



西安交通大学
XI'AN JIAOTONG UNIVERSITY

基于GAN的图片自动着色

2022年4月

Contents Title



一、背景与现状

二、简笔画着色

三、漫画页着色

图片自动着色

图片自动着色是对**黑白图片进行合理染色**的技术，可以应用与于恢复历史黑白图片、修复电影等领域。

着色在技术上是将三维**RGB颜色信息**分配给灰度图像的每个像素，使其在视觉上是合理的、可接受的。

为了降低的任务的复杂度，**亮度-色度的色彩空间**是比RGB更加常用的，如YUV和CIELAB。这两种色彩空间都将亮度从色彩空间中分离出来，使用已有的亮度信息预测剩余的两个颜色通道更加容易。

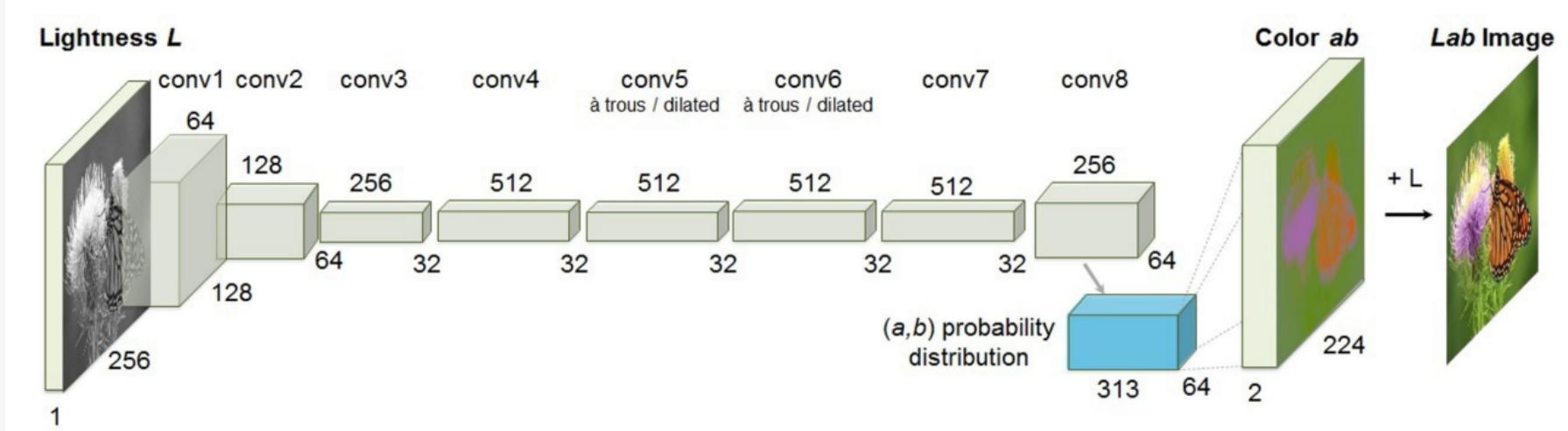
深度学习技术发展使我们能够用大量的源图像训练一个人工神经网络，因此可以自动学习真实物体自然对应的颜色。



基于CNN的图片着色

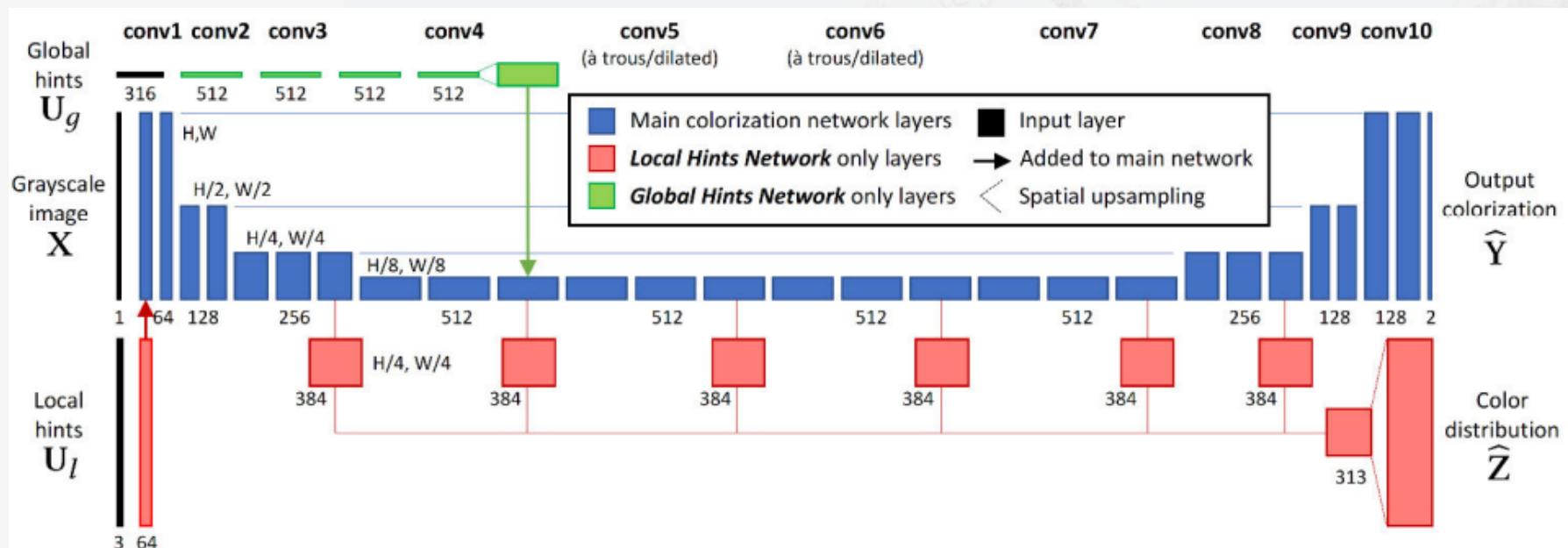
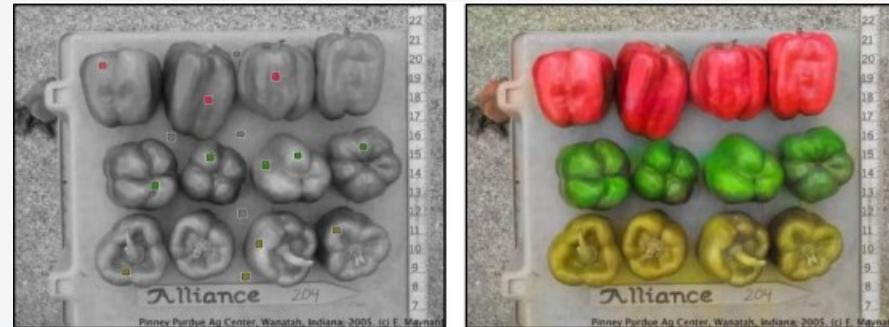
最为直接的方法是使用卷积神经网络。网络结构由多个卷积模块组成。该方法将CIELAB色彩空间的ab两通道编码为313个部分，通过每个像素亮度L预测颜色属于313个(a,b)元组之一的概率。

该方法认识到自然图像中不饱和色彩值的像素数量要高于饱和色彩值的像素数量，因此根据饱和和不饱和的像素分布调整了损失函数。



用户指导的着色神经网络

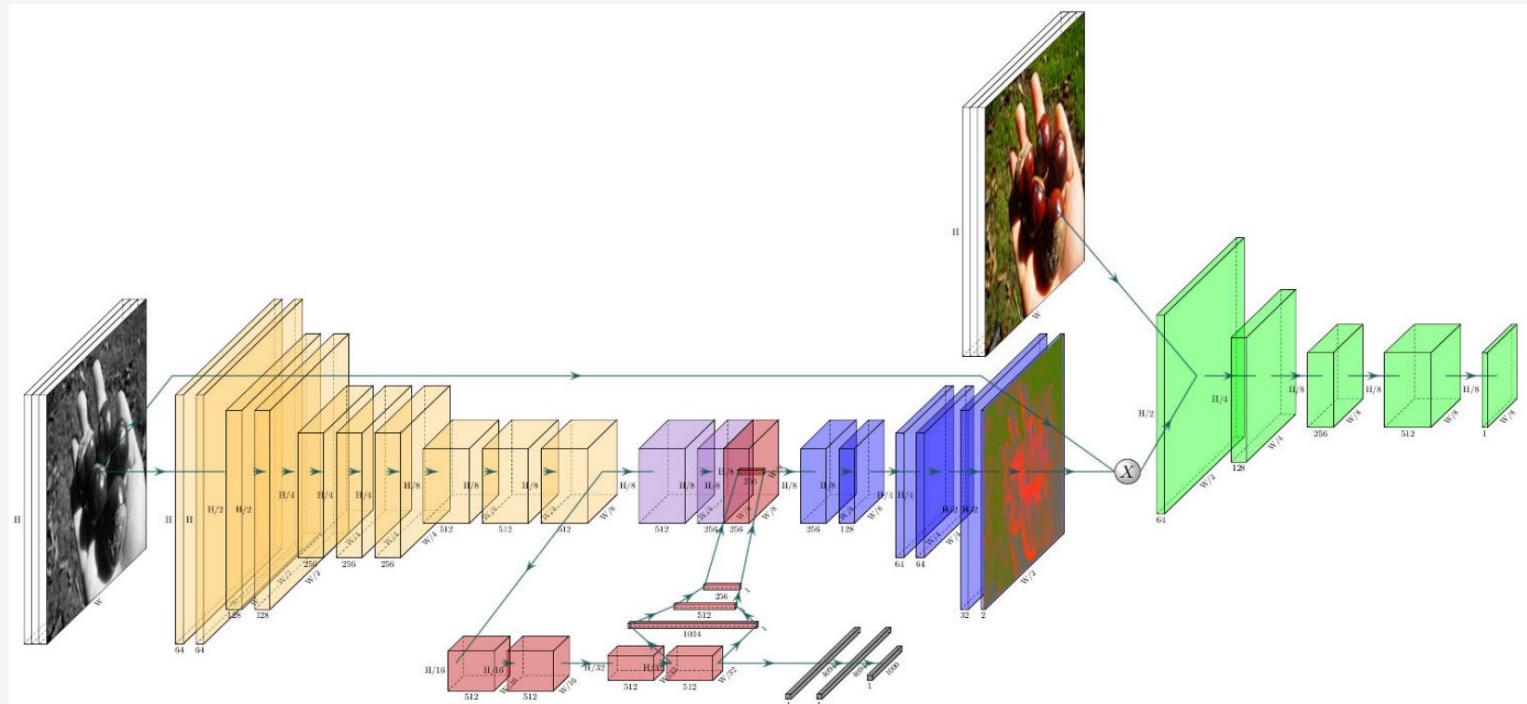
用户指导的着色神经网络将自动方法与交互方法结合。局部提示网络处理用户给出的染色提示预测图片的颜色分布；全局提示网络则处理所有图片的统计信息，将结果整合到主框架中。



R. Zhang, J.-Y. Zhu, P. Isola, X. Geng, A. S. Lin, T. Yu, and A. A. Efros, “Real-time user-guided image colorization with learned deep priors,” ACM Trans. Graph., vol. 36, no. 4, pp. 1–11, Jul. 2017.

基于GAN的图片着色

绿色的是鉴别器网络；左边是生成器网络，分为两部分：第一部分（黄色，紫色，红色，蓝色），预测图片的色度信息，第二部分（黄色，红色，灰色），预测图片的语义信息，输出类别分布向量，使得图片染色能够结合语义信息。

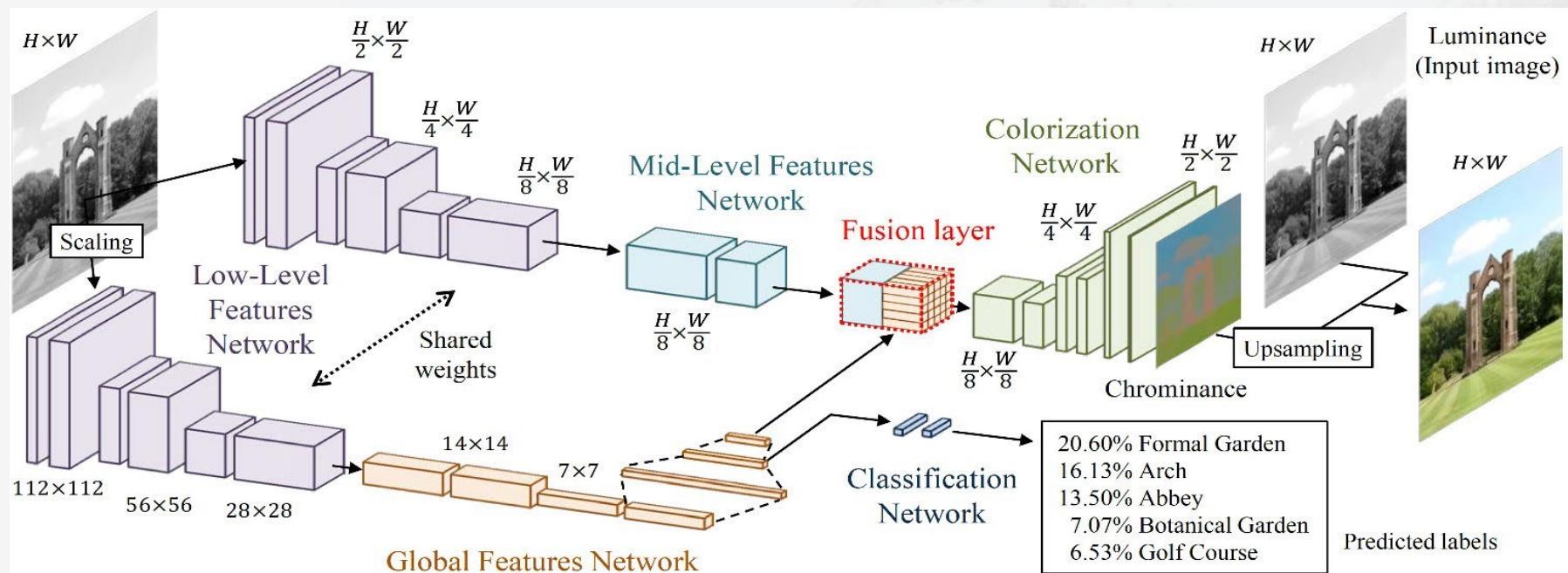


P. Vitoria, L. Raad, and C. Ballester, “ChromaGAN: Adversarial picture colorization with semantic class distribution,” in Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV), Mar. 2020, pp. 2445–2454.

多路径着色神经网络

Iizuka等人提出的多路径着色神经网络在一个大数据集上进行端到端训练，用于场景识别，并进行联合着色和分类丢失。其结构由四个部分组成：低级特征网络、中级特征网络、全局特征网络和着色网络。

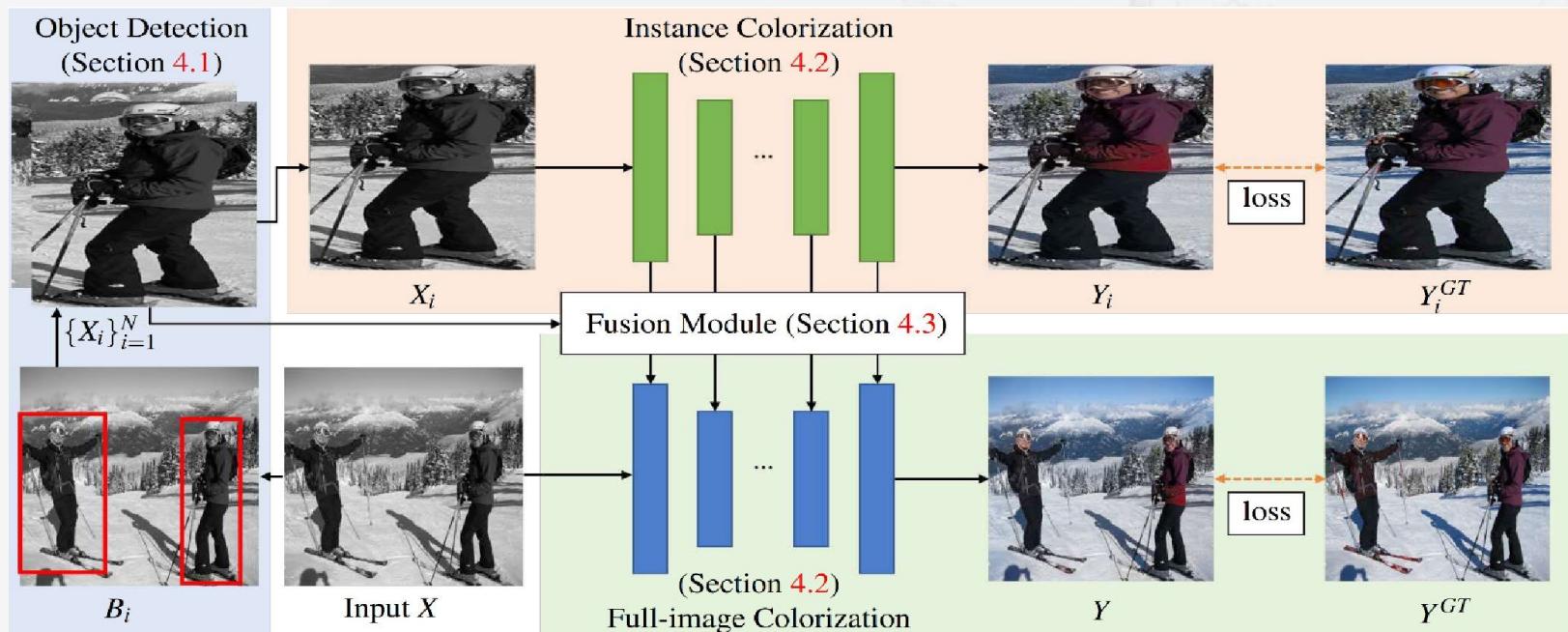
local信息和**global**信息融合在一起，低级用户输入与高级语义信息在网络中进行分割和后期融合，从不同路径获得着色特征。



基于范例的着色神经网络

现有图像着色模型在**背景杂乱的多目标图像**上效果不佳。

Su等人提出的基于范例的着色神经网络，利用现成的目标检测器来获取裁剪后的目标图像，并使用实例着色网络来提取**目标级别的特征**。我们使用另一个着色网络来提取**完整图像特征**，最后将目标级特征和图像级特征用**融合模块**相结合，获取最终颜色。



Contents Title



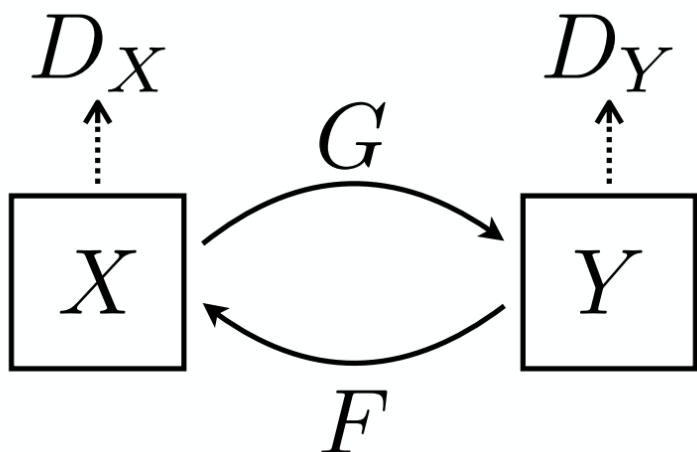
一、背景与现状

二、简笔画着色

三、漫画页着色

- 简笔画着色：
 - 基于Cycle GAN：具有两个判别器，两个生成器，这和传统意义上的GAN的结构不太一样，但是两者训练过程其实是大同小异的。
 - 模型一方面训练判别器 D_x 和 D_y ，使判别器对映射图像和真实图像的判别能力不断增强；另一方面训练生成器 F, G ，使其对 $X \rightarrow Y$ 的图像变换和 $Y \rightarrow X$ 的图像变换越来越成熟，变换图像和目标图像也越来越像，同时使得 $X \rightarrow Y \rightarrow X, Y \rightarrow X \rightarrow Y$ 的过程 loss 也越来越小。
 - 算法的整体损失函数Loss包含三部分： $X \rightarrow Y$ 的生成对抗损失， $Y \rightarrow X$ 的生成对抗损失以及循环一致性损失。
- 漫画页着色：
 - 基于GAN，模型包括四个网络：全局特征网络、低层特征网络、中层特征网络和彩色化网络。
 - 将参考图像作为输入，从每个参考图像中提取颜色特征，利用颜色特征，参考图像的颜色信息被自动合并以对输入图像进行着色。

Cycle GAN是一种图像风格迁移网络。传统的风格迁移模型（如Pix2Pix）需要训练数据成对，而GAN可以令数据集放宽，不需要两个域的数据有严格对应关系。图片着色就是将草稿转换为彩色图像，本质上是一种风格迁移任务。

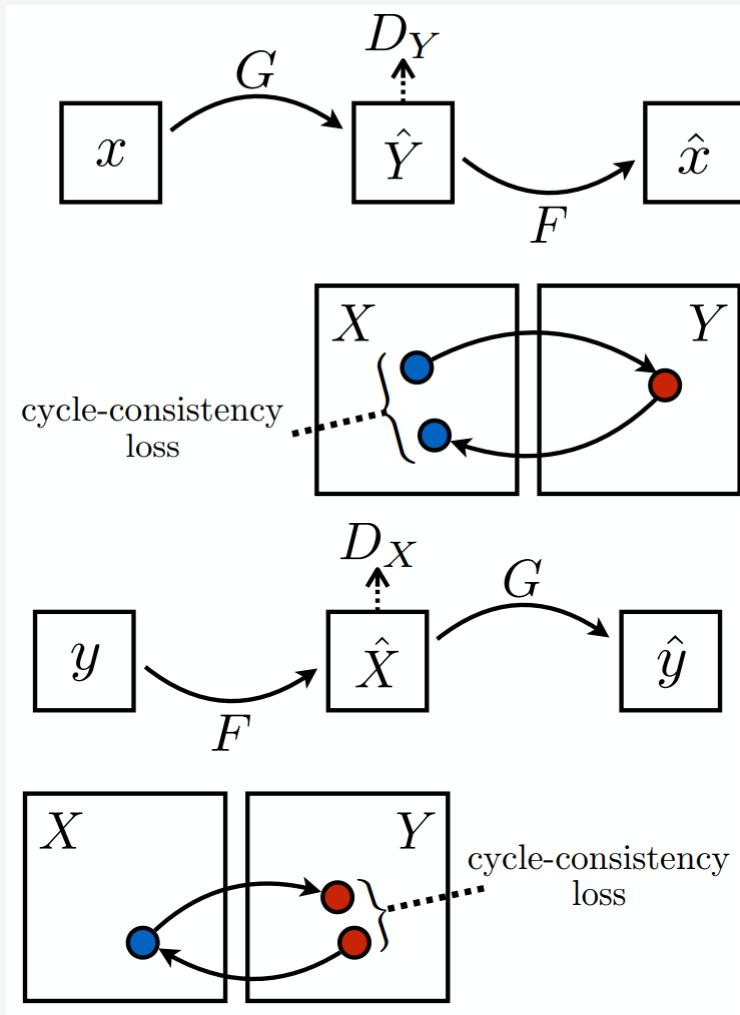


- GAN

训练出一个生成器\$G\$，输入一个\$X\$域的图像，输出一个\$Y\$域的图像，即\$G(x) = y', x \in X\$；训练出另一个生成器，输入\$Y\$域的图像，输出一个\$X\$域的图像，\$F(y) = x', y \in Y\$。为了达到这个目的，我们还需要训练两个判别器\$D_X\$和\$D_Y\$，分别判断两个生成器生成图像风格是否正确。

$$Loss_{GAN} = \mathcal{L}_{GAN}(G, D_Y, X, Y) + \mathcal{L}_{GAN}(F, D_X, X, Y)$$

Cycle GAN



判别器可以确保生成器生成 X 或 Y 域风格的图像，但其内容不一定与原图像相同。因此，作者提出通过使 $F(G(x)) = x$ 来确保 $G(x) = x$ 。

- Cycle

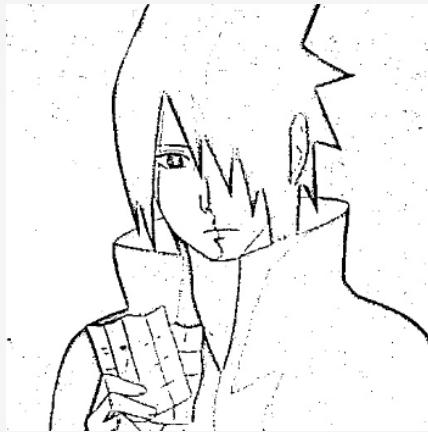
将 x 输入到生成器 G 生成的图像 y' 再放入生成器 F 中，产生的新图像 x' 和最开始的图尽可能的相似。

$$\begin{aligned} Loss_{cycle} = & \mathbb{E}_{x \sim p_{\text{data}}(x)} [\|F(G(x)) - x\|_1] \\ & + \mathbb{E}_{y \sim p_{\text{data}}(y)} [\|G(F(y)) - y\|_1] \end{aligned}$$

实验结果

➤ 复现示例

输入



输出



原图



Zhu J Y , Park T , Isola P , et al. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks[J]. IEEE, 2017.

Contents Title



一、背景与现状

二、简笔画着色

三、漫画页着色



之前方法的问题

“颜色模糊”：漫画每一页由多个角色分镜组成，必须在所有页面为相同角色染上相同的颜色。但是不同角色可能会有相同颜色，相同角色有不同颜色。



弱划分

强划分

参考着色

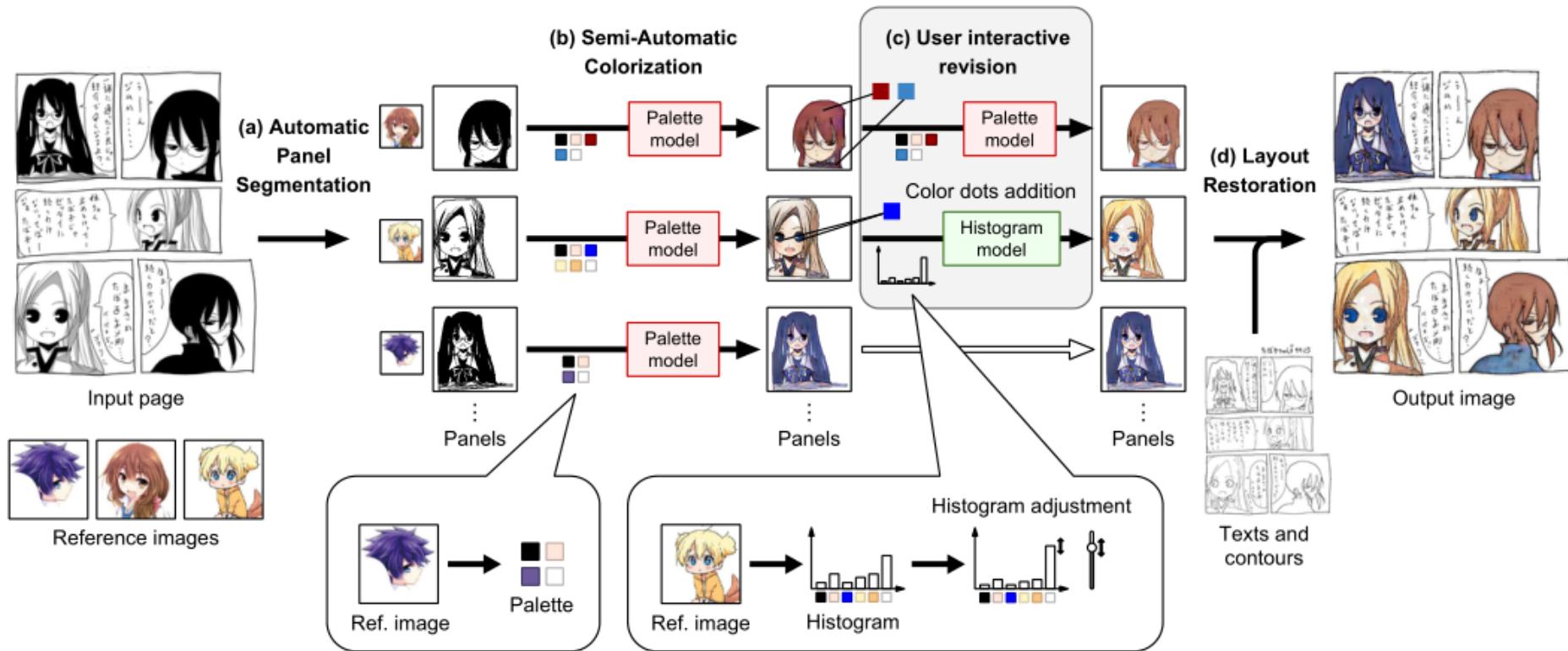
- Trade-off

着色越精细，同一角色差异越大；
同角色差异越小，细节越模糊。

- 参考着色法

1. 对整幅页面进行划分，识别角色。
2. 将参考图像作为附加输入。
3. 提取参考图像颜色特征进行迁移。
4. 可交互式修改。

方法框架

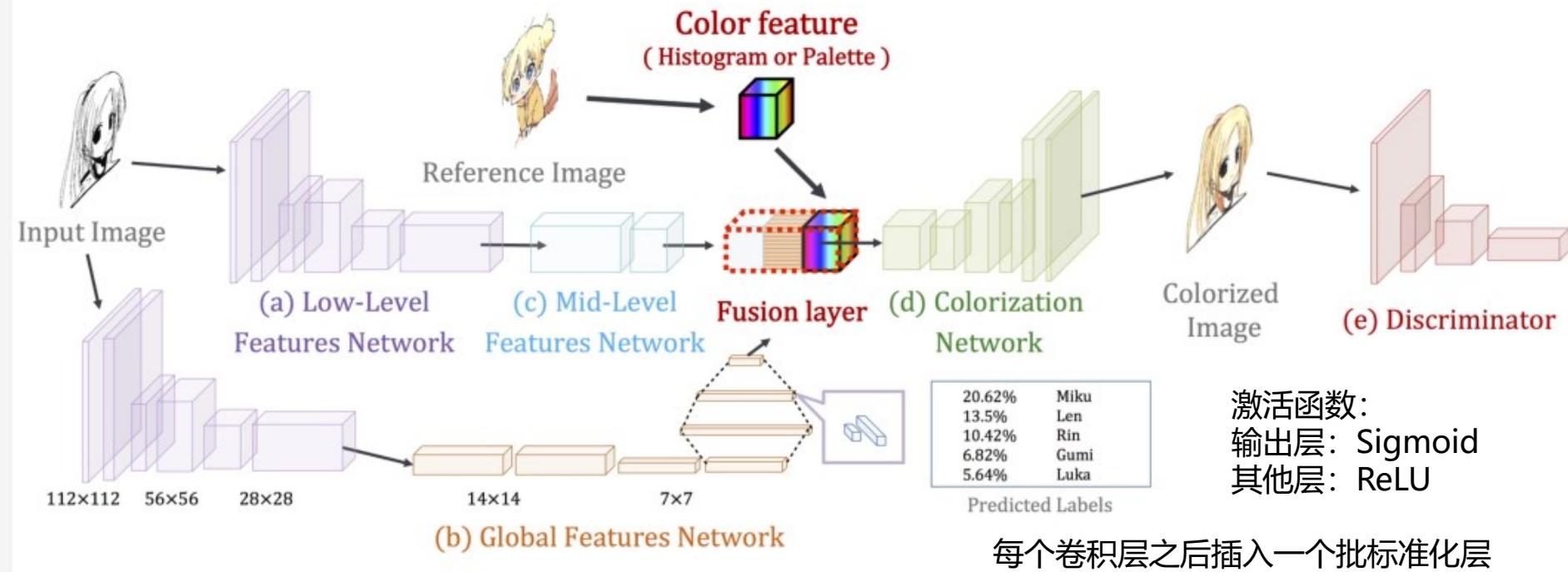


- A. 将漫画页面图像分割成分镜人物图像。
- B. 使用输入参考图自动着色（基于GAN）。
- C. 用户根据实际情况进行调整。
- D. 彩色分镜通过调整大小恢复原始布局。

- 1. 提取参考图像颜色特征。
- 2. 依靠名称训练角色分类。
- 3. 使用对抗性损失自动着色。

训练过程

着色步骤的网络架构



- A. 低级特征网络：提取低级视觉信息。
- B. 全局特征网络：提取角色分类信息。
- C. 中级特征网络：提取头发衣服等特征。
- D. 着色网络（CNN）。
- E. 鉴别网络：生成对抗损失训练着色模型。

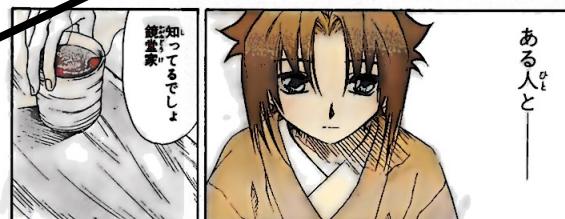
实验结果

复现示例

参考



输出



原图



Chie Furusawa, Kazuyuki Hiroshima, Keisuke Ogaki, and Yuri Odagiri. 2017. Comicolorization: semi-automatic manga colorization. In SIGGRAPH Asia 2017 Technical Briefs (SA '17).

实验结果

➤ 转换参考图

输出1



参考



输出1





西安交通大学
XI'AN JIAOTONG UNIVERSITY

谢谢大家

