

Uncolorable Examples: Preventing Unauthorized AI Colorization via Perception-Aware Chroma-Restrictive Perturbation

Yuki Nii*, Futa Waseda*, Ching-Chun Chang[†], Isao Echizen*[†]

* The University of Tokyo, Japan

E-mail: {yuki-nii, futa-waseda}@g.ecc.u-tokyo.ac.jp

[†] National Institute of Informatics, Japan

E-mail: {ccchang, iechizen}@nii.ac.jp

Abstract—AI-based colorization has shown remarkable capability in generating realistic color images from grayscale inputs. However, it poses risks of copyright infringement—e.g., the unauthorized colorization and resale of monochrome manga and films. Despite these concerns, no effective method currently exists to prevent such misuse. To address this, we introduce the first defensive paradigm, *Uncolorable Examples*, which embed imperceptible perturbations into grayscale images to invalidate unauthorized colorization. To ensure real-world applicability, we establish four criteria: effectiveness, imperceptibility, transferability, and robustness. Our method, Perception-Aware Chroma-Restrictive Perturbation (PACHroma), generates Uncolorable Examples that meet these four criteria by optimizing imperceptible perturbations with a Laplacian filter to preserve perceptual quality, and applying diverse input transformations during optimization to enhance transferability across models and robustness against common post-processing (e.g., compression). Experiments on ImageNet and Danbooru datasets demonstrate that PACHroma effectively degrades colorization quality while maintaining the visual appearance. This work marks the first step toward protecting visual content from illegitimate AI colorization, paving the way for copyright-aware defenses in generative media.

I. INTRODUCTION

Recent advances in AI colorization [1], [2] has demonstrated remarkable capability in generating realistic color images from grayscale inputs. However, these advancements raise significant ethical and legal concerns. In Japan, for instance, a man was arrested for selling unauthorized colorized versions of the famous animation “Godzilla” [3]. With the increasing accessibility of powerful colorization models, malicious users can easily colorize manga or movies without the creator’s consent and resell them, leading to copyright infringement. Yet, no method currently exists to prevent such unauthorized colorization, highlighting the urgent need for protection.

In this paper, we present the first defensive paradigm against unauthorized image colorization, termed **Uncolorable Examples**, and establish four key criteria for practical applicability: effectiveness, imperceptibility, transferability, and robustness. To meet these criteria, Uncolorable Examples are generated using our proposed method, **Perception-Aware Chroma-Restrictive Perturbation (PACHroma)**. PACHroma utilizes the idea of adversarial examples by embedding im-

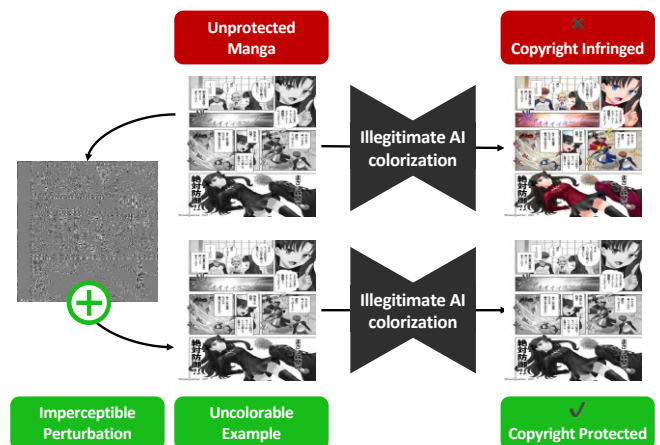


Fig. 1. Overview of Uncolorable Examples. Without protection, grayscale images can be illegitimately colorized by AI colorization models. Our method, PACHroma, generates Uncolorable Examples by adding human-imperceptible perturbations to the input, effectively invalidating unauthorized colorization.

perceptible perturbations into grayscale images to invalidate AI colorization (Fig. 1). These perturbations are designed to concentrate on high-frequency regions by leveraging a Laplacian filter to target coarse areas. Furthermore, to enhance transferability across diverse colorization models and improve robustness to common image transformations (e.g., resizing, compression), we optimize the perturbations via iterative input transformations. Our method achieves significant suppression of colorization quality with minimal visual change compared to the unprotected image. We evaluate our method on both natural and manga image datasets, demonstrating consistent effectiveness across multiple state-of-the-art colorization models (Fig. 2). Our contributions are summarized as follows:

- **Novel defensive paradigm:** We introduce *Uncolorable Examples*, the first defense against unauthorized colorization via imperceptible perturbations to grayscale images.
- **Definition of defense criteria:** We establish four key criteria, *effectiveness*, *imperceptibility*, *transferability*, and *robustness*, for practical colorization defenses.

- **Practical method (PACHroma):** We propose Perception-Aware Chroma-Restrictive Perturbation (PACHroma), which generates Uncolorable Examples by balancing the four criteria through Laplacian filtering and input transformations during optimization.
- **Comprehensive empirical validation:** Experiments on ImageNet and Danbooru demonstrate that PACHroma effectively invalidates state-of-the-art colorization models without changing the visual quality of the image.

II. RELATED WORK

Automatic Colorization. Early work on automatic image colorization focused on CNN-based methods (e.g., DeOldify [4]) which predict plausible color channels from grayscale inputs. Later, GAN-based methods (e.g., BigColor [8] and GCP [2]) emerged, aiming to produce more diverse and vibrant colorizations using learned priors. More recently, transformer-based approaches (e.g., DDColor [6]) leverage global attention for generating semantically consistent and photorealistic results. In this work, we aim at invalidating colorization models—including CNN-, GAN-, and transformer-based models.

Adversarial Examples. Adversarial examples are carefully crafted inputs that cause neural networks to produce incorrect results. Szegedy et al. [9] first revealed the vulnerability of AI models to small perturbations. Among the most well-known attacks, Projected Gradient Descent (PGD) [10] is regarded as a strong first-order adversary and is widely used. While most adversarial examples are used to attack classification models [11], [12], we repurpose them as a defense mechanism by crafting Uncolorable Examples—grayscale images with imperceptible perturbations that look unchanged to humans, but block AI models from adding unauthorized colors.

Image Translation Protection. Yeh et al. [13] proposed an attack against GAN-based image translation (e.g., deepfake synthesis) by adding adversarial perturbations to the input image to nullify or distort the model’s result. Consequently, subsequent studies [14], [15] have focused on disrupting deepfake generation. Motivated by this, we aim at preventing unauthorized colorization, which has been unexplored. We introduce a novel defense paradigm, define key criteria, then propose a practical defense.

III. A DEFENSE PARADIGM: UNCOLORABLE EXAMPLES

We introduce a novel defensive paradigm, *Uncolorable Examples*, which embeds imperceptible adversarial perturbations into grayscale images to invalidate unauthorized colorization. Following [13], there are two possible strategies for invalidating colorization: (1) nullifying it to produce a grayscale result, and (2) distorting it to produce unnatural colors. The diversity of plausible colorizations makes unnatural outputs unreliable as a defense. We instead steer results toward grayscale to suppress colorization.

Unlike natural images, manga often contains large, flat regions with minimal detail, making perturbations—especially in backgrounds or speech bubbles—visually conspicuous. This highlights the importance of imperceptible defenses. Moreover,

Algorithm 1 Perception-Aware Chroma-Restrictive Perturbation (PACHroma)

Input: Colorization model $G(\cdot)$; colorfulness loss L_{CF} ; Laplacian mask M ; grayscale image x_l ; max perturbation ϵ ; number of iterations T ; decay factor μ ; block split number s ; number of transformations N

Output: Final adversarial image x_T^{adv}

- 1: Initialize: $\alpha = \epsilon/10$, $g_0 = 0$, $x_0^{\text{adv}} = x_l$
 - 2: **for** $t = 0$ to $T - 1$ **do**
 - 3: Generate a set of transformed inputs $\mathcal{X} = \{x_i^{\text{tran}}\}_{i=1}^N$ using structure-invariant transformations
 - 4: **for** $i = 1$ to N **do**
 - 5: Compute colorized output: $x_i^{\text{rgb}} = G(x_i^{\text{tran}})$
 - 6: Compute gradient: $g^{(i)} = \nabla_x L_{CF}(x_i^{\text{rgb}})$
 - 7: **end for**
 - 8: Compute averaged gradient: $\bar{g}_{t+1} = \frac{1}{N} \sum_{i=1}^N g^{(i)}$
 - 9: Update momentum: $g_{t+1} = \mu g_t + \frac{\bar{g}_{t+1}}{\|\bar{g}_{t+1}\|_1}$
 - 10: Compute perturbation step: $\Delta = M \cdot \alpha \cdot \text{sign}(g_{t+1})$
 - 11: Update adversarial image: $x_{t+1}^{\text{adv}} = \text{Clip}(x_t^{\text{adv}} + \Delta)$
 - 12: **end for**
 - 13: **return** x_T^{adv}
-

colorization models vary widely in architecture, demanding transferable defense, and simple post-processing can easily remove perturbations [16], [17], further emphasizing the need for robustness. To ensure practical applicability, we propose four criteria an effective defense should satisfy:

- **Effectiveness.** Perturbations should disable the model’s ability to add color, resulting in grayscale outputs.
- **Imperceptibility.** Perturbations should be visually imperceptible to the human eye.
- **Transferability.** Perturbations should remain effective across different colorization models.
- **Robustness.** Perturbations should remain effective under common image transformations (e.g., resizing, cropping and JPEG compression).

IV. METHOD

A. Overview

Given a grayscale input image $x_l \in \mathbb{R}^{H \times W}$, we add imperceptible perturbation $\delta \in \mathbb{R}^{H \times W}$ to nullify the colorization generator $G : \mathbb{R}^{H \times W} \rightarrow \mathbb{R}^{H \times W \times 3}$. The perturbation is optimized to minimize the colorfulness score [18], which quantifies visual vividness:

$$L_{CF} = \text{Colorfulness}(G(x_l + \delta)) \quad (1)$$

To improve imperceptibility, transferability, and robustness, we optimize the perturbation using a Laplacian filter and diverse input transformations (see Algorithm 1).

B. Perception-Aware Chroma-Restrictive Perturbation

We propose Perception-Aware Chroma-Restrictive Perturbation (PACHroma), a method that produces Uncolorable Examples. PACHroma applies the Momentum Iterative Fast Gradient

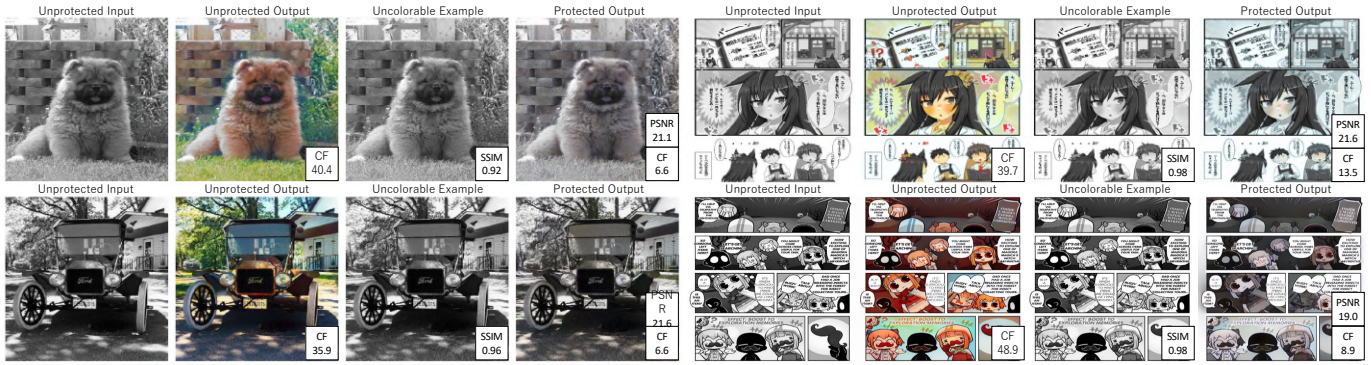


Fig. 2. **Uncolorable Examples from PACHroma invalidate colorization via imperceptible perturbations.** Top-left: DeOldify [4]; top-right: ACDO [5]; bottom-left: DDColor [6]; bottom-right: MC-V2 [7]. Each image is shown with its CF score, SSIM between the input, or PSNR between the output.

Sign Method (MI-FGSM) [11] as its core optimization loop, incorporating the input transformation strategy of Structural Invariant Attack (SIA) [19] and a continuous Laplacian mask during each iteration (Algorithm 1).

1) *Input Transformation*: To enhance transferability and robustness, PACHroma applies structure-preserving augmentations [19]. The input is divided into blocks (e.g., 3×3), and random transformations are applied independently to each. The transformations include geometric changes (shift, flip, rotation), intensity modifications (scaling, jitter, noise), frequency-domain filtering (DCT), resolution changes (resizing), and spatial dropout ($p=0.1$). At each step, N transformed inputs are used to compute a loss that encourages generalization.

2) *Continuous Laplacian Mask*: To enhance imperceptibility, we guide gradient updates with a continuous Laplacian mask, exploiting reduced distortion visibility along edges due to contrast masking in human vision [20]. The continuous Laplacian map M of the input x_l is computed via convolution with a standard Laplacian kernel:

$$M = |\nabla^2 x_l| = |x_l * K_{\text{Laplacian}}|, \quad (2)$$

where $K_{\text{Laplacian}}$ is defined as $[0, 1, 0; 1, -4, 1; 0, 1, 0]$, and $*$ denotes convolution. The resulting map is normalized to range $[0, 1]$ and applied as a multiplicative weighting mask on the gradient during each update step.

By integrating input transformations with a Laplacian mask, PACHroma produces Uncolorable Examples that are simultaneously effective, imperceptible, transferable, and robust. Visualization results of PACHroma are presented in Fig. 2.

V. EXPERIMENTS

A. Experimental Setup

1) *Colorization Models*: We evaluate our method on three natural image colorization models—CNN-based **DeOldify** [4], GAN-based **BigColor** [8], and transformer-based **DD-Color** [6]—using official 12M ImageNet weights. For manga, we use two domain-specific models: **AnimeColorDeOldify (ACDO)** [5] and **Manga Colorization V2 (MC-V2)** [7].

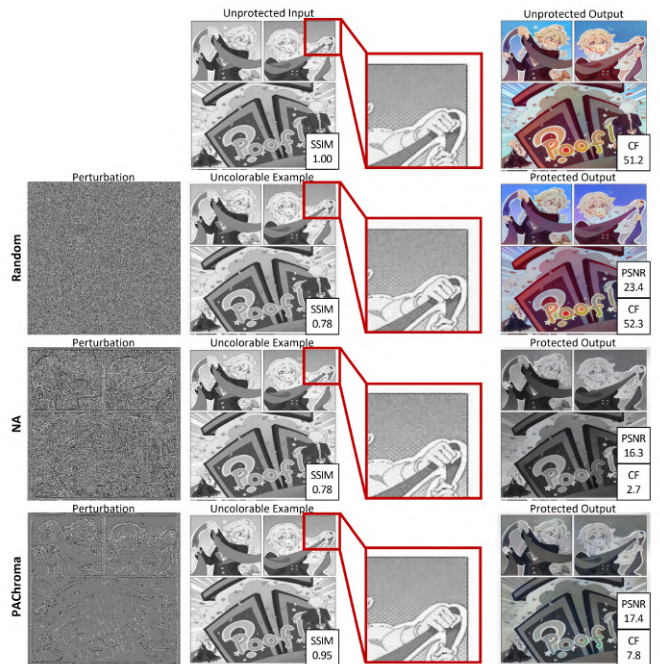


Fig. 3. **Effectiveness and Imperceptibility of random noise, Nullifying Attack (NA), and PACHroma.** PACHroma preserves grayscale structure while preventing colorization, outperforming NA in imperceptibility. Each image includes CF score, SSIM between the input, or PSNR between the output.

2) *Datasets*: **ImageNet**: We sample 100 images from the ImageNet validation set [21], repeated with five random seeds. All images are resized to 256×256 . **Danbooru**: Following [22], we collect 4,367 “manga”-tagged images from Danbooru, resize them to 576×576 , and randomly sample 30 images with five random seeds.

3) *Evaluation Protocol*: We evaluate our method using standard metrics in image colorization [6], [8]. Effectiveness is measured by PSNR and Colorfulness (CF) between unprotected and protected outputs; imperceptibility by PSNR and SSIM between the input. We also evaluate black-box transferability and robustness to post-processing (JPEG 75%,

TABLE I

PERFORMANCE OF UNCOLORABLE EXAMPLES ON NATURAL IMAGES ACROSS COLORIZATION MODELS. PACHROMA EFFECTIVELY SUPPRESSES COLORIZATION WHILE PRESERVING VISUAL QUALITY, BALANCING THE FOUR CRITERIA. GRAY HIGHLIGHTED ROWS INDICATE WHITE-BOX SETTINGS.

Source Model	Attack Model	Attack Type	Unprotected CF	Effectiveness		Imperceptibility		Robustness			
				Protected CF↓	PSNR (Output)↓	PSNR (Input)↑	SSIM (Input)↑	JPEG 75% CF↓	JPEG 50% CF↓	RRC CF↓	
DeOldify	DeOldify	Random	34.15	26.99 (-20.90%)	25.70	28.96	0.75	25.00	24.34	24.72	
		NA	34.15	6.53 (-80.90%)	23.19	28.41	0.74	7.74	12.71	8.62	
		NA-Mask (ours)	34.15	6.65 (-80.51%)	24.88	43.23	1.00	20.65	24.39	20.38	
		PACHroma (ours)	34.15	7.38 (-78.40%)	24.05	32.59	0.95	10.71	14.10	9.06	
	BigColor	NA	34.15	23.22 (-32.00%)	23.14	27.10	0.71	21.09	20.14	21.05	
		NA-Mask (ours)	34.15	27.20 (-20.30%)	28.34	32.54	0.95	26.16	25.36	25.64	
		PACHroma (ours)	34.15	24.50 (-28.30%)	25.66	30.28	0.91	23.17	23.04	23.50	
	DDColor	NA	34.15	25.13 (-26.40%)	24.75	28.33	0.74	23.65	22.92	22.98	
		NA-Mask (ours)	34.15	32.52 (-4.80%)	36.80	40.37	0.99	30.26	28.80	30.52	
		PACHroma (ours)	34.15	27.42 (-19.70%)	28.19	32.60	0.95	26.70	26.47	26.82	
	BigColor	DeOldify	NA	29.91	21.29 (-28.80%)	23.03	27.20	0.70	22.16	22.50	21.26
			NA-Mask (ours)	29.91	28.78 (-3.80%)	31.80	34.84	0.97	26.56	25.49	28.48
PACHroma (ours)			29.91	25.84 (-13.60%)	26.74	30.79	0.92	24.58	23.89	26.01	
BigColor		Random	29.91	27.06 (-9.50%)	22.77	28.96	0.73	24.79	24.02	23.86	
		NA	29.91	0.79 (-97.40%)	22.45	28.28	0.72	1.41	3.68	2.28	
		NA-Mask (ours)	29.91	1.17 (-96.10%)	24.02	37.00	0.98	6.74	11.61	7.39	
		PACHroma (ours)	29.91	5.15 (-82.80%)	23.64	32.48	0.94	6.29	7.83	6.17	
DDColor		NA	29.91	20.07 (-32.90%)	23.10	27.20	0.70	20.08	21.14	20.48	
		NA-Mask (ours)	29.91	28.58 (-4.50%)	31.29	34.39	0.97	26.49	25.45	28.53	
		PACHroma (ours)	29.91	24.44 (-18.30%)	26.60	30.72	0.92	23.36	22.97	24.65	
DDColor		DeOldify	NA	36.84	36.99 (-0.40%)	22.59	28.41	0.74	28.68	26.54	31.46
			NA-Mask (ours)	36.84	35.97 (-2.40%)	37.73	43.23	1.00	31.29	29.30	37.06
	PACHroma (ours)		36.84	31.32 (-15.00%)	26.18	32.59	0.95	28.31	27.55	32.11	
	BigColor	NA	36.84	31.47 (-14.60%)	21.35	27.10	0.71	26.43	25.96	30.77	
		NA-Mask (ours)	36.84	33.71 (-8.50%)	26.59	32.54	0.95	30.49	28.91	35.17	
		PACHroma (ours)	36.84	27.71 (-24.80%)	24.27	30.28	0.91	26.55	25.79	29.31	
		Random	36.84	39.50 (+7.20%)	22.42	28.96	0.75	29.68	27.65	33.52	
	DDColor	NA	36.84	1.43 (-96.10%)	21.14	28.33	0.74	7.41	16.49	12.99	
		NA-Mask (ours)	36.84	2.47 (-93.30%)	22.24	40.37	0.99	20.87	24.41	25.22	
		PACHroma (ours)	36.84	7.60 (-79.40%)	22.02	32.60	0.95	11.81	14.72	12.16	

50%, and random resized cropping).

4) *Defense Settings*: We evaluate PACHroma alongside two baselines: the Nullifying Attack (NA) [13] and NA with Laplacian Mask (NA-Mask). Perturbations are ℓ_∞ -bounded and are optimized using the loss defined in Eq. 1. The default hyperparameters are: $\epsilon = \frac{16}{255}$, $\alpha = \frac{1.6}{255}$, number of iterations $T = 100$, and number of transformed images $N = 20$.

B. Results

Effectiveness. PACHroma achieves sufficient reduction of CF, producing results that differ from the unprotected outputs. As shown in Tab. I, the baseline Nullifying Attack (NA) achieves CF reduction of 80.90%–97.40% and PSNR of 21.14–23.19. While PACHroma shows slightly lower CF reduction (78.40%–82.80%) and comparable PSNR (22.02–24.05), it still effectively disables colorization and yields outputs perceptually distinct from the unprotected ones—meeting the goal of protection. Qualitative comparisons with random noise, NA, and PACHroma are presented in Fig. 3.

Imperceptibility. PACHroma produces Uncolorable Examples that remain visually close to the unprotected inputs while effectively preventing colorization. As shown in Fig. 3, NA introduces visible artifacts—especially on smooth regions like manga backgrounds and faces—whereas PACHroma maintains a natural appearance with minimal distortion. Quantitatively,

Tab. I shows improved perceptual similarity for PACHroma, with PSNR increasing from 28.28 to 32.48 and SSIM from 0.72 to 0.94 on BigColor. Despite a modest drop in CF reduction (from 97.40% to 82.80%), colorization is still strongly suppressed. These trends are consistent across models, indicating that the Laplacian mask enhances imperceptibility without sacrificing defense performance—crucial for manga, where large smooth areas like speech bubbles are common.

Transferability. PACHroma achieves higher transferability in the black-box setting compared to NA with Laplacian masking (NA-Mask). As shown in Fig. 4 and Tab. I, NA-Mask yields only modest CF reduction (2.40%–20.30%), while PACHroma improves this to 13.60%–28.30%. Although the CF reduction is modest compared to the white-box setting, PACHroma still produces perceptibly low-quality colorizations (Fig. 4), effectively hindering malicious users from creating high-quality media for resale and other misuse. These results suggest that PACHroma generalizes better across models by incorporating an input transformation strategy compared to NA-Mask, offering a more practical defense.

Robustness. PACHroma consistently outperforms NA-Mask in terms of robustness to post-processing. As seen in Fig. 5 and Tab. I, NA-Mask is highly vulnerable to common transformations like JPEG compression and random resized cropping

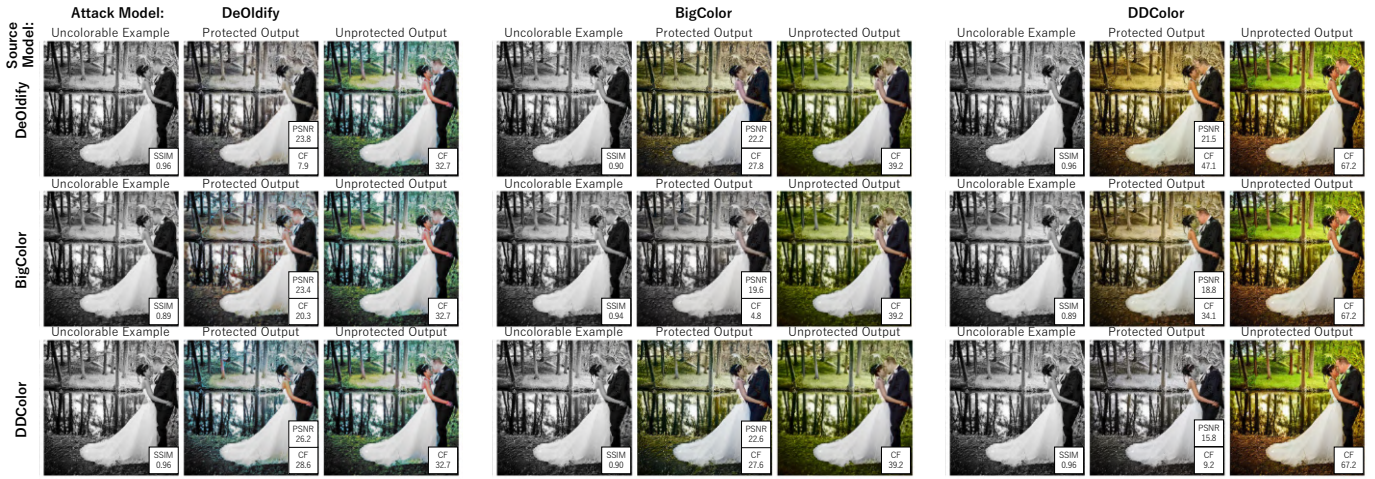


Fig. 4. Transferability of PACHroma among DeOldify, BigColor, and DDColor. Shown with CF, SSIM between inputs, or PSNR between outputs.

TABLE II

PERFORMANCE OF UNCOLORABLE EXAMPLES ON MANGA IMAGES ACROSS COLORIZATION MODELS. PACHROMA ACHIEVES HIGH EFFECTIVENESS, IMPERCEPTIBILITY, AND ROBUSTNESS. HIGHLIGHTED ROWS ARE WHITE-BOX SETTINGS.

Source Model	Attack Model	Attack Type	Effectiveness			Imperceptibility			Robustness		
			Unprotected CF	Protected CF↓	PSNR (Output)↓	PSNR (Input)↑	SSIM (Input)↑	JPEG 75% CF↓	JPEG 50% CF↓	RRC CF↓	
ACDO	ACDO	Random	45.25	59.41 (+31.29%)	22.79	29.66	0.76	60.47	57.87	56.39	
		NA	45.25	5.86 (-87.05%)	21.28	29.42	0.80	9.29	14.49	21.88	
		PACHroma (ours)	45.25	7.72 (-82.95%)	21.93	36.64	0.99	17.73	12.74	26.90	
	MC-V2	NA	45.25	54.06 (+19.48%)	24.35	29.48	0.79	54.48	51.51	49.52	
		PACHroma (ours)	45.25	45.90 (+1.44%)	34.24	34.48	0.99	46.21	41.94	43.00	
		PACHroma (ours)	45.25	45.28 (+0.08%)	31.40	32.52	0.97	45.53	41.46	42.32	
MC-V2	ACDO	NA	54.65	41.63 (-23.84%)	22.24	29.42	0.80	38.56	40.32	40.21	
		PACHroma (ours)	54.65	53.22 (-2.63%)	35.94	36.64	0.99	53.23	50.06	53.64	
		PACHroma (ours)	54.65	51.93 (-4.99%)	29.73	32.82	0.98	51.37	48.95	52.25	
	MC-V2	Random	54.65	46.46 (-15.00%)	22.18	29.65	0.76	40.39	36.21	39.24	
		NA	54.65	2.95 (-94.60%)	16.41	29.48	0.79	8.16	16.05	16.65	
		PACHroma (ours)	54.65	6.01 (-89.01%)	17.34	34.48	0.99	28.60	25.43	29.94	
PACHroma (ours)	54.65	15.94 (-70.84%)	17.71	32.52	0.97	20.93	20.96	22.00			

TABLE III

COMPARISON OF IMAGE COLORIZATION PREVENTION METHODS



Fig. 5. Robustness of NA, NA-Mask, and PACHroma to JPEG compression ($Q=X\%$) and Random Resized Cropping (RRC). Each image is shown with its CF score on the bottom corner.

Method	Effectiveness	Imperceptibility	Transferability	Robustness
Random Noise	-	-	-	-
NA	~90%	-	-	-
NA-Mask (ours)	~90%	✓	-	-
PACHroma (ours)	~80%	✓	✓	✓

(RRC), often resulting in partial colorization recovery. In contrast, PACHroma retains significantly lower CF scores after post-processing—e.g., under JPEG 50%, DDColor’s CF is 14.72 for PACHroma versus 24.41 for NA-Mask. Although NA occasionally yields stronger suppression, PACHroma provides a better balance of imperceptibility and robustness, offering more dependable protection across models.

Manga Domain. Tab. II demonstrates that PACHroma achieves strong effectiveness, imperceptibility, and robustness against manga colorization models. However, its transferability across models remains limited, likely due to the unique char-

acteristics of manga—namely, large flat regions with minimal texture, which constrain perturbation flexibility and hinder generalization. Higher image resolution compared to natural image datasets may further amplify this challenge.

Moreover, our method currently targets fully automatic colorization. Extending to user-guided approaches (e.g., scribble- or text-based) as well as to higher-resolution and more computationally efficient settings remains an open challenge.

VI. CONCLUSIONS

We introduce *Uncolorable Examples*—grayscale images with imperceptible perturbation that resist AI colorization. Generated by our method *PAChroma*, which combines input transformations and a Laplacian mask, they suppress colorization while preserving appearance. *PAChroma* is effective and imperceptible in white-box settings, with moderate transferability and robustness (Tab. III), laying the foundation for protecting content from unauthorized generative manipulation.

VII. ACKNOWLEDGMENTS

This work was partially supported by JSPS KAKENHI Grants JP21H04907 and JP24H00732, by JST CREST Grant JPMJCR20D3 including AIP challenge program, by JST AIP Acceleration Grant JPMJCR24U3, and by JST K Program Grant JPMJKP24C2 Japan.

REFERENCES

- [1] Z. Huang, N. Zhao, and J. Liao, *Unicolor: A unified framework for multi-modal colorization with transformer*, 2022. arXiv: 2209.11223.
- [2] Y. Wu, X. Wang, Y. Li, H. Zhang, X. Zhao, and Y. Shan, *Towards vivid and diverse image colorization with generative color prior*, 2022. arXiv: 2108.08826.
- [3] T. Mainichi. “Man arrested in japan for selling ai-colored pirated 1954 ‘godzilla’ film.” (Jun. 2025).
- [4] J. Antic, *DeOldify: Deep learning for image colorization and restoration*, GitHub, <https://github.com/jantic/DeOldify>, 2021.
- [5] Dakini, AIEMMU, and A. Regmi, *Anime-colordeoldify: Deoldify-based colorization for anime, sketch, and manga*, GitHub, <https://github.com/Dakini/AnimeColorDeOldify>, 2024.
- [6] X. Kang, T. Yang, W. Ouyang, P. Ren, L. Li, and X. Xie, *Ddcolor: Towards photo-realistic image colorization via dual decoders*, <https://arxiv.org/abs/2212.11613>, 2023. arXiv: 2212.11613.
- [7] Manga Colorization V2, *Manga colorization v2*, <https://github.com/qweasdd/manga-colorization-v2>, 2022.
- [8] G. Kim, K. Kang, S. Kim, *et al.*, *Bigcolor: Colorization using a generative color prior for natural images*, 2022. arXiv: 2207.09685.
- [9] C. Szegedy, W. Zaremba, I. Sutskever, *et al.*, “Intriguing properties of neural networks,” *arXiv preprint arXiv:1312.6199*, 2013.
- [10] A. Madry, A. Makelov, L. Schmidt, D. Tsipras, and A. Vladu, “Towards deep learning models resistant to adversarial attacks,” in *International Conference on Learning Representations (ICLR)*, 2018.
- [11] Y. Dong, F. Liao, T. Pang, *et al.*, *Boosting adversarial attacks with momentum*, 2018. arXiv: 1710.06081.
- [12] I. J. Goodfellow, J. Shlens, and C. Szegedy, *Explaining and harnessing adversarial examples*, 2015. arXiv: 1412.6572.
- [13] C.-Y. Yeh, H.-W. Chen, S.-L. Tsai, and S.-D. Wang, “Disrupting image-translation-based deepfake algorithms with adversarial attacks,” in *2020 IEEE Winter Applications of Computer Vision Workshops (WACVW)*, 2020, pp. 53–62. DOI: 10.1109/WACVW50321.2020.9096939.
- [14] H. Huang, Y. Wang, Z. Chen, *et al.*, *Cmua-watermark: A cross-model universal adversarial watermark for combating deepfakes*, <https://arxiv.org/abs/2105.10872>, 2021. arXiv: 2105.10872.
- [15] R. Wang, Z. Huang, Z. Chen, L. Liu, J. Chen, and L. Wang, *Anti-forgery: Towards a stealthy and robust deepfake disruption attack via adversarial perceptual-aware perturbations*, <https://arxiv.org/abs/2206.00477>, 2022. arXiv: 2206.00477.
- [16] C. Xie, J. Wang, Z. Zhang, Z. Ren, and A. Yuille, *Mitigating adversarial effects through randomization*, <https://arxiv.org/abs/1711.01991>, 2018. arXiv: 1711.01991.
- [17] C. Guo, M. Rana, M. Cisse, and L. van der Maaten, *Countering adversarial images using input transformations*, <https://arxiv.org/abs/1711.00117>, 2018. arXiv: 1711.00117.
- [18] D. Hasler and S. Suesstrunk, “Measuring colourfulness in natural images,” *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 5007, pp. 87–95, Jun. 2003. DOI: 10.1117/12.477378.
- [19] C. Shen, Y. Dong, H. Su, and J. Zhu, “Structure-preserving transformation for adversarial example generation,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2021, pp. 6890–6900.
- [20] A. B. Watson and J. A. Solomon, “Model of visual contrast gain control and pattern masking,” *JOSA A*, vol. 14, no. 9, pp. 2379–2391, 1997.
- [21] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “Imagenet: A large-scale hierarchical image database,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009, pp. 248–255.
- [22] M. Xie, C. Li, X. Liu, and T.-T. Wong, “Manga filling style conversion with screentone variational autoencoder,” *ACM Transactions on Graphics (SIGGRAPH Asia 2020 issue)*, vol. 39, no. 6, 226:1–226:15, Dec. 2020.