塗り残しに対して、AIが適切な色を提案
一時的にオーバーレイ表示（「**色選択**」の省略）

(c) 拡大プレビューによる確認（ホバー操作）
マーカー上にカーソルがある間、局所的に拡大表示
拡大不要で色予測を確認可能（「**拡大操作**」の省略）

(d) 色予測結果の修正（カラーピッカー操作）
予測が外れていた場合、マーカー**内**でドラッグ開始
近傍の色を選択して直感的に修正可能

(e) 予測結果の一括適用（ブラシ操作）
予測が正しい場合、マーカー**外**でドラッグ開始
なぞったマーカーへ予測色を一括適用



【色予測手法】

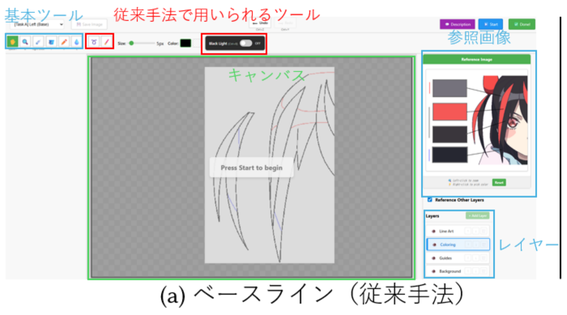
アプローチ：領域間対応の学習による間接的色予測
- アニメ画像特有の「**ベタ塗り（単色）**」の有効活用
- 塗り残し発生の**非色依存性**に着目
- 色を直接予測せず、**領域間の対応関係**を学習（**頑健**）
- 「**周辺のどの大領域と同色か**」を推定・**間接的**色予測

ネットワーク構造：U-Net [3] ベース

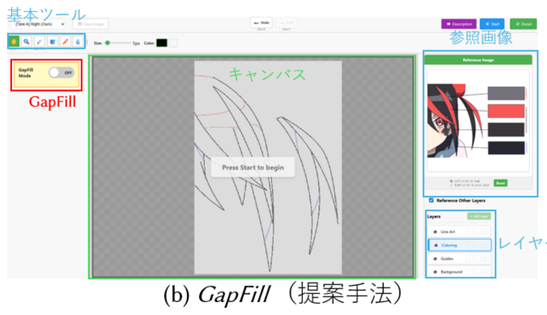
入力：「**線画マスク**」＋「**塗り残し領域マスク**」
出力：対象領域と同色の確率を示す「**空間尤度マップ**」
色決定：マップ中で、**最も尤度が高い領域**の色を参照

【学習データ】
- プロのアニメ画像から線画抽出・閉領域分割
- 閾値以下の小閉領域を「潜在的な塗り残し」と見なす

実験・結果



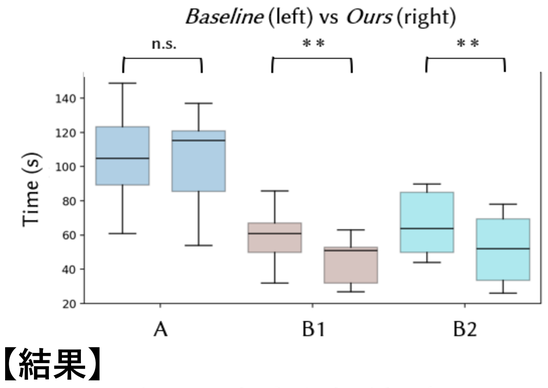
(a) ベースライン（従来手法）

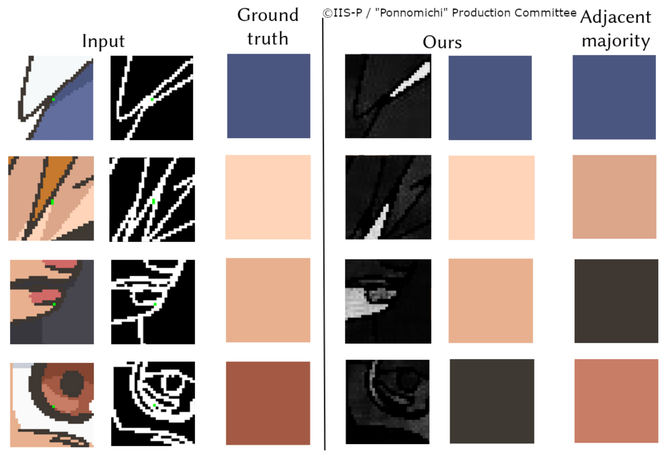


(b) *GapFill*（提案手法）

【実験概要】
- 被験者：13名のプロのアニメ彩色スタッフ
- 環境：業界標準（CSP）の基本機能と制作環境を模した自作ペイントソフト
- 使用画像：髪を含んだ**左右対称**なアニメ画像を**半分に分割**
- タスク：既存ツールとGapFillの**比較（所要時間・見落とし数）**
  - **タスクA（通常彩色）**：実線画を**一から彩色**
  - **タスクB（最終確認）**：ほぼ彩色済みな画像の**塗り残し検出・修正特化**
- 定性的評価：機能単位の分析・アンケート・半構造化インタビュー

- 色予測手法の評価
  - **主観的**：閾値以下の閉領域に一括適用後、**有用性を目視評価（タスクC）**
  - **客観的**：未知データセットに対して、閾値以下の閉領域の色を予測





【結果】
- 所要時間：**有意な短縮（タスクB）**
- 見落とし数：**0** (提案手法) vs 見落とし有り（従来手法）（タスクA・B）
- 有用性（定性）：概ね好評（特に**最終確認工程**で強い支持）
- 機能単位の分析：**自動検出**や拡大プレビュー等の効率的な操作性を評価
- 色予測（主観）：実務視点の評価にて、**肯定的評価**（中央値 5/7）を獲得
- 色予測（客観）：**81.68%** (提案手法) vs 37.02%（周辺最頻色採用の貪欲法）

【考察】
- 有用性には**色予測精度以外の要素の寄与**も大きく、**AIへの信頼**等も影響
- 塗り残しの色選択に**一意の正解がない**場合もある
- 本ツールが既存ツールを補完する形で**柔軟に組み合わせられる**可能性

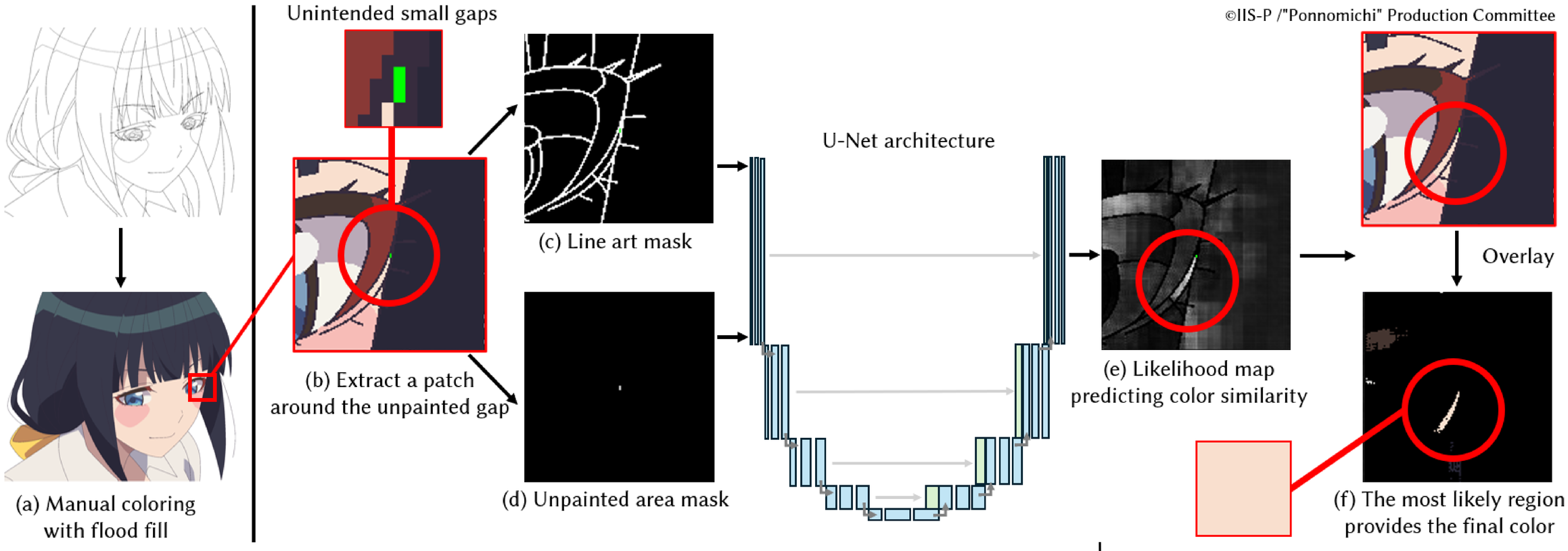
[1] A. Maejima, S. Shinagawa, H. Kubo, T. Funatomi, T. Yotsukura, S. Nakamura, and Y. Mukaigawa. Continual few-shot patch-based learning for anime-style colorization. Computational Visual Media, 10(4):705–723, 2024.
[2] Celsys. CLIP STUDIO PAINT. Website, 2025. Retrieved January 28, 2026 from https://www. clipstudio.net.
[3] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention, pp. 234–241. Springer, 2015.

Predicting Colors in Unpainted Gaps for Anime-Style Illustration

Masahiro Kono[1], Akinobu Maejima[2], Yuki Koyama[1], Takeo Igarashi[1]
[1]The University of Tokyo, [2]OLM Digital, Inc. / IMAGICA GROUP Inc.

SIGGRAPH 2025
Vancouver+ 10-14 August

Method



■ Color Prediction via Region Prediction

We propose a **U-Net** [1] based method to predict flat colors for **unpainted gaps** in anime-style line drawings

1. Context Extraction:
Crop a small image patch centered on the unpainted gap **(b)**

2. Input Encoding:
- **Line art** mask (1 for lines) **(c)**
- **Unpainted area** mask (1 for target region) **(d)**

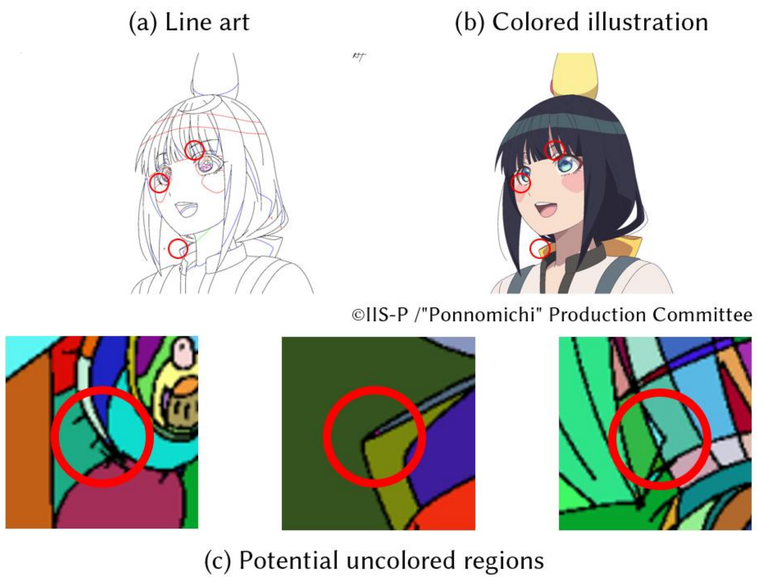
3. Model Inference:
The model predicts a **likelihood map** showing how likely each pixel belongs to the same color region **(e)**

4. Predicted Color Selection:
The final color is chosen from the region with the **highest average likelihood (f)**
⇒ Learns the **correspondence between regions** rather than the colors themselves (more **robust**)
- The occurrence of such gaps is **independent of the color being used**
- Utilizes the characteristic of anime-style illustrations **(flat coloring)**

■ Dataset Creation

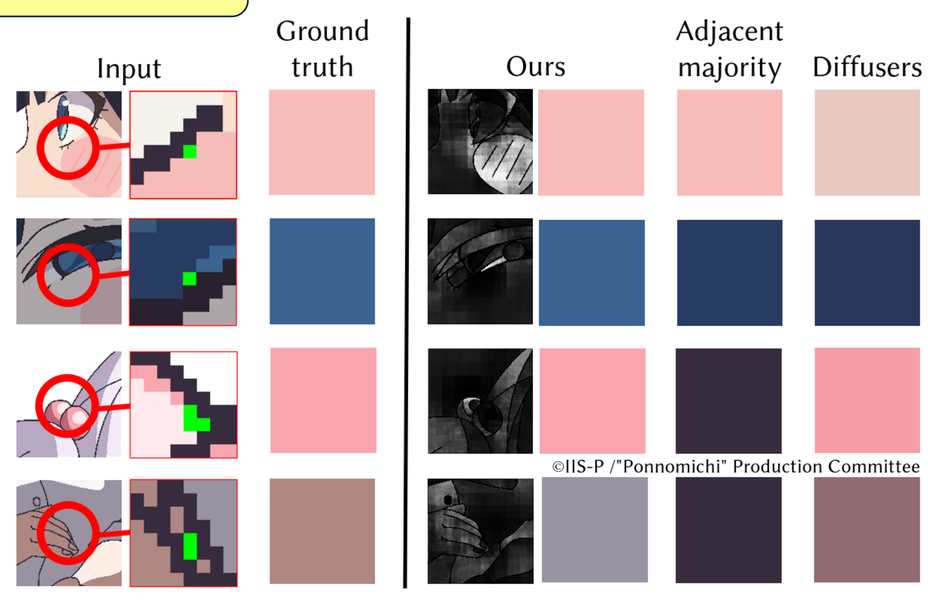
- **Source**:
  One full episode of **professional animation**
- **Gap Definition for Synthetic Data**:
  - Simulate flood fill colorization by detecting **closed regions** using BFS
  - Define **potential "gaps"** as regions with pixel area below a certain threshold
- **Total**:
  234,340 target regions
- **Preparation**:
  - Extract patches and apply preprocessing
  - Augmentation by rotation and flipping

Background



- ***Flood Fill* (*Paint Bucket*) tools** are commonly used to fill enclosed areas with colors
- **Problem: Unpainted "gaps"** often remain due to unintended line intersections
  - These gaps are **hard to detect** and **time-consuming to fix manually**
  - They occur even in professional anime production workflow
- **Related work**
  - **Existing automatic colorization methods** [2,3] require a reference or user hints
  - **Inpainting methods** using U-Net exist [4], but they target natural images
    - Do not fit the **flat-color nature (without gradients)** of anime art
- **Key observation**:
  Gaps should be filled with one of the colors from their **neighboring regions**

Result



- **Comparison Methods**:
  - **Rule-based**: the majority color from 8-connected adjacent pixels
  - **Diffusers** [5]: general-purpose deep learning inpainting model
- **Accuracy on unseen data (full episode)**:
  - **Our method**: 62.57% (90,265 / 144,258 cases)
  - **Rule-based**: 22.38% (32,292 / 144,258 cases)
  - **Diffusers:** often produced **gradients**, unnatural for flat coloring
- **Future Work**:
  - Improve accuracy **parameter tuning** and **other strategies**
  - Design **UI** for integration into real-world production pipelines

Acknowledgements

This work is based on the results obtained from GENIAC (Generative AI Accelerator Challenge), a project implemented by the Ministry of Economy, Trade and Industry (METI) and the New Energy and Industrial Technology Development Organization (NEDO)

References

[1] Ronneberger et al, U-net: Convolutional networks for biomedical image segmentation. (2015)
[2] Sato et al, Reference-based manga colorization by graph correspondence using quadratic programming. (2014)
[3] Cao et al, AnimeDiffusion: Anime diffusion colorization. (2024)
[4] Zhang et al, Image inpainting based on deep learning: A review (2023)
[5] Platen et al, Diffusers: State-of-the-art diffusion models. (2022)

Predicting Colors in Unpainted Gaps for Anime-Style Illustration

Masahiro Kono
The University of Tokyo
Tokyo, Japan
ms-kono@animins.org

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

Yuki Koyama
The University of Tokyo
Tokyo, Japan
koyama@pe.t.u-tokyo.ac.jp

Takeo Igarashi
The University of Tokyo
Tokyo, Japan
takeo@acm.org

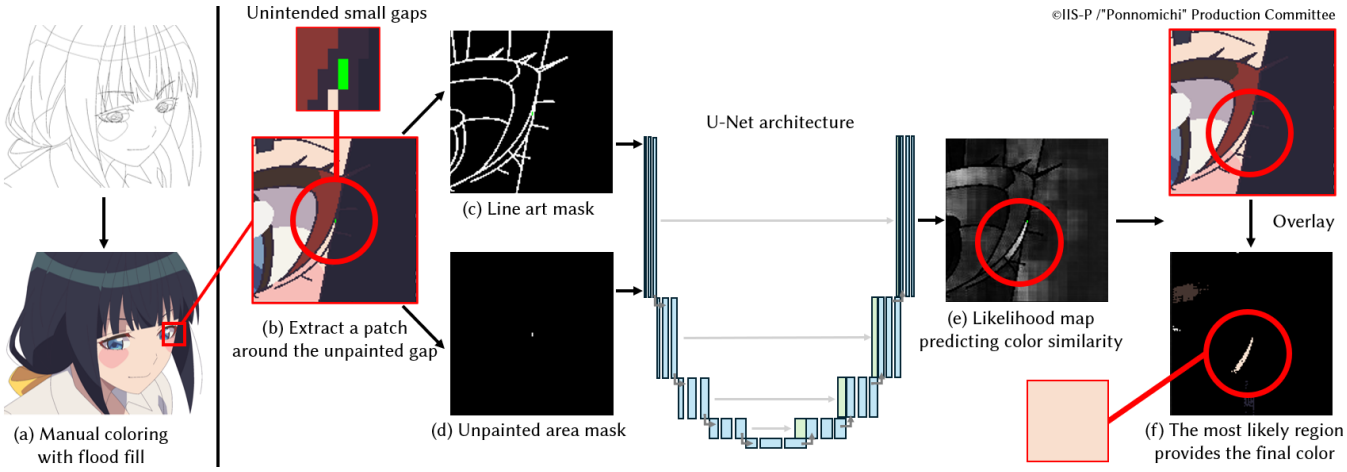


Figure 1: (a) Manual colorization of line drawings with flood fill tends to leave unintended small regions (gaps) unpainted. To address this problem, we propose a novel color prediction method: (b) we extract a patch around the target gap as a context for inference. (c) A *line art mask* and (d) an *unpainted area mask* are input to U-Net, which outputs (e) a likelihood map showing how likely each pixel matches the target color. (f) The final color is then chosen from the region with the highest likelihood.

CCS Concepts

• **Computing methodologies** → **Image processing**.

Keywords

Anime; colorization; digital painting; deep learning; inpainting

ACM Reference Format:
Masahiro Kono, Akinobu Maejima, Yuki Koyama, and Takeo Igarashi. 2025. Predicting Colors in Unpainted Gaps for Anime-Style Illustration. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Posters (SIGGRAPH Posters '25), August 10–14, 2025, Vancouver, BC, Canada.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3721250.3743027

1 Introduction

Anime-style digital illustrations, particularly those rooted in Japanese pop culture, have become a prominent form of creative production. The typical workflow involves drawing binary line art (Figure 2a), filling segmented regions with base colors, and then refining the image by adding details such as shadows and highlights (Figure 2b). To fill regions with colors, the "Paint Bucket" (or "Flood Fill") tool is commonly used, which employs region-growing algorithms to fill enclosed areas bounded by lines [Burtsev and Kuzmin 1993] (Figure 1a). However, small unpainted regions—referred to as *gaps*—often remain due to unintended line intersections, especially at sharp corners such as hair tips (Figure 2c). These gaps are often difficult to detect and time-consuming to fix manually. They frequently occur even in professional anime production environments.

Many studies have explored automatic colorization for anime-style images, ranging from early graph-matching methods [Sato et al. 2014] to recent diffusion-based approaches [Cao et al. 2024]. However, most of these rely on a fully colorized reference image or user-provided hints as additional input, which is not applicable in our setting. A related task is image inpainting, for which many U-Net-based methods have been proposed [Zhang et al. 2023], primarily targeting natural images. In contrast, anime-style illustrations are characterized by flat color regions, and uncolored gaps should often be colored by one of the colors of their spatially neighboring areas, requiring a different approach.

Our contributions are twofold. First, we present the first academic attempt to address the unpainted gap problem in anime-style colorization workflows. Second, we formulate it as a color prediction task and propose a foundational solution with evaluations.

SIGGRAPH Posters '25, Vancouver, BC, Canada

ACM ISBN 979-8-4007-1549-5/25/08
https://doi.org/10.1145/3721250.3743027

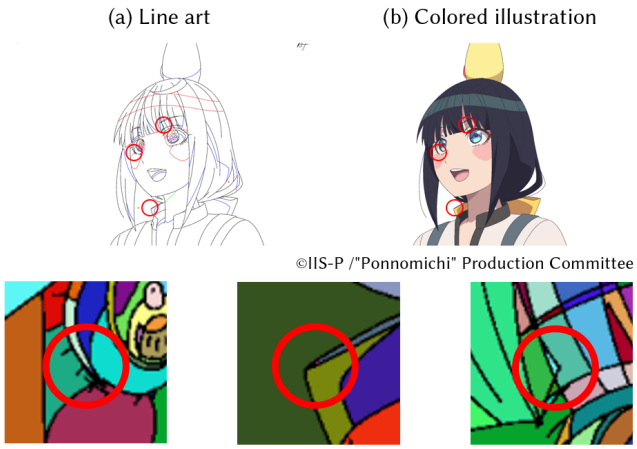


Figure 2: An example of (a) a line art and (b) a final colored illustration. (c) During flood fill operations, unintended small enclosed regions can potentially become unpainted gaps.

2 Method

2.1 Creating Dataset

Based on the characteristics of anime-style illustration, we construct synthetic training data by applying BFS-based fill operations to line drawings and defining potential unpainted gaps as areas with pixel counts below a certain threshold. Then, for each such region extracted from a full episode of professional animation (234, 340 targets in total), we crop a small image patch centered on it (Figure 1b), apply preprocessing and data augmentations, and finally use it as input to the model (Figure 1c,d).

2.2 Color Prediction via Region Prediction

To address this novel problem setting, we propose a method for predicting appropriate colors for unfilled regions. Since the occurrence of such regions is independent of the specific color being used, our approach does not aim to learn the colors themselves, but rather the correspondence between regions. This enables us to build a robust model that does not rely on specific color distributions and to predict flat colors without gradients, which is typical in anime-style images. We design a deep learning model based on U-Net [Ronneberger et al. 2015], which has demonstrated strong performance in capturing features of neighboring regions. Figure 1 illustrates the overall prediction framework. The model takes a two-channel binary input: (Figure 1c) a *line art mask* where line pixels are set to 1; and (Figure 1d) an *unpainted area mask* where the target unpainted region is encoded as 1. The model then predicts a spatial likelihood map indicating the probability that each pixel within the patch shares the same color as the target area (Figure 1e). The final color is determined by selecting the color from the region with the highest average predicted likelihood (Figure 1f).

3 Evaluation and Conclusion

The effectiveness of our method was evaluated by comparisons with two methods, as shown in Figure 3: a rule-based approach

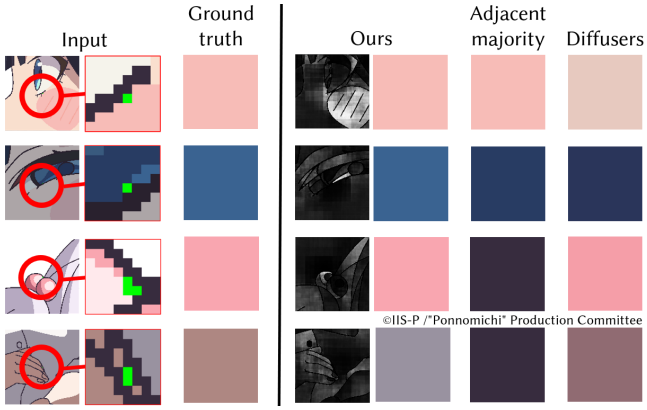


Figure 3: Comparison of color predictions between our method, a rule-based adjacent majority approach, and *diffusers*, a widely used deep learning inpainting model.

that selects the majority color among eight-connected adjacent pixels of the target region, and *diffusers* [von Platen et al. 2022], a widely used deep learning inpainting model for general-purpose use. Our method achieved an accuracy of **62.57%** (= 90, 265/144, 258 cases) on unseen data from a different full episode, where accuracy is defined as the proportion of cases in which the predicted color exactly matches the ground truth. On the other hand, the rule-based approach achieved an accuracy of **22.38%** (= 32, 292/144, 258 cases). *Diffusers* often produced colors with noticeable gradients, which are atypical in anime-style illustrations. These results demonstrate the robustness and superiority of our approach over naive alternatives.

In summary, our approach establishes a promising foundation for supporting the colorization process in anime-style digital illustrations. Future directions include improving accuracy through parameter tuning and training strategies, as well as designing user interfaces for integration into real-world production pipelines.

Acknowledgments

This work is based on the results obtained from GENIAC (Generative AI Accelerator Challenge), a project implemented by the Ministry of Economy, Trade and Industry (METI) and the New Energy and Industrial Technology Development Organization (NEDO).

References

SV Burtsev and Ye P Kuzmin. 1993. An efficient flood-filling algorithm. *Computers & graphics* 17, 5 (1993), 549–561.

Yu Cao, Xiangqiao Meng, PY Mok, Tong-Yee Lee, Xueting Liu, and Ping Li. 2024. AnimeDiffusion: Anime diffusion colorization. *IEEE Transactions on Visualization and Computer Graphics* 30, 10 (2024), 6956–6969.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 2015. U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*. Springer, 234–241.

Kazuhiro Sato, Yusuke Matsui, Toshihiko Yamasaki, and Kiyoharu Aizawa. 2014. Reference-based manga colorization by graph correspondence using quadratic programming. In *SIGGRAPH Asia 2014 Technical Briefs*. 1–4.

Patrick von Platen, Suraj Patil, Anton Lozhkov, Pedro Cuenca, Nathan Lambert, Kashif Rasul, Mishig Davaadorj, Dhruv Nair, Sayak Paul, William Berman, Yiyi Xu, Steven Liu, and Thomas Wolf. 2022. Diffusers: State-of-the-art diffusion models. https://github.com/huggingface/diffusers.

Xiaobo Zhang, Donghai Zhai, Tianrui Li, Yuxin Zhou, and Yang Lin. 2023. Image inpainting based on deep learning: A review. *Information Fusion* 90 (2023), 74–94.