# PalGAN: Image Colorization with Palette Generative Adversarial Networks


上海人工智能实验室
Shanghai Artificial Intelligence Laboratory

**Yi Wang[1], Menghan Xia[2], Lu Qi[3], Jing Shao[4], and Yu Qiao[1]**

[1]Shanghai AI Laboratory [2]Tencent AI Lab [3]UC Merced [4]SenseTime Research


ECCV TEL AVIV 2022

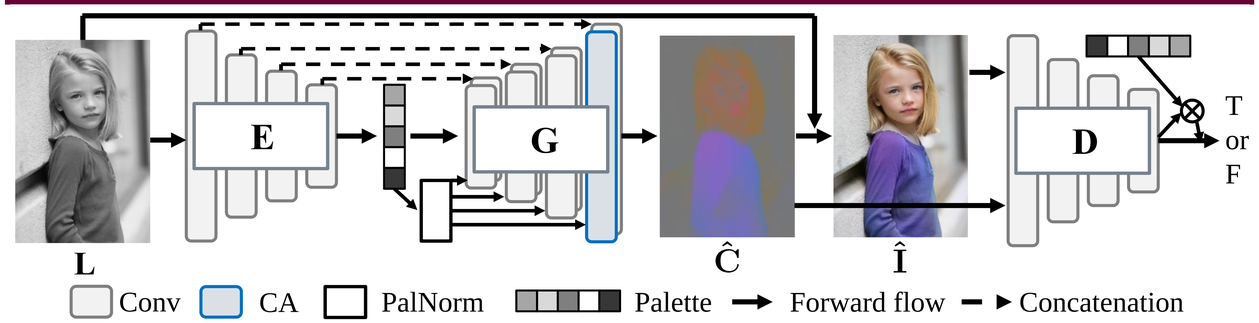
## Introduction

Key Challenges
- Multimodal ambiguity
- Color bleeding

Our method PalGAN
- Palette estimation
  - Predict a probabilistic palette from a gray image
  - Conduct color assignment based on the palette
- Chromatic attention (CA): study color affinities by considering semantic and intensity correlation.

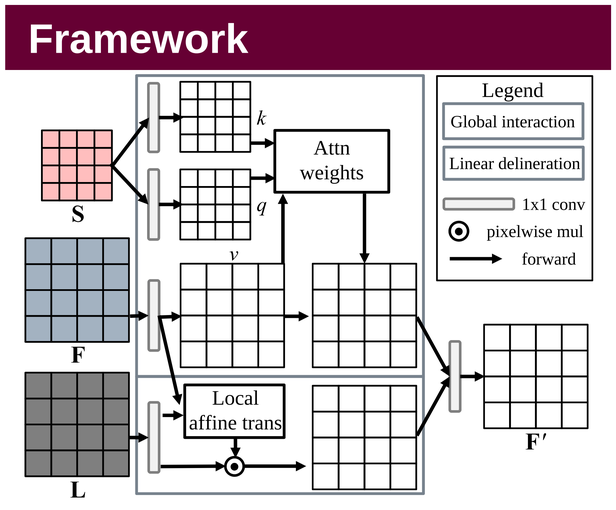
The performance of PalGAN:
- It gets SOTA quantitative results in ImageNet / COCO.
- It gives diverse, contrastive, and edge-preserving appearances.

## Framework




- Our PalGAN formulates colorization as a palette prediction and assignment problem.
- A palette generator E: estimates the global palette probabilities from the given gray image
- A palette assignment generator G: conducts color assignment task via conditional image generation.
  - Chromatic attention
- A color discriminator D: improves the result from adversarial training

## Framework



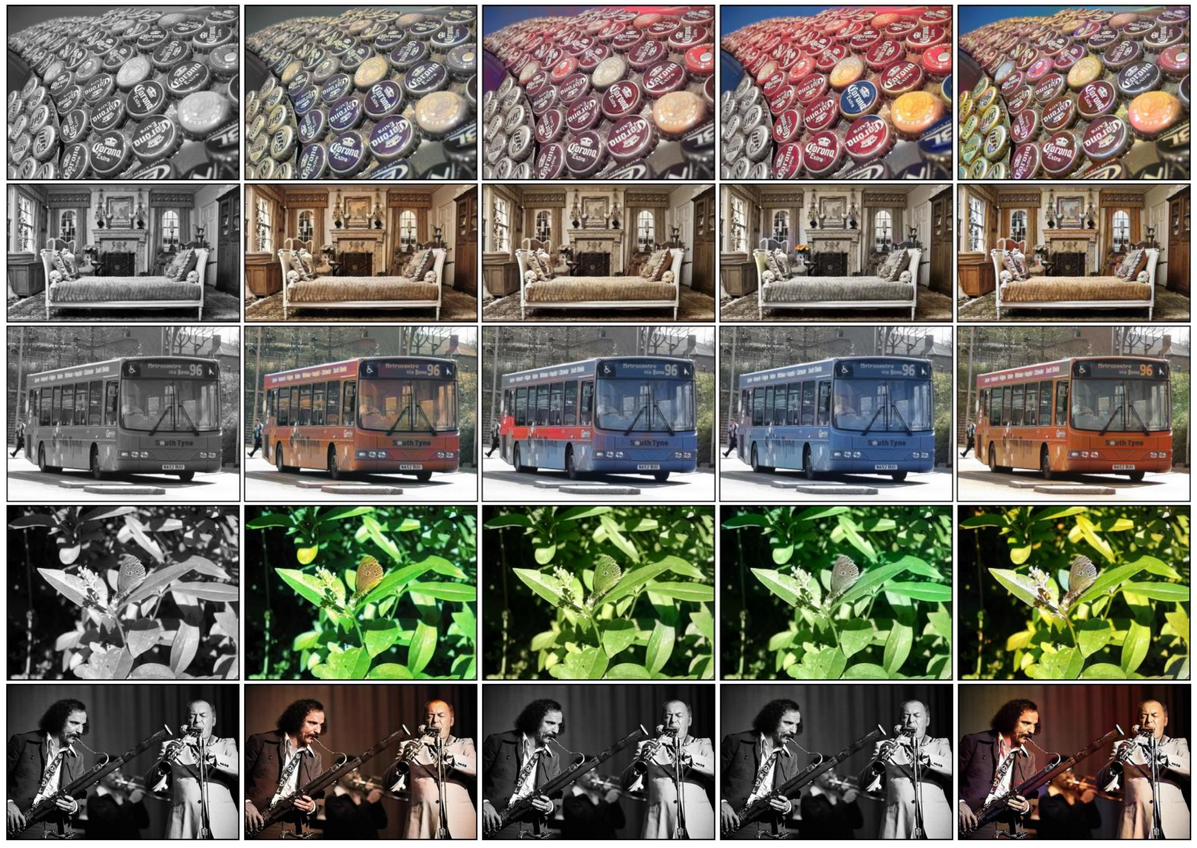

Motivation
- incorporates both semantic and low-level affinities into constructing color relations.

Components: global interaction & local delineation

(a) (b) (c) (d) (e)

Ablation studies of chromatic attention (CA). (a) input, (b) wo CA, (c) w Global, (d) w Local, (e) full CA. Please zoom in.

## Visual Quality



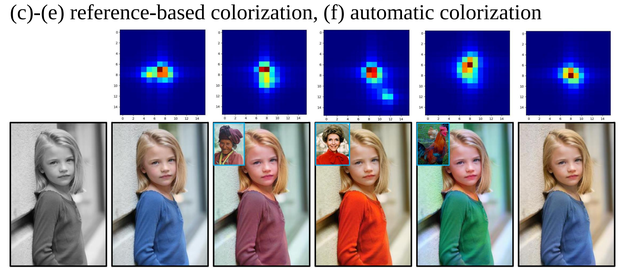
Input Deoldify UGColor InstColor Ours

## Experiment

Quantitative results on the validation sets from different methods.

| Method | ImageNet (ctest10k) PSNR ↑ | SSIM ↑ | LPIPS ↓ | FID ↓ | ImageNet (val50k) PSNR ↑ | SSIM ↑ | LPIPS ↓ | FID ↓ | COCO-Stuff PSNR ↑ | SSIM ↑ | LPIPS ↓ | FID ↓ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CIColor [54] | 22.30 | 0.902 | 0.221 | 12.20 | 22.26 | 0.902 | 0.221 | 9.39 | 21.84 | 0.895 | 0.234 | 22.32 |
| UGColor [56] | 24.26 | 0.918 | 0.174 | 7.49 | 24.26 | 0.919 | 0.173 | 4.60 | 24.34 | 0.924 | 0.165 | 14.74 |
| Lei *et. al.* [26] | 24.52 | 0.917 | 0.202 | 12.60 | 24.03 | 0.918 | 0.189 | 6.35 | 24.59 | 0.922 | 0.191 | 23.10 |
| Deoldify [2] | 23.54 | 0.914 | 0.187 | 5.78 | 22.97 | 0.911 | 0.185 | 3.87 | 23.98 | 0.939 | 0.161 | 12.75 |
| ColTrans [23] | 21.81 | 0.892 | 0.218 | 6.37 | 22.12 | 0.894 | 0.216 | 3.81 | 22.11 | 0.898 | 0.210 | 11.65 |
| Ours[1] | 24.19 | 0.917 | **0.161** | **4.60** | 24.25 | 0.917 | **0.161** | **2.78** | 24.56 | 0.924 | **0.148** | **7.70** |
| Ours[2] | **24.66** | **0.920** | 0.170 | 5.24 | **24.54** | **0.920** | 0.168 | 3.62 | **24.72** | **0.944** | 0.156 | 8.93 |
| InstColor* [38] | 23.03 | 0.909 | 0.191 | 7.35 | 23.06 | 0.910 | 0.190 | 4.94 | 22.35 | 0.838 | 0.238 | 12.24 |
| GPColor* [49] | 21.66 | 0.871 | 0.230 | 5.46 | 21.81 | 0.880 | 0.230 | 3.62 | N/A | N/A | N/A | N/A |
| Ours* | 27.75 | 0.932 | 0.110 | 4.20 | 27.53 | 0.913 | 0.118 | 2.42 | 28.28 | 0.936 | 0.105 | 7.21 |

Visualizations of palettes (1st row, shown in jet colormap) and how they work on colorization (2nd row). (a) Input, (b) the ground truth, (c)-(e) reference-based colorization, (f) automatic colorization



(a) (b) (c) (d) (e) (f)

Ablation studies of chromatic attention (CA). (a) input, (b) wo CA, (c) w Global, (d) w Local, (e) full CA. Please zoom in.

(a) (b) (c) (d) (e)

User study. Each entry gives the percentage of cases where colorization results are favored compared with GT.

| Method | Ours | Coltrans | GPCol | InstCol | Deoldify | UGCol |
|---|---|---|---|---|---|---|
| Rate | **47.20%** | 41.50% | 39.25% | 37.50% | 41.13% | 42.50% |

Quantitative results on COCO-Stuff using different structures.

| Structure | PSNR ↑ | SSIM ↑ | LPIPS ↓ | FID ↓ |
|---|---|---|---|---|
| AE | **25.89** | **0.928** | **0.146** | 14.15 |
| VAE | 23.21 | 0.905 | 0.179 | 11.76 |
| UGC w CA | 24.52 | 0.923 | 0.162 | 11.38 |
| PalGAN w rand ref | 20.88 | 0.883 | 0.240 | 9.64 |
| PalGAN w SA | 22.68 | 0.892 | 0.175 | 9.02 |
| PalGAN w PatchD [19,41] | 23.07 | 0.895 | 0.183 | 8.44 |
| PalGAN w BN | 22.36 | 0.895 | 0.209 | 9.97 |
| PalGAN w SPADE [35] | 24.06 | 0.916 | 0.167 | 7.90 |
| PalGAN wo $E(\hat{\mathbf{h}})$ | 24.58 | 0.924 | 0.149 | 8.17 |
| PalGAN | 24.56 | 0.924 | 0.148 | **7.70** |

Quantitative results on COCO-Stuff by ablating chromatic attention.

| G | L | PSNR ↑ | SSIM ↑ | LPIPS ↓ | FID ↓ |
|---|---|---|---|---|---|
| ✗ | ✗ | 21.93 | 0.902 | 0.203 | 9.90 |
| ✗ | ✓ | 24.52 | **0.924** | **0.146** | 9.97 |
| ✓ | ✗ | 23.32 | 0.907 | 0.174 | 8.34 |
| ✓ | ✓ | **24.56** | **0.924** | 0.148 | **7.70** |