

基于 PhotoWCT² 和多重注意力机制的 印花图案配色模型

汤逸烽¹, 蒋俊豪¹, 张 津², 朱 威¹

(1. 浙江工业大学 信息工程学院, 杭州 310023; 2. 杭州宏华数码科技股份有限公司, 杭州 310057)

摘要: 针对现有风格迁移算法无法满足印花图案配色工作对纹理保留和颜色风格迁移效果的要求, 设计了一种基于 PhotoWCT² 和多重注意力机制的印花图案配色模型; 该模型首先对 PhotoWCT² 网络添加了多尺度通道-局部注意力模块, 增强了模型对印花图案复杂纹理特征的提取能力; 其次引入频率注意力模块, 加强了频域中的颜色信息, 提升模型对颜色风格的迁移能力; 最后采用基于对比学习的内容损失函数, 增强模型对特征空间的建模能力, 提升模型对印花图案复杂纹理特征的重建能力; 实验结果表明, 与原始网络模型相比, 所设计模型在 SSIM、PSNR 和 Gram Loss 这 3 个客观指标上均表现更优; 此外, 与其它网络模型相比, 所设计模型的印花图案配色结果也具有更好的纹理保留和颜色风格迁移效果。

关键词: 风格迁移; 深度学习; 注意力机制; 对比学习; 印花图案配色

Color Matching Model for Printed Pattern Based on PhotoWCT² and Multiple Attention Mechanisms

TANG Yifeng¹, JIANG Junhao¹, ZHANG Jin², ZHU Wei¹

(1. College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China;
2. Hangzhou Atexco Digital Technology Co., Ltd., Hangzhou 310057, China)

Abstract: Aimed at the problem that existing style transfer models cannot meet the requirements of texture retention and color transfer for color matching for printed patterns, a color matching model for printed patterns based on PhotoWCT² and multiple attention mechanism is designed; Firstly, a Multi-scale Channel and Local Attention Module is introduced into the PhotoWCT² model to enhance the model's ability to extract complex texture features of printed patterns; Secondly, a Frequency-domain Channel Attention Module is introduced to strengthen the color information in frequency domain and to improve the model's ability to transfer color styles; Finally, the contrastive learning-based content loss function is used to enhance the model's ability to model feature space and improve the model's ability to reconstruct complex texture features of printed patterns. Experimental results show that compared with the original model, the designed model is better in three objective metrics: SSIM, PSNR, Gram Loss; In addition, the designed model is superior in printed patterns with texture retention and color style transfer to other models.

Keywords: style transfer; deep learning; attention mechanism; contrastive learning; color matching for printed patterns

0 引言

中国是全球纺织品制造强国, 在产业规模与供应链完整性方面具有显著战略地位, 近年来在纺织品产业始终保持着出口和竞争优势。印花图案作为纺织产品价值

提升的关键要素, 兼具艺术性与实用性, 其设计质量直接影响纺织品附加值和市场销量^[1]。在目前的印花图案设计行业中, 印花图案配色主要是在保留原始图案形态特征的前提下, 设计师使用图像编辑软件等工具通过调

收稿日期:2025-03-06; 修回日期:2025-05-27。

基金项目:杭州市重大科技创新项目(2022AIZD0077);国家自然科学基金联合基金重点项目(U24A20270)。

作者简介:汤逸烽(2000-),男,硕士研究生。

通讯作者:朱 威(1982-),男,博士,副教授,硕士生导师。

引用格式:汤逸烽,蒋俊豪,张 津,等. 基于 PhotoWCT² 和多重注意力机制的印花图案配色模型[J]. 计算机测量与控制, 2025, 33(10):251-258.

整颜色组合策略,实现单图案多色系的视觉风格转变,但这种方式效率低且配色结果受设计师主观影响大。近些年人工智能技术的快速发展,特别是基于深度学习的图像风格迁移算法的兴起,为优化传统印花图案配色工作流程提供了创新路径。

传统的风格迁移算法通过捕捉图像的颜色、纹理、笔触等特征,然后使用渲染算法将这些特征转移到新的图像上^[2],能够实现某些特定风格的迁移,但在风格表达的灵活性、多样性和图像结构完整性上存在局限性。随着深度学习理论的快速发展,基于深度学习的风格迁移模型开始被广泛应用印花图案设计行业。基于深度学习的风格迁移算法^[3]可以分为两类:第一类是基于图像优化的风格迁移算法;第二类是基于模型优化的风格迁移算法^[4],因其计算效率高,实时性好,目前逐步应用于印花图案设计行业,帮助设计师完成印花图案配色工作。早期基于模型优化的风格迁移算法也存在生成图像风格有限的问题,但随着文献[5]提出的自适应实例归一化方法(AdaIN, adaptive instance normalization),该方法使用特征的均值和方差进行风格转化,解决了单风格的风格迁移模型在扩展风格时需要重新训练新模型的问题,实现了任意风格的风格迁移算法。文献[6]提出了基于零相位成分分析(ZCA, zero-phase component analysis)的白化和着色(WCT, whitening and coloring transforms)模块,该方法在白化步骤使用ZCA消除风格图冗余风格,在着色步骤使用颜色变换^[7]对齐内容特征图和风格特征图的协方差矩阵,完成内容图和风格图特征的融合和转化,实现了内容和风格更好的分离,保留了更多的内容结构。为解决任意风格迁移算法中存在的伪影和空间失真问题,文献[8]提出了PhotoWCT风格迁移模型,该模型在WCT工作的基础上将编码器中下采样层的掩码跳接至解码器上采样层,并引入高斯亲和矩阵^[9]约束像素对、区域对之间的风格关系,从而优化了模型的风格迁移效果,但存在计算成本高昂,单次迁移耗时长等问题。文献[10]提出了WCT²风格迁移模型,该模型将编码器的上采样层和下采样层更换为小波上采样层和小波下采样层,利用小波变换和小波反变换的可逆性精确重建原始信号,解决了特征压缩重建的损失不可恢复的问题,在小波下采样层中低频信息传递给下一级编码器,高频信息跳接至解码器的小波上采样层,从而最大程度上恢复图像的原始特征,但计算成本和存储成本高昂。文献[11]提出CAP-VSTNet风格迁移模型,该模型由可逆残差网络^[12]和无偏线性变换模块组成,并加入了拉普拉斯损失^[13]进一步优化因线性变换引起的像素亲和度损失问题,在纹理保留和风格迁移效果取得了较好提升。文献[14]提出了PhotoWCT²风格迁移模型,在WCT²风

格迁移模型基础上实现了更好的风格迁移效果,该模型将WCT²的小波上采样层被替换为高频跳跃残差模块(平均池化层和上采样层),将小波上采样层的操作减少到原来的四分之一。此外,PhotoWCT²将WCT转换嵌入到解码器的多个层级中实现由粗到细的风格迁移并且采用分块训练的策略^[15],因此该模型相较于其它模型具有更好的风格迁移效果。

尽管现有风格迁移算法在照片图像风格化等方面取得了显著的进展,但面对内容纹理复杂、颜色风格多样的印花图案,这些算法在纹理保留和颜色风格迁移效果上还是存在局限性。为了解决上述问题,本文设计了一种基于PhotoWCT²和多重注意力机制的印花图案配色模型。该模型通过引入多尺度通道-局部注意力模块(MCLA, multi-scale channel and local attention)提升纹理特征提取能力;引入频率通道注意力模块(FCAM, frequency-domain channel attention module)增强了颜色风格迁移能力;引入基于对比学习的内容损失函数提高重建质量,进一步加强纹理保留。

1 PhotoWCT²模型与注意力机制

1.1 PhotoWCT²模型

PhotoWCT²是一种基于深度学习的风格迁移网络模型,其将风格迁移任务分解为图像重建和风格迁移两个部分。

1) 图像重建: PhotoWCT²使用VGG19网络构建编码器-解码器结构。编码器通过多层级卷积(Relu1_1至Relu4_1)提取图像全局和局部结构特征,并插入高频残差跳跃模块分离高频信息并跳接至对应解码器的上采样层,低频信息传递给下一层解码器,增强了特征表达的完整性。解码器则采用与编码器对称的卷积层与上采样层,结合编码器传递的高频信息逐级重建图像。

2) 风格迁移: PhotoWCT²模型使用基于ZCA的WCT模块实现内容图与风格图的特征解耦与统计对齐。基于ZCA变换的白化与上色模块首先通过对输入内容图和风格图的协方差矩阵进行奇异值计算获得白化后的内容图矩阵 F_c 和风格图着色矩阵 C ,如公式(1)所示:

$$\begin{aligned} F_c &= U_c \Lambda_c^{-1/2} U_c^T \bar{f}_c \\ C &= U_s \Lambda_s^{1/2} U_s^T \end{aligned} \quad (1)$$

其中: Λ_c, Λ_s 表示对应内容图和风格图特征值对角矩阵, U_c, U_s 表示对应内容图和风格图特征向量矩阵, $\bar{f}_c$ 表示中心化后的内容图特征向量。

最后加上风格图特征均值 μ_s 得到风格转化后的特征 F_s ,如公式(2)表示:

$$F_s = C F_c + \mu_s \quad (2)$$

1.2 注意力机制

传统的卷积神经网络因感受野受限和固定的权重分配导致存在全局信息获取不足、特征冗余等问题^[16]。为解决上述问题, 注意力机制被引入卷积神经网络, 注意力机制通过动态调整不同区域或通道的权重, 使模型能关注关键信息, 从而提升图像分类、检测等任务的精度和效果。卷积注意力模块 (CBAM, convolutional block attention module)^[17] 目前在深度神经网络领域应用较为广泛, 其通过通道注意力建模通道之间的相互依赖性; 通过空间注意力建模像素之间的联系。通过对通道和空间的混合注意力, CBAM 实现了对重要信息的捕捉和关注。除了 CBAM 之外, 还有关注全局的自注意力机制^[18] 以及关注局部信息的局部注意力机制^[19] 等的注意力模块被应用于卷积神经网络。

2 本文印花图案配色模型

本文设计模型的网络结构如图 1 所示。

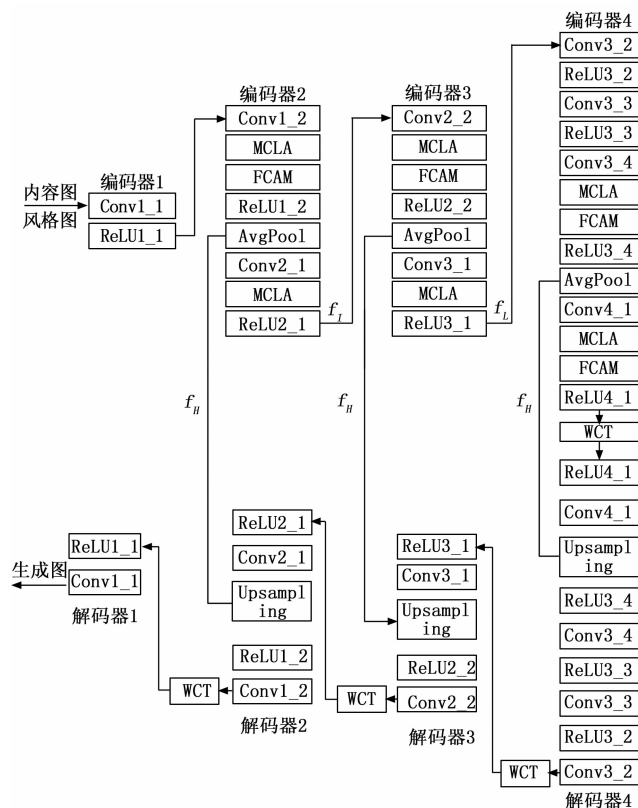


图 1 印花图案配色模型网络结构

2.1 总体结构

如图 1 所示, 为了防止出现过拟合以及不必要的计算开销, 所设计的模型将 MCLA 模块插入到每层编码器的下采样卷积层和 ReLU 层之间。这样设计的好处是, MCLA 模块可以直接处理原始卷积输出同时避免 ReLU 的截断效应影响注意力机制的效果, 更有利于调整特征的分布, 并且在多个层级插入 MCLA 模块可以

在多尺度融合时保留更多细节, 加强改进后的模型对印花图案复杂纹理特征、纹理层次等的提取和保留能力。模型将 FCAM 模块插入到高频跳跃残差模块之前, 可以有效避免平均池化操作引发的空间分辨率降低以及图像过度平滑导致的低频信息减弱, 保证对主要颜色信息的有效加强以及对低频颜色信息与高频纹理信息关联性的学习。此外, 每层编码器都插入 FCAM 模块能有效利用频域分析捕捉跨通道的长程依赖关系, 并且提高网络的鲁棒性。

2.2 多尺度通道一局部注意力模块

印花图案往往由花朵、装饰纹理等不同元素组合构造, 元素大小不一, 纹理信息复杂且多采用拼接等方式构图, 存在较多冗余信息; 并且在纹理结构上, 印花图案纹理丰富, 层次感较强。传统的卷积神经网络由于受感受野的限制, 不能充分捕捉甚至丢失印花图案复杂的纹理信息; 此外大量冗余信息也因固定权重导致重要信息得不到加强, 最终导致生成印花图案存在纹理保留不足, 如出现图像伪影和边缘杂色问题。针对以上问题, 本文设计了 MCLA 模块, 该模块由多尺度特征提取融合模块、通道注意力模块和局部注意力模块组成, 具体结构如图 2 所示。

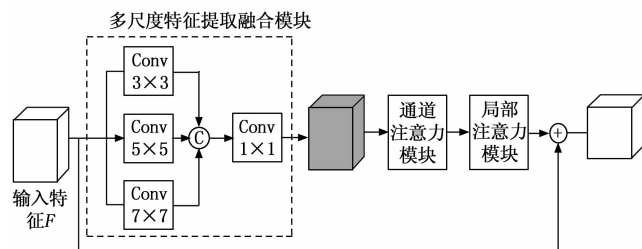


图 2 MCLA 模块结构

针对卷积神经网络固定卷积核感受野受限问题, MCLA 模块使用多尺度特征提取融合模块, 增强模型对印花图案不同大小的复杂纹理信息的捕获和表达能力和对纹理层次的处理能力。MCLA 使用 3×3 , 5×5 , 7×7 的卷积核组合提取输入特征获得多尺度特征。该卷积核组合不仅能够扩大感受野, 并且能使模型获得最佳感受野。为避免传统多尺度特征融合方法 (特征拼接、加权求和) 可能带来的信息冗余、计算开销大和不够灵活的问题, 使用 1×1 卷积对拼接起来的多尺度特征进行融合, 通过训练时对卷积参数的调整实现对特征融合权重的调整, 输出融合特征。

如 1.2 节所述, 针对印花图案可能存在的大量冗余信息, MCLA 模块使用通道注意力模块调整融合特征的通道权重, 过滤削弱冗余无用的特征, 增强主要纹理特征信息, 并将经过调整后的特征 F 输入局部注意力模块。

在传统的 CBAM 模块中, 空间注意力模块通过平均池化与最大池化操作对输入特征进行处理。具体来说, CBAM 空间注意力模块通过公式 (3) 进行处理:

$$\begin{aligned} \text{AvgPool}(F) &= \frac{1}{H \cdot W} \sum_{i=1}^H \sum_{j=1}^W F_{i,j} \\ \text{MaxPool}(F) &= \max_{i,j} (F_{i,j}) \end{aligned} \quad (3)$$

其中: AvgPool 和 MaxPool 分别为平均池化操作和最大池化操作, $F_{i,j}$ 表示输入特征的每个像素点, H 和 W 图像的高度和宽度。然后, 空间注意力模块将这些池化结果进行融合, 得到空间注意力图, 并将其与输入特征进行加权操作。该方法虽有效, 但在处理细小纹理特征时可能丢失高频信息, 且无法捕捉到局部细节。

为了解决这一问题, MCLA 模块用局部注意力模块替代了空间注意力模块。局部注意力模块通过引入 Softmax 操作和温度参数获取局部重要区域以强化对细小纹理和边缘纹理的重要性, 具体公式如下:

$$S(F) = \text{Conv}_{3 \times 3} \left(\frac{e^{\lceil x - \max(F) \rceil / T}}{\sum_i e^{\lceil x_i - \max(F) \rceil / T}} \right)_{\times 2} \quad (4)$$

其中: $S(F)$ 表示 Softmax 归一化后的权重, T 是温度参数, 也是可训练参数, i 表示像素位置, x_i 表示对应位置的像素值, $\text{Conv}_{3 \times 3}$ 表示卷积核为 3 的卷积操作, $\times 2$ 表示进行两次卷积操作。通过 Softmax 和可训练温度参数 T , 局部注意力模块能够更加精确地聚焦于细小的纹理和边缘信息, 而不依赖于平均池化或最大池化所带来的全局信息损失。同时凭借卷积操作, 实现了对像素局部领域中更强提取和更加细致的调整, 进一步细化模块对局部区域重要性的评分。此外由于减少了池化操作, 模型在训练过程中能够更加稳定, 收敛速度也得到提升。

2.3 频率通道注意力模块

PhotoWCT² 风格迁移模型虽然有较强的风格转化能力, 但在处理印花图案时易出现迁移后颜色变化较小, 边缘杂色, 颜色迁移错误等问题。其原因在于高频残差跳跃模块在空间域分离高低频信息时, 使用平均池化导致图像特征过度平滑使颜色风格在跨尺度传递中减弱了局部适应性, 同时也削弱了低频信息中颜色分布与高频信息中纹理走向的关联性, 降低了模型的颜色风格迁移能力。

离散余弦变换 (DCT, discrete cosine transform) 是针对这一问题的一种可行解决方案。离散余弦变换通过一组不同频率和幅值的余弦函数和近似表达一副图像, 将图像空间域信息转化为频域信息, 为颜色风格信息提供了更优的频域表示。二维离散余弦变换公式如式 (5) 所示:

$$X(u, v) = \alpha(u)\alpha(v)$$

$$\sum_{x=0}^{H-1} \sum_{y=0}^{W-1} F(x, y) \cos \left[\frac{\pi u}{H} \left(x + \frac{1}{2} \right) \right] \cos \left[\frac{\pi v}{W} \left(y + \frac{1}{2} \right) \right] \quad (5)$$

其中: $X(u, v)$ 为变换后的频率谱, $F(x, y)$ 为输入特征图的像素值, H 和 W 是输入特征的宽和高, u 和 v 对应 H 和 W 为频率谱的水平坐标和垂直坐标, $\alpha(u)$ 和 $\alpha(v)$ 为归一化因子, 值分别为 $\sqrt{\frac{2}{H}}$ 和 $\sqrt{\frac{2}{W}}$ 。

从公式 (5) 可以看出, 特征图的每个像素值可以视为不同频率余弦基函数的加权和, 实现了特征图空间域向频域的转化, 当 u 和 v 较小时, 基函数频率较低主要捕获图像中的低频信息, 如图像的整体颜色和形状。随着 u 和 v 的增大, 基函数频率增高, 捕获更多图像中的高频信息, 如图像的细节和纹理。由于图像颜色变化通常较为平滑, 低频分量通常包含了大部分能量, 约占图像的 95% 以上, 而高频分量则主要携带图像的细节信息。

此外这种频域分离使得低频分量能够更好地保留图像颜色信息, 确保颜色的一致性和连续性, 同时高频部分保持了图像的细节和纹理结构, 有助于图像风格迁移时的精确调整。通过离散余弦变换, 颜色信息与纹理特征得以有效分离, 避免了颜色信息和纹理特征的丢失, 从而增强了模型对颜色信息和纹理特征相关性的表达能力。

二维离散余弦变换的逆变换如公式 (6) 所示:

$$\begin{aligned} F(x, y) &= \alpha(u)\alpha(v) \\ \sum_{u=0}^{H-1} \sum_{v=0}^{W-1} X(u, v) \cos \left[\frac{\pi u}{H} \left(i + \frac{1}{2} \right) \right] \cos \left[\frac{\pi v}{W} \left(j + \frac{1}{2} \right) \right] \end{aligned} \quad (6)$$

从公式 (6) 可以看出, 二维离散余弦逆变换使用的是与二维离散余弦变换相同的余弦基函数, 只是频率方向反转, 所有的频率成分都能得到准确的还原。通过对频域各个分量的加权叠加, 逆变换能够高效恢复出图像的空间域特征, 确保了图像的整体信息和细节特征不被丢失。

基于上述分析, 本论文设计了 FCAM 模块, 其结构如图 3 所示。该模块首先对输入特征图向量图 F 进行通道独立的二维离散余弦变换, 将空间域特征投影至频域空间获得频域特征向量图 F_{freq} 。其次将 F_{freq} 输入设计的轻量级注意力模块快速高效生成颜色通道权重, 轻量级注意力模块首先通过平均池化加强主要颜色特征的响应值; 然后通过 4 个 1×1 卷积构成的瓶颈结构, 通过先升维再降维减少计算量的同时跨通道提取特征信息, 并对重要的颜色频域信息进一步增强同时加强模型对颜色与纹理相关性学习; 最后经过 Sigmoid 激活函数并于频域特征向量图相乘获得增强获得频域特征向量图。最

后将频域特征向量图进行逆二维离散余弦变换后与原始输入 F 残差连接, 加强主要颜色信息的同时保证了印花图案颜色信息不会因过分加强而失真。

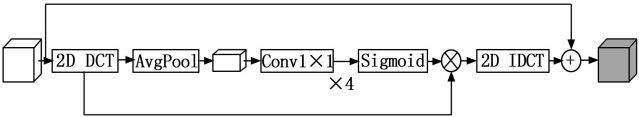


图 3 FCAM 模块结构

2.4 基于对比学习的内容损失函数

原始 PhotoWCT² 网络模型主要依赖于像素级重建损失与感知损失共同优化模型参数, 从而提升图像重建效果。然而, 两种损失都存在一定局限性。首先, 像素级损失在捕捉图像结构关联方面能力有限, 在处理印花图案时易造成纹理信息及其层次结构丢失; 其次, 感知损失通常依赖固定预训练模型提取图像特征, 这些特征本质上是对图像统计属性的泛化, 其表示空间可能与图像重建任务的目标域存在分布差异, 在印花图案风格迁移任务中具体表现为图案边缘伪影和杂色问题。

本文采用对比学习的内容损失函数^[20]来解决原始 PhotoWCT² 模型在损失函数方面存在的局限性。该方法的核心在于通过双路特征提取结构将输入图像与重建结果映射至统一的隐空间, 并建立多层次对比关系, 以实现重建图像内容表达能力的增强。具体来说, 在局部尺度上, 通过随机采样构建正样本对, 引导模型保持图像细节纹理一致性; 在全局尺度上, 利用记忆库维护历史重建样本的负例队列, 通过最大化正负样本间距隐式学习重建任务的本质特征分布。模型训练时, 图像 I 以及增强图像 I^+ (随机调整图像大小、裁剪和旋转) 输入到双路特征提取结构获得 K 维向量 q, q^+ ; 同时, 数据集中的视为负样本的其他印花图案图像形成对比 I^- 通过同样操作获得 K 维向量 q^- 。为防止出现特征表征空间坍塌, 对向量 q^-, q, q^+ 做归一化处理。基于对比学习的内容损失函数如公式 (7) 所示:

$$Loss_{contra} = - \sum_{i=1}^M \log \frac{\exp(q_i \cdot q_i^+ / \tau)}{\exp(q_i \cdot q_i^+ / \tau) + \sum_{j=1}^N \exp(q_i \cdot q_j^- / \tau)}$$

(7)

其中: $\cdot$ 表示两个向量的点积, $\exp$ 为指数运算。 τ 为温度比例因子, 在本文中设置为 0.07。

双路特征提取结构是对比学习内容损失的核心, 包含风格提取网络与映射网络。风格提取网络负责提取图像浅层纹理特征到深层语义特征, 其采用 VGG19 为骨干网络, 并将 VGG19 前 70 层拆分为 6 个子模块 (从 ReLU1_1 至 ReLU5_1), 各层级模块提取特征经过平均池化与最大池化处理后拼接, 再通过各层级模块对应的 1×1 卷积层对特征进行通道压缩生成紧凑风格编码。

映射网络负责将风格提取网络生成的紧凑风格编码映射到高维特征空间, 因此针对风格提取网络的 6 个子模块分别设计了独立的全连接映射网络将特征映射到统一的 2 048 维空间后再对向量做归一化操作, 用于后续对比学习。

3 实验结果与分析

3.1 数据集

本实验数据集包含 400 000 张印花图案, 均为合作印花数据企业设计生产中使用到的印花图案; 数据集中涵盖组合类花卉、小碎花、几何、抽象等多种类别图像, 由设计师统一标注, 各类别印花图案分布如表 1 所示, 表中的其他类别是指除组合类花卉、小碎花、几何、抽象外的其他类别。数据集图像平均分辨率为 $1\,500 \times 1\,000$, 训练时训练集和测试集比例为 8 : 2。

表 1 数据集印花图案类别分布情况

印花图案类别	比例/%
组合类花卉	29.22
抽象	12.79
小碎花	8.92
装饰纹样	6.06
几何	5.82
其他	37.91

3.2 实验环境与评价指标

实验环境为 Linux 服务器, 操作系统使用 Ubuntu 22.04 LTS; 使用 Pytorch 搭建并训练模型, Pytorch 版本为 2.0.1。服务器的 CPU 为酷睿 i9-14900K, 共有 32 个线程和 24 个核心; 图形处理器使用 RTX 4090 显卡, 显存大小为 24 G, CUDA 版本为 12.2; 服务器运行内存为 128 G。

模型训练过程中, 将输入印花图案尺寸缩放为 512×512 并随机裁剪至尺寸为 256×256 ; 训练时优化器使用 Adam 优化器, 批次大小为 8, 学习率为 0.000 1, 训练方式采用分块训练方式, 每个块训练轮数为 10。

为了评估本文设计模型的有效性, 使用风格迁移算法常用的结构相似度 (SSIM, stuctural similartiy index measure), 峰值信噪比 (PSNR, peak signal-to-noise ratio) 和格拉姆矩阵损失 (Gram Loss) 3 个客观指标进行评估。

PSNR 是通过计算两幅图像的像素级均方误差, 结合信号的最大可能功率与噪声功率的比值, 量化图像的保真度, 其定义如公式 (8) 所示:

$$PSNR = 10 \times \log_{10} \left(\frac{MAX^2}{MSE} \right)$$

(8)

其中: MAX 为图像像素值的最大范围, MSE 为两幅图像像素值的均方误差。PSNR 反映了模型处理图

像质量的变化,在视觉任务中广泛应用,但是 PSNR 无法充分反映人类视觉系统对图像质量的感知^[21],仅用 PSNR 无法完全衡量风格迁移后印花图案的质量。因此加入 SSIM,使用 SSIM 和 PSNR 一起衡量评价风格迁移前后印花图案内容损失。

SSIM 是通过比较两幅图像的亮度分布、对比度和结构信息来评估它们之间的相似度。其定义如公式 (9) 所示:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (9)$$

其中: x, y 为输入的两幅图像, μ_x, μ_y 为对应图像的均值, σ_x, σ_y 为对应图像的标准差, σ_{xy} 为图像 x 和图像 y 的协方差, C_1, C_2 为常数项。

Gram Loss^[22]是通过计算两张图像特征的格拉姆矩阵之间的平均绝对值误差来衡量图像风格上的差异,使用 Gram Loss 评价风格迁移后印花图案风格迁移效果,其定义如公式 (10) 所示:

$$Gram Loss = \frac{1}{n} \sum_{i=1}^n |G_{s,i} - G_{cs,i}| \quad (10)$$

其中: G_s 为风格图特征的格拉姆矩阵, G_{cs} 为配色结果特征的格拉姆矩阵,格拉姆矩阵计算公式如式 (11) 所示:

$$G = \frac{1}{HW} F_{flat} \cdot F_{flat}^T \quad (11)$$

其中: H, W 为特征图的长和宽, F_{flat} 为展平后的特征, F_{flat}^T 为展平后的特征的转置。

3.3 消融实验

为了验证本文 MCLA 和 FCAM 模块的有效性,进行消融实验,实验结果如图 4 所示。在 (a) 组结果中,本文设计模型与原始模型 PhotoWCT² 相比,图像边缘伪影问题得到解决,花朵花瓣上色更加符合风格图中的颜色逻辑,颜色迁移效果更好且花朵纹理更加清晰。在 (b) 组结果中,本文设计模型风格迁移后得到的整体图像伪影问题得到解决,花朵、花瓣和花蕊颜色上色效果相较于原始模型 PhotoWCT² 更好。综上所述,本文设

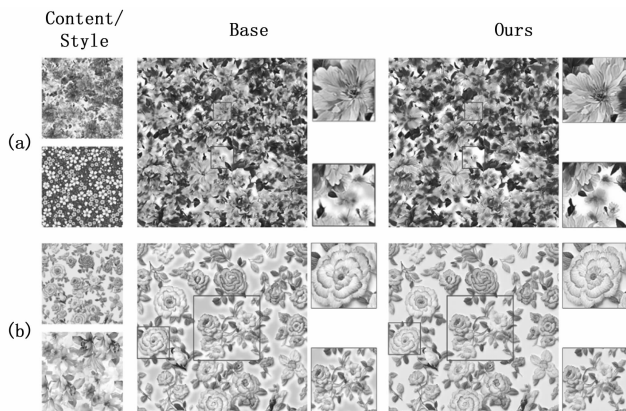


图 4 消融实验

计模型在增加这两个模块之后,图案边缘的伪影和杂色问题得到缓解,颜色风格更加偏向于风格图,有更好的纹理保留度和颜色风格迁移效果。

为了进一步验证 MCLA 模块的有效性,将使用 MCLA 模块的 Photowct², 记为 w/MCLA。将正常架构的 Photowct², 记为 w/o MCLA, 在其余参数和网络结构均相同的情况下进行比较,实验结果如图 5 所示。从实验结果可以看出, w/MCLA 相较于 w/o MCLA 有更好的细节和纹理效果。w/MCLA 红框中的外部花纹样式以及内部花卉装饰相较于 w/o MCLA 红框中对应的图案样式边缘纹理保持更好,视觉效果上更清晰。w/MCLA 蓝框中的几何装饰边缘相较于 w/o MCLA 更清晰,同时几何装饰包裹的花纹装饰相较于 w/o MCLA 边缘杂色更少。

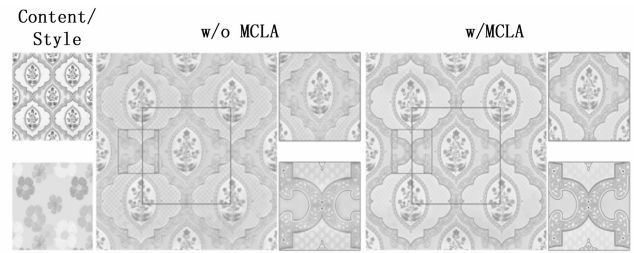


图 5 MCLA 模块消融实验

为进一步验证 FCAM 模块的有效性,将使用 FCAM 模块的 Photowct², 记为 w/ FCAM。将正常架构的 Photowct², 记为 w/o FCAM, 在其余参数和网络结构均相同的情况下进行比较,实验结果如图 6 所示。从实验结果可以看出, w/ FCAM 相较于 w/o FCAM 有更好的颜色风格迁移效果。w/ FCAM 红框中的整体背景颜色的迁移效果相较于 w/o FCAM 红框中对应部分的迁移效果更好,视觉效果上在符合内容图背景织布粗糙纹理颜色分布的情况下,将风格图的背景颜色较好的迁移到结果图中。w/ FCAM 蓝框中的仿国风花样相较于 w/o MCLA 颜色更符合风格图中主体内容的颜色,同时得益于背景颜色迁移效果更好,整体颜色层次和颜色视觉效果上更好。

以上 3 个消融实验通过主观对比,证明了 MCLA 模块和 FCAM 模块的有效性,为进一步证明两个模块有效性,从客观指标上对两个模块进行评估。本文选择

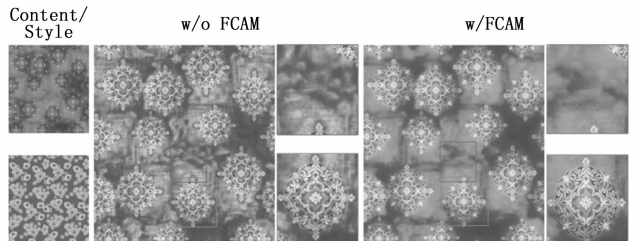


图 6 FCAM 模块消融实验

SSIM, PSNR 和 *Gram Loss* 三个定量指标, 其中 SSIM 和 PSNR 的计算基于输入印花图案内容图与风格迁移后的图像之间的比较, *Gram Loss* 的计算基于输入印花图案风格图与风格迁移后的图像之间的比较。训练集采用 3.1 节所提印花图案数据集进行训练, 输入印花图案统一缩放至 512×512 像素并随机裁剪至 256×256 像素, 输入批次大小为 16; 各个模型均采用 Adam 优化器, 学习率为 0.001, 进行 50 轮训练。实验结果如表 2 所示。

表 2 不同风格迁移模型指标对比

模型	SSIM	PSNR	<i>Gram Loss</i>
PhotoWCT ²	0.714	10.934	9.483
w/MCLA	0.755	11.030	9.484
w/FCAM	0.713	10.922	9.750
OURS	0.756	11.033	9.753

实验结果显示, w/MCLA 在 SSIM 和 PSNR 两个指标上均强于 PhotoWCT², 表明 MCLA 模块能够捕捉到印花图案复杂和多层次的纹理信息, 处理纹理信息之间联系, 增强模型对内容保留能力。w/ FCAM 在 *Gram Loss* 这个指标上强于 PhotoWCT², 表明 FCAM 模块能够增强模型对风格图颜色的提取能力和保持能力, 提升模型风格迁移效果, 增强模型颜色风格迁移能力。

3.4 客观指标对比

为验证本文设计的模型在纹理保留和颜色风格迁移效果的有效性, 各个模型在 SSIM, PSNR 和 *Gram Loss* 三个客观指标上进行比较, 其中 SSIM 和 PSNR 的计算基于输入印花图案内容图与风格迁移后的图像之间的比较, *Gram Loss* 的计算基于输入印花图案风格图与风格迁移后的图像之间的比较。各模型训练超参数参考 3.3 节设置, 实验结果如表 3 所示。

表 3 不同风格迁移模型指标对比

模型	SSIM	PSNR	<i>Gram Loss</i>
WCT ²	0.659	10.668	9.262
PhotoWCT ²	0.714	10.934	9.483
CAP-VSTNet	0.691	11.433	9.561
OURS	0.756	11.033	9.753

实验结果显示, 在印花图案风格迁移结果中, 本文设计模型的 SSIM 为 0.756, PSNR 为 11.433 dB, *Gram Loss* 为 9.753, 都高于 PhotoWCT² 和 WCT² 模型。与 CAP-VSTNet 模型相比, 本文设计模型仅 PSNR 低了 0.4 dB, 但 SSIM 和 *Gram Loss* 都明显更高, 具有更好的纹理结构和风格迁移效果。

3.5 主观效果对比

如 3.3, 3.4 实验结果所示, 本文设计模型在 3 个客

观指标上均强于 WCT² 模型, 在 PSNR 上稍弱于 CAP-VSTNet 模型; 在客观指标和主观效果上, 与原始模型 PhotoWCT² 相比, 纹理保留和颜色风格迁移效果更强。

为进一步评估本文设计的模型, 实验选取 2 组印花图案进行风格迁移, 与 CAP-VSTNet 模型的效果进行对比, 对比效果如图 7 所示。从图 7 可以看出, 在整体效果上, 本文设计模型与 CAP-VSTNet 模型相比具有更好的纹理保留和颜色风格迁移能力, 在颜色风格转换和纹理保留之间取得了良好的平衡。从图 (a) (b) 两组结果来看, 本文设计的模型仅在图 (a) 组花朵边缘出现少量伪影, 另一组未出现伪影。此外, 相较于 CAP-VSTNet 模型结果, 本文设计的模型颜色风格更偏向于风格图, 整体颜色协调未出现过曝, 颜色风格表现更好。这得益于引入的 FCAM 模块增强了模型对主要颜色信息的感知, 获取和加强能力, 使其能够实现更好的颜色风格迁移效果。同时引入的 MCLA 模块对复杂纹理信息的提取和增强能力。

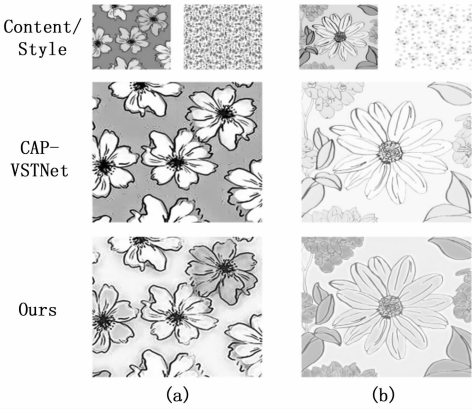


图 7 不同风格迁移模型对比试验

4 结束语

针对印花图案具有纹理复杂、风格多样的特点, 本文设计了一种基于 PhotoWCT² 和多重注意力机制的印花图案配色模型, 通过多尺度通道一局部注意力模块、频率通道注意力模块及对比学习内容损失函数, 提升了模型的纹理保留和颜色风格迁移能力。定量实验结果表明, 本文所设计的模型在 SSIM、PSNR 和 *Gram Loss* 三个定量指标上的优势。主观效果对比实验进一步证明, 本文所设计的模型证明了有更好的纹理保留和颜色风格迁移能力。未来研究将集中于超高分辨率印花图案配色及精准的印花图案配色技术。

参考文献:

[1] 张为海. 纺织品印花图案智能设计 [J]. 丝网印刷, 2021 (9): 27-33.
[2] KYPRIANIDIS J E, COLLOMOSSE J, WANG T, et al.

- State of the “art”: A taxonomy of artistic stylization techniques for images and video [J]. IEEE Transactions on Visualization and Computer Graphics, 2012, 19 (5): 866 – 885.
- [3] IOANNOU E, MADDOCK S. Evaluation in neural style transfer: a review [J]. Computer Graphics Forum, 2024, 43 (6): e15165.
- [4] JING Y, YANG Y, FENG Z, et al. Neural style transfer: a review [J]. IEEE Transactions on Visualization and Computer Graphics, 2019, 26 (11): 3365 – 3385.
- [5] HUANG X, BELONGIE S. Arbitrary style transfer in real-time with adaptive instance normalization [C] //Piscataway: IEEE, 2017: 1501 – 1510.
- [6] LI Y, FANG C, YANG J, et al. Universal style transfer via feature transforms [J]. Advances in Neural Information Processing Systems, 2017, 30 (30): 386 – 396.
- [7] HOSSAIN M. Whitening and coloring transformations for multivariate gaussian data [J]. A Slecture Partly Based on the ECE662 Spring, 2014, 4 (5): 4 – 14.
- [8] LI Y, LIU M Y, LI X, et al. A closed-form solution to photorealistic image stylization [M]. Berlin: Springer, 2018: 453 – 468.
- [9] SHI J, MALIK J. Normalized cuts and image segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22 (8): 888 – 905.
- [10] YOO J, UH Y, CHUN S, et al. Photorealistic style transfer via wavelet transforms [C] //Piscataway: IEEE, 2019: 9036 – 9045.
- [11] WEN L, GAO C, ZOU C. CAP-VSTNet: content affinity preserved versatile style transfer [C] //Piscataway: IEEE, 2023: 18300 – 18309.
- [12] GOMEZ A N, REN M, URTASUN R, et al. The reversible residual network: Backpropagation without storing activations [J]. Advances in Neural Information Processing Systems, 2017, 30: 2214 – 2224.
- [13] NIKLAUS S, LIU F. Context-aware synthesis for video frame interpolation [C] //Piscataway: IEEE, 2018: 1701 – 1710.
- [14] CHIU T Y, GURARI D. Photowct2: Compact autoencoder for photorealistic style transfer resulting from block-wise training and skip connections of high-frequency residuals [C] //Piscataway: IEEE, 2022: 2868 – 2877.
- [15] CHIU T Y, GURARI D. Pca-based knowledge distillation towards lightweight and content-style balanced photorealistic style transfer models [C] //Piscataway: IEEE, 2022: 7844 – 7853.
- [16] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation [C] //Piscataway: IEEE, 2015: 3431 – 3440.
- [17] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module [M]. Berlin: Springer, 2018: 3 – 19.
- [18] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [J]. Advances in Neural Information Processing Systems, 2017: 30.
- [19] WANG Y, LI Y, WANG G, et al. PlainUSR: chasing faster ConvNet for efficient super-resolution [M]. Berlin: Springer, 2024.
- [20] ZHANG Y, TANG F, DONG W, et al. Domain enhanced arbitrary image style transfer via contrastive learning [C] //New York: ACM, 2022: 1 – 8.
- [21] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity [J]. IEEE Transactions on Image Processing, 2004, 13 (4): 600 – 661.
- [22] GATYS L A, ECKER A S, BETHGE M. A neural algorithm of artistic style [J/OL]. ArXiv, 2015. [2025 – 02 – 20]. <https://arxiv.org/pdf/1508.06576>.
- 术发展趋势分析 [J]. 系统仿真技术及其应用, 2015, 16 (3): 207 – 211.
- [20] 胡鑫武, 罗鹏程, 张笑楠, 等. 基于体系仿真大数据的效能评估方法 [J]. 火力与指挥控制, 2020, 45 (1): 7 – 11.
- [21] 姜忠钦, 张明智, 杨镜宇, 等. 武器装备体系仿真分析支撑环境研究 [J]. 计算机仿真, 2009, 26 (1): 11 – 14.
- [22] 陈英敏. GPS 导航卫星仿真系统设计 [J]. 微计算机信息, 2011, 27 (11): 123 – 125.
- [23] 舒适, 房建成, 张伟, 等. 基于 MSCMG 大型遥感卫星高精度姿态控制方法 [J]. 中国惯性技术学报, 2017, 25 (4): 422 – 431.
- [24] 王 鹏. 卫星仿真元模型及其设计研究 [J]. 计算机仿真, 2020, 165 (12): 35 – 38.
- (上接第 250 页)
- [14] 唐 忠, 薛永奎, 刘 丽. 美军作战仿真系统综述 [J]. 航天电子对抗, 2014, 30 (4): 45 – 52.
- [15] 李 进, 钱大庆, 陈 豪. 美军大型仿真系统 [J]. 国防科技, 2009, 30 (1): 87 – 91.
- [16] 黄晓冬, 谢孔树. 高性能分布式面向对象仿真引擎研究与实现 [J]. 系统仿真学报, 2021, 23 (9): 46 – 57.
- [17] 卿杜政, 李伯虎, 孙 磊, 等. 基于组件的一体化建模仿真环境 (CISE) 研究 [J]. 系统仿真学报, 2008, 20 (4): 900 – 904.
- [18] 阮开智, 袁晴晴, 翟文华, 等. 基于 Xsim 平台的防空导弹武器系统仿真平台设计 [J]. 系统仿真学报, 2020, 32 (1): 142 – 148.
- [19] 于雪松, 张一彬, 胡叶楠, 等. 面向装备体系的仿真技