

基于 DreamBooth 的傣锦图案人工智能生成模型

纪乐福, 王永江, 李启正*

[浙江理工大学 纺织科学与工程学院(国际丝绸学院), 浙江 杭州 310018]

摘要:为推动传统文化引领下的现代纺织图案设计发展,提出了一种利用人工智能生成民族纺织图案的模型。以傣族织锦图案为例,对傣族织锦实物图案进行矢量化处理,为每张图片编写对应文本标签,并以此为训练集;选取适用于傣锦的预训练模型,使用 DreamBooth 方法微调现有的文本-图像模型。对模型训练效果进行分析,得出 V1 模型是一个具有良好拟合度和图像生成效果的文本-图像生成模型。

关键词: 模型微调;传统纺织图案;傣锦图案;DreamBooth 方法;文本-图像生成模型

中图分类号: TS 941.26 **文献标志码:** A **文章编号:** 2096-1928(2024)05-0433-10

Artificial Intelligence Generation Model for Dai Brocade Pattern Based on DreamBooth

Ji Lefu, WANG Yongjiang, LI Qizheng*

[College of Textile Science and Engineering (International Institute of Silk), Zhejiang Sci-Tech University, Hangzhou 310018, China]

Abstract: In order to promote the development of modern textile pattern design under the guidance of traditional culture, a model of introducing ethnic textile patterns into the field of artificial intelligence creation was proposed. Taking the Dai brocade pattern as an example, the physical images of Dai brocade were vectorized, and the corresponding text labels were written for each image, and the training set was used. The research selected a pre-trained model suitable for Dai brocade, and used the DreamBooth method to fine-tune the existing text image models. By analyzing the training effect of the model, it is concluded that V1 model is a text-image generation model with good fitting degree and image generation effect.

Key words: model fine-tuning, traditional textile patterns, Dai brocade patterns, DreamBooth method, text-image generation model

傣锦是一种主要流传于傣族的传统手工艺品,具有丰富的图案和色彩,反映了傣族人民的生活、信仰和文化,是傣族的文化瑰宝^[1]。随着现代工业的发展和机械化生产模式的广泛应用,傣锦面临一系列的挑战,如市场竞争激烈、生产成本上升等^[2]。主要依赖手工制作的傣锦产业正遭受破坏并逐渐被替代,传承与发展傣锦制作技艺迫在眉睫,而数字化技术的应用可以为传统手工艺品注入新的活力和创新力^[3]。

当前,人工智能(artificial intelligence, AI)图像

生成技术崭露头角,利用该技术可以生成高质量图像^[4]。目前,主流的文本-图像生成模型有3类:①以Bing图像创建器为代表的AI图像生成系统所使用的DALL-E 2模型。该模型是OpenAI公司开发的图像生成模型,其较为成熟,适合企业使用,但企业和个人用户都不能自主对模型进行微调。②Midjourney公司自行训练的模型。该模型对提示词的要求不高,在使用不够贴切的提示词时也可以生成较优质的图像。企业可以根据用户使用反馈进一步优化模型,但是用户不能微调模型。③Stable

收稿日期:2024-03-10; 修订日期:2024-08-31。

基金项目:教育部产学合作协同育人项目(202102654037)。

作者简介:纪乐福(1997—),男,硕士研究生。

*通信作者:李启正(1981—),男,副教授,硕士生导师。主要研究方向为学术平台建设与编辑出版。

Email:liqizheng@139.com

Diffusion(以下简称 SD)模型。其对提示词要求较高,需要本地算力支持,但作为开源的模型,每个用户都可以对模型进行调整。

现有图像生成类模型大多具有特定的图像风格,但在纺织非遗图案领域,尚未采用此类模型进行相关训练。为了推动传统纺织行业的发展,同时验证 AI 可以理解并学习抽象的纺织图案,文中设计了一种基于 DreamBooth 方法的傣锦图案 AI 生成模型。

1 模型训练效果比较及训练思路

1.1 图像相似性比较

将定向训练后的傣锦图案 AI 生成模型以系列编号命名,其中生成效果最贴合提示文本的模型命名为 V1。为验证傣锦模型输出图案与傣锦实物图案的相似性,文中在使用相同提示词的条件下,将 V1 模型输出图案与实物图案、训练集中相关图案进行对比,具体对比情况见表 1。

表 1 傣锦矢量图、实物图案和 AI 生成图案对比
Tab.1 Comparison among Dai brocade vector images, physical brocade images and AI-generated patterns

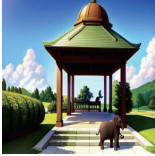
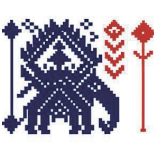


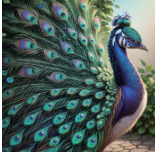


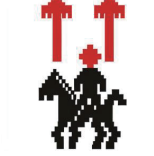


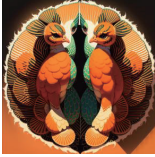



傣锦矢量图	傣锦实物图案	AI生成图案
		
		
		
		
		
		

从色彩、元素种类、造型 3 个方面对傣锦图案进行分析。由表 1 可以看出,在图案色彩方面,V1 模型输出的图案以单色作为整体的配色基调,色相与傣锦主要色彩相符;在图案元素种类方面,V1 模型可以在提示词的作用下,输出契合该元素特征的图案;在图案造型方面,V1 模型输出的图案高度还原了傣锦像素化的风格。由此证明,V1 模型对傣锦主题拟合程度良好。

1.2 模型训练效果比较

为验证傣锦图案生成模型的训练效果,文中尝试在使用相同提示词的条件下,将预训练模型输出图案与 V1 模型输出图案进行对比,图案对比情况及使用提示词见表 2。

表 2 模型训练前后输出图案
Tab.2 Output images before and after model training

提示词	模型	
	Anything V5	V1
dai_style,1elephant, 1gazebo		
dai_style,1elephant, 1gazebo,1human		
dai_style,1peacock		
dai_style,1human, 1horse,		
dai_style,1gazebo		
dai_style,2vultures, symmetry		
dai_style,1building		

由图 2 可以看出,原始模型经过傣锦数据集微调以后,在特殊字符作用下,输出的图案风格发生了巨大变化。预训练模型输出的图案有一定的 3D 效果,色彩艳丽,画风细腻,多种色彩共同构成图案;V1 模型输出的图案以傣锦经典配色为主基调,具有强烈的像素化风格。由此证明,预训练模型风格的改变是成功的。

1.3 模型训练思路

DreamBooth 方法由 Hugging Face 公司的 PATIL S 和 CUENCA P 研发,于 2022 年 11 月 7 日在 GitHub 上发布。DreamBooth 是一种使用专门的微调形式训练 SD 的新技术,其集成了多种模型和算法。该系统的核心在于将自动生成类特定先验保留损失和特定标识符相关联,以微调预训练模型,

生成个性化文本及图像。

DreamBooth 可以读取特定唯一标识符和语义先验提示词,利用嵌入在模型中的语义先验提示词生成类特定先验图像,并将这些图像与真实图像进行比较,计算损失函数;通过反向传播算法,可以最小化损失函数,以更新模型参数,实现新概念的学习,一旦该特定标识符嵌入到模型的输出域中,就可以在不同场景中输出包含该特定元素的全新图像^[5]。图 1 为采用 DreamBooth 训练 AI 生成傣锦纺织图案的思路。由图 1 可以看出,傣锦图案训练步骤为:①构建傣锦数据集,包括图案数据集和对应的文本数据集;②比较多个预训练模型,选择对经典傣锦元素回应准确的模型作为基底模型;③调节 DreamBooth 训练相关参数,优化模型训练结果。

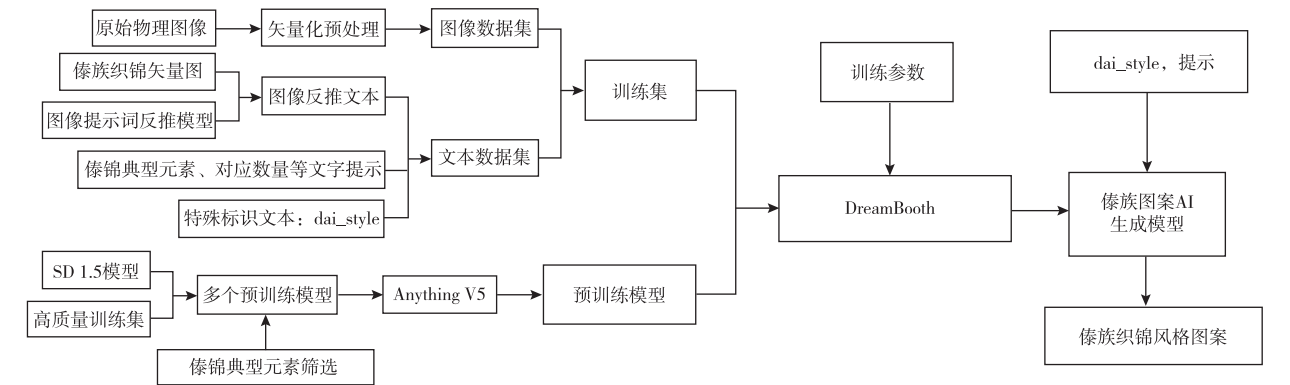


图 1 DreamBooth 训练生成傣锦图案思路

Fig. 1 Process of training for generating Dai brocade images through DreamBooth

DreamBooth 这种使用专门的微调形式训练 SD 的新方法需要算力支持,文中训练时使用了 AutoDL 平台提供的云端算力,配置的 CPU 为 15 vCPU AMD EPYC 7543 32-Core Processor, GPU 为 RTX A5000 (24GB),使用的软件环境为 Python 3.10、PyTorch 1.12.1、CUDA 10.0,操作系统为 Linux。文中所采用的操作软件基于开源项目 SD web UI。

2 模型训练数据集构建

2.1 傣锦图案特征分析

笔者通过对云南西双版纳、德宏两地傣锦的考察,以及对西双版纳聚匠非遗文化传播交流中心、西双版纳傣族自治州文化馆、西双版纳民族博物馆、景洪市嘎洒镇曼乱点村国家级传承人玉儿甩工作坊、勐腊县勐腊镇曼龙代村委会曼龙代村省级传承人咪香应工作坊、德宏州博物馆、德宏州文化馆、德宏州图书馆、芒市芒市镇等相村省级传承人雷安

座傣锦坊的调研,搜集了 1 500 余张傣锦相关实物图片以及大量文献资料。在此基础上,分析不同支系傣锦的纹样特征并将其进行矢量化呈现。总结得出傣锦图案有以下 3 个特点:

- 1)纹样色彩协调且鲜艳明亮。每幅傣锦图案通常由 3~5 种颜色组成,包括原色、纯色、互补色和对比色^[6]。西双版纳傣锦和德宏傣锦的色彩特征存在区别^[7]。西双版纳傣锦在色彩搭配上保留了原汁原味的民族特色,通常以白纱线作通经,红、黑色纬线织花,大多以白色为底色,再配以黑色或红色纹样;德宏地区的傣锦则色彩更丰富,以黑色为底色,搭配红、绿、黄、蓝等高纯度的对比色,视觉效果强烈。此外,深绿、草绿、桃红、大红、明黄、橘红等颜色也常用于纹样中,并以褐色勾勒纹边,白色作为分隔符,使颜色更加协调^[8]。
- 2)纹样种类繁多。傣锦纹样大多取自自然景物、人类生活场景以及一些事物的抽象表达。例如,因傣族深受佛教信仰的影响,所以傣锦纹样中

经常出现佛寺、佛塔、孔雀、菩提花等有较浓佛教文化象征的元素。傣锦纹样按题材主要分为动物纹样、几何纹样、植物纹样、建筑纹样、人物纹样、其他纹样 6 类。傣锦的组织形式主要为单独式和连续式,单独式纹样一般分为对称和非对称 2 种,并以对称式纹样为主;连续式纹样一般分为二方连续和四方连续 2 种。傣锦图案大多为连续式纹样。

3)纹样造型具有典型的像素特征。整体而言,傣锦传统纹样具有连续、重复的对称美和高度几何化的形式美。它通过点、线的变形和组合排列形成造型各异的面状纹样。多数傣锦纹样为直线与斜线交错的几何图案,并采用循环或轴对称等方式进行组合,形成完整的纹样^[9]。傣锦纹样呈由点及线、由线及面的抽象化表现形式,整体纹样简洁凝练且层次丰富,呈现出强大的视觉美感^[10]。图 2 为典型傣锦纹样。



图 2 经典傣锦实物纹样
Fig. 2 Classic Dai brocade physical images

2.2 傣锦矢量化图案数据集构建

文中傣锦资料多来源于文物博物馆、图书馆、档案馆以及现存非遗文化传承人作坊保存的实物影像资料,因此需要对实物图案进行预处理。由于傣锦具有织物组织的特点,其图案比一般的位图更复杂,含有更多的干扰信息,图案线条存在一定的扭曲变形,所以在智能识别时效果不佳。因此,笔者从 1 334 张傣锦实物图片中整理得到 394 幅典型

傣锦图案,进行纹样提取,得到 276 种傣锦的典型纹样,并将典型纹样矢量化。将纹样按照动物纹样、几何纹样、植物纹样、建筑纹样、人物纹样、其他纹样 6 类进行分类,得到动物纹样 89 种(其中大象纹 22 种,孔雀纹 8 种,马纹 22 种,其他动物纹 37 种),几何纹样 20 种,植物纹样 72 种,建筑纹样 22 种,人物纹样 27 种,其他纹样 46 种。


数据集的创建是 DreamBooth 训练中的关键环节^[5]。傣锦纺织图案 AI 生成模型建立在 SD 1.5 模型的基础上,因此在训练中,数据集格式尽可能贴合 SD 的训练数据集会得到更好的训练效果,所以在进行傣锦图案标注前需进行预处理,包括统一尺寸、统一风格。按照 SD 1.0 和 SD 2.0 的要求,模型采用的训练集图像尺寸为 256 × 256、512 × 512、768 × 768 像素,其中以 512 × 512 像素居多,所以在不破坏图案原始比例的条件下,将傣锦图案缩放至 512 × 512 像素。

从矢量化图案中选取适用于 DreamBooth 方法的图案,按图案元素进行分类,构建数据集,结果见表 3。为增强识别效果、提升数据集的代表性,将分类的主体从图案中分离,为每个类别设置单独的数据集,结果见表 4。

表 3 完整纹样分类数据集

Tab. 3 Complete pattern classification dataset		
类别	图案数/个	实物图例
动物纹样 (象纹)	22	
植物纹样	28	
人物纹样	20	
动物纹样 (马纹)	11	
动物纹样 (孔雀纹)	8	
建筑纹样	6	
几何纹样	20	

表 4 单一元素纹样分类数据集

Tab. 4 Single element pattern classification dataset		
类别	图案数/个	实物图例
动物纹样 (象纹)	22	
植物纹样	83	
人物纹样	27	
动物纹样 (马纹)	22	
动物纹样 (孔雀纹)	20	
建筑纹样	54	
几何纹样	23	

2.3 模型文本数据集构建

DreamBooth 的微调扩散模型对提示词和图像数量有较高的要求,包括提示词应该简单易懂,能够描述所需的场景或环境,并且应该足够清晰和具体,以有效地指导图像生成过程。目前,DreamBooth 的文本提示有 2 种方式:①使用自然语言,即使用一句完整的语句来描述图像;②使用多个单词描述图像。考虑到实验中需多次修改提示词文本,文中采用方式②编写提示词。

傣锦文本-图像数据集中,文本的来源除特定标识符——dai_style(自行设定,在原本模型中无特定含义即可)外,还包括以下 2 个方面:

1) 傣锦图案中具体元素的文本。傣锦文本-图像数据集按特征可分为 2 类:①单一元素图案数据集;②完整图案数据集。对应这 2 个分类,有不同的注释方式。对于单一元素图案,只需标注具体元素及数量;对于完整图案,还需关注元素顺序,将图案主体元素置于其余元素之前。

2) 图像反推文本。在 SD web UI 中使用插件 Waifu Diffusion 1.4 Tagger。Waifu Diffusion 1.4 Tagger 可以用于提取图像特征提示词,其使用 ViT

模型标记图像。ViT 模型由 Google 开发,其使用 Transformer 架构,可以将图像分成小块,再将小块转换为向量,这些向量可以用于分类、检测或标记图像。使用 Waifu Diffusion 1.4 Tagger 对所有傣锦图案提取特征提示词,并清洗与图案无明显联系的提示词。以象驼宝塔诵经厅为例,表 5 为置信度 20% 以上的提示词及其置信度。清洗后,添加到提示词文件中的有 solo、full body、1girl、silhouette、white_background、simple_background、monochrome、braid。

表 5 象驼宝塔诵经厅特征提示词及置信度

Tab. 5 Feature keywords and confidence levels for the elephantine pagoda recitation hall	
特征提示词	提示词置信度/%
solo	81
full body	68
1girl	66
silhouette	59
white_background	58
simple_background	53
shide	43
monochrome	39
twin_braids	36
braid	25

3 预训练模型对比与训练参数讨论

3.1 文本-图像扩散模型对比

DreamBooth 方法最初用来对 SD 模型进行微调,随着技术发展,出现了 SD 1.5 和 SD 2.0 模型,3 个模型的区别在于模型架构和分辨率。SD 1.5 默认支持 512×512 像素分辨率,目前的训练大多基于 SD 1.5 模型。Models 和 Anything V5 都是选取 SD 1.5 为基底模型,由模型创建者使用自己准备的数据集训练得到的。与原本 SD 1.5 和 SD 2.0 模型相比,Models 和 Anything V5 在输出图案时具有更好的准确性。考虑到傣锦图案的像素化风格,选取像素风格模型 Pixelart Ultramerge 进行模型训练。

在选择傣锦模型训练底模时,以傣锦图案中蕴含的元素为标准,测试模型输出元素的完整度。测试时,所有模型使用相同提示词,每组提示词各输出 100 个图案,统计其中完整且准确描述提示词的图案出现频率。模型准确率对比情况见表 6。由表 6 可知,Anything V5 模型对典型傣锦元素输出的准确率最高,因此文中选择以该模型为基底。

表 6 多模型输出图案准确率对比

Tab. 6 Accuracy comparison of multiple model output images

底模	准确率/%			
	人、大象	人、马	大象、鸟	平均值
SD 1.5	54	32	67	51
SD 2.0	75	26	34	45
Models	65	91	87	81
Anything V5	85	89	94	89
Pixelart Ultramerge	52	78	45	58

3.2 模型训练参数讨论

在傣锦风格模型的训练中,选用合适的参数可以提高训练效果。RUIZ N 等^[5]所用模型的测试数据集包括 3~5 张图片,其在学习率为 1×10^{-5} 、权重衰减系数为 1 时,进行 200 个 epoch 就足以获得良好的结果。在实际训练中,训练效果会根据数据集的变化而改变,需要进行多次参数调整以追求更满意的效果。

Batch_size、epoch、学习率是深度学习中的重要参数,其设置会影响模型的训练效果。Batch_size 的大小会影响模型的训练时间和稳定性,较大的 batch_size 可以减少训练时间,同时梯度的计算更加稳定;epoch 影响实验总步数,步数越多,模型的精度越高,但是过多的步数可能会导致过拟合;学习率越大,模型收敛速度越快,但是可能会导致震荡或者不收敛。文中所有训练过的模型中,效果最优的为 V1,基底模型为 Anything V5,epoch 为 100, batch_size 为 8,固定学习率为 1×10^{-5} 。

表 7 不同训练参数模型损失值对比

Tab. 7 Comparison of loss values for different training parameters

模型编号	底模	epoch	batch_size	学习率	损失值
V1	Anything V5	100	8	1×10^{-5}	0.014
V2	Anything V5	100	8	2×10^{-5}	0.025
V3	Anything V5	100	8	3×10^{-5}	未收敛
V4	Anything V5	100	8	9×10^{-6}	0.018
V5	Anything V5	100	8	8×10^{-6}	0.024
V6	Anything V5	100	8	7×10^{-6}	0.027
V7	Anything V5	100	6	1×10^{-5}	0.015
V8	Anything V5	100	8	1×10^{-5}	0.014
V9	Anything V5	100	1	1×10^{-5}	0.015
V10	Anything V5	100	1	1×10^{-5}	0.014
V11	Anything V5	30	8	1×10^{-5}	未收敛
V12	Anything V5	50	8	1×10^{-5}	未收敛
V13	Anything V5	70	8	1×10^{-5}	0.032
V14	Anything V5	90	8	1×10^{-5}	0.019
V15	Anything V5	110	8	1×10^{-5}	0.014
V16	Anything V5	120	8	1×10^{-5}	0.015

4 傣锦模型训练效果分析

4.1 损失值分析

4.1.1 损失函数构建 在卷曲神经网络训练中,损失值越小,学习效果越好^[11]。损失值是每次训练中预测值与实际值的差异,通过反向传播可以改变网络的各层权重,从而优化网络参数,提高网络性能。通过分析损失值可以得到对模型训练效果的大致判断^[12]。AI 图像生成模型训练与传统深度学习略有不同,通常取得良好训练效果的损失曲线呈下降趋势,且在训练临近结束时围绕某个数值上下波动。

DreamBooth 使用个性化的类别先验损失函数,它主要关注生成的图像在语义上与输入图像相似的主题特征^[13]。该损失函数包括 2 个子项:上下文保留项和类别先验保留项。上下文保留项旨在保留输入图像的上下文信息,以确保生成的图像与输入图像在视觉上相似;类别先验保留项用于生成与给定人物主题相似的新图像,并确保生成的图像与其类别保持一致。

通过使用损失函数,DreamBooth 可以生成符合给定主题的多种图像,并且这些图像可以在不同的环境中自然互动,同时保留主题的关键视觉特征。对模型训练效果产生影响的重要参数为 epoch、batch_size、学习率,通过对训练参数的对比优化,可以得到相对有效的傣锦风格模型。将代表性模型相关参数、最后一个 epoch 收敛情况以及平均损失值进行对比,结果见表 7。

从表 7 可以看出, $\text{epoch} \leq 50$ 时,底模与其他参数的变化均不能改变损失值曲线不能收敛的情况; epoch 为 100 时,损失函数普遍收敛,且更大的 epoch 值会使模型过拟合概率增加,因此 epoch 值大于 100 意义不大。由此得出,底模为 Anything V5, epoch 为 100,学习率为 1×10^{-5} 时,算法收敛情况最好,此时模型拟合程度良好,训练取得了较好的结果。

4.1.2 学习率对傣锦模型训练效果的影响 通过分析实际训练效果及损失曲线,发现学习率高于 1×10^{-4} 或低于 1×10^{-5} 时,都会导致模型训练不拟合。将 1×10^{-4} 和 1×10^{-5} 分别作为学习率的上下限,取一定的间隔分别进行训练,设定 epoch 为 100, batch_size 为 8,得到的损失曲线如图 3 所示。

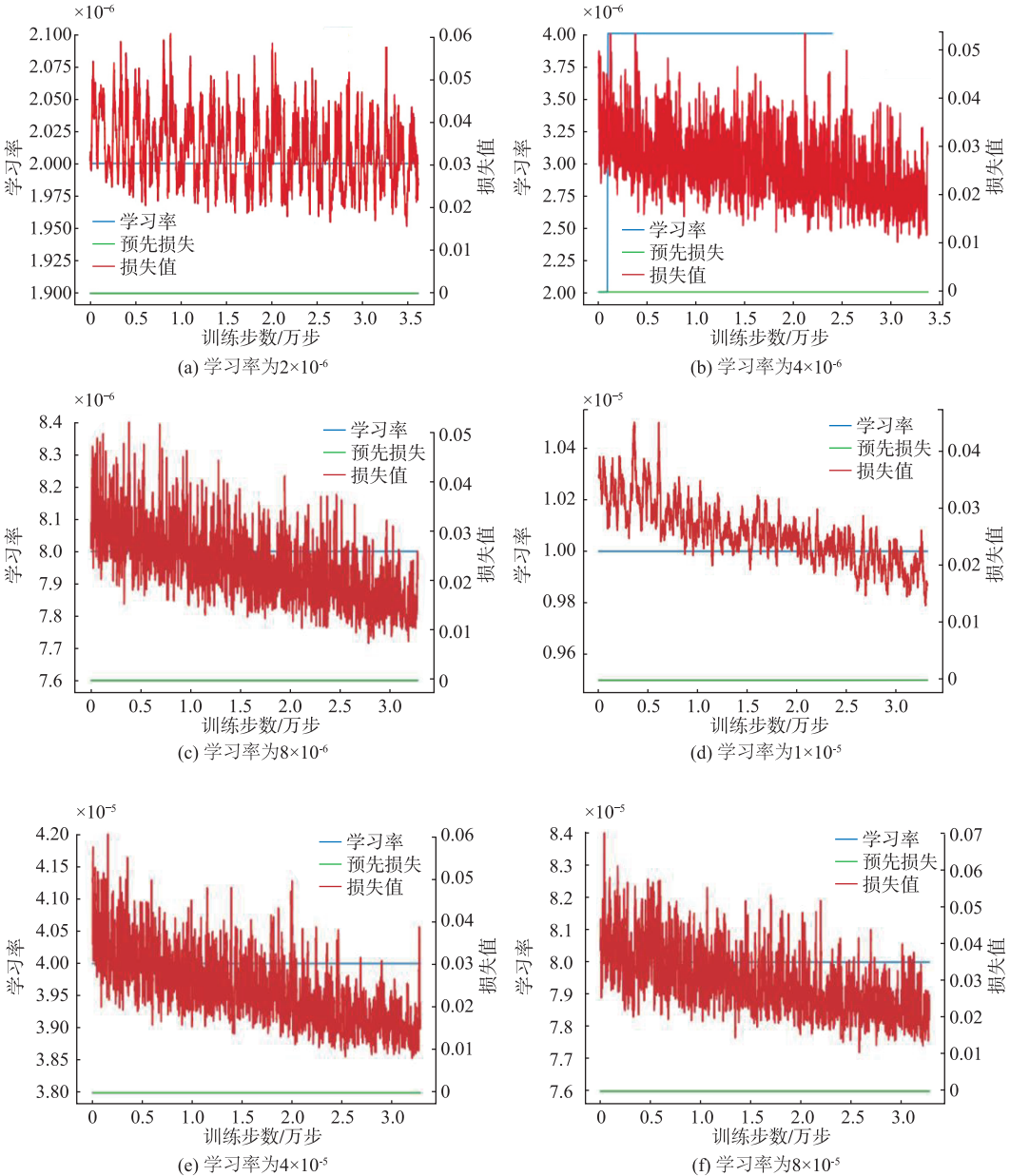


图 3 不同学习率的模型训练损失曲线

Fig.3 Model training loss curves with different learning rates

由图 3 可以看出,学习率为 1×10^{-5} 时,损失值波动相对较小,损失曲线有明显的下降趋势,说明模型训练取得了良好的效果。

4.1.3 Batch_size 对傣锦模型训练效果的影响 因硬件算力有限,RTX A5000 最大可设定 batch_size

为 8,根据实际训练效果及后台数据分析结果,得出相对有效的 batch_size 介于 1 与 8 之间。将 1 与 8 分别作为 batch_size 的上下限,取一定的间隔分别进行训练,设定 epoch 为 100,学习率为 1×10^{-5} ,得到损失曲线如图 4 所示。

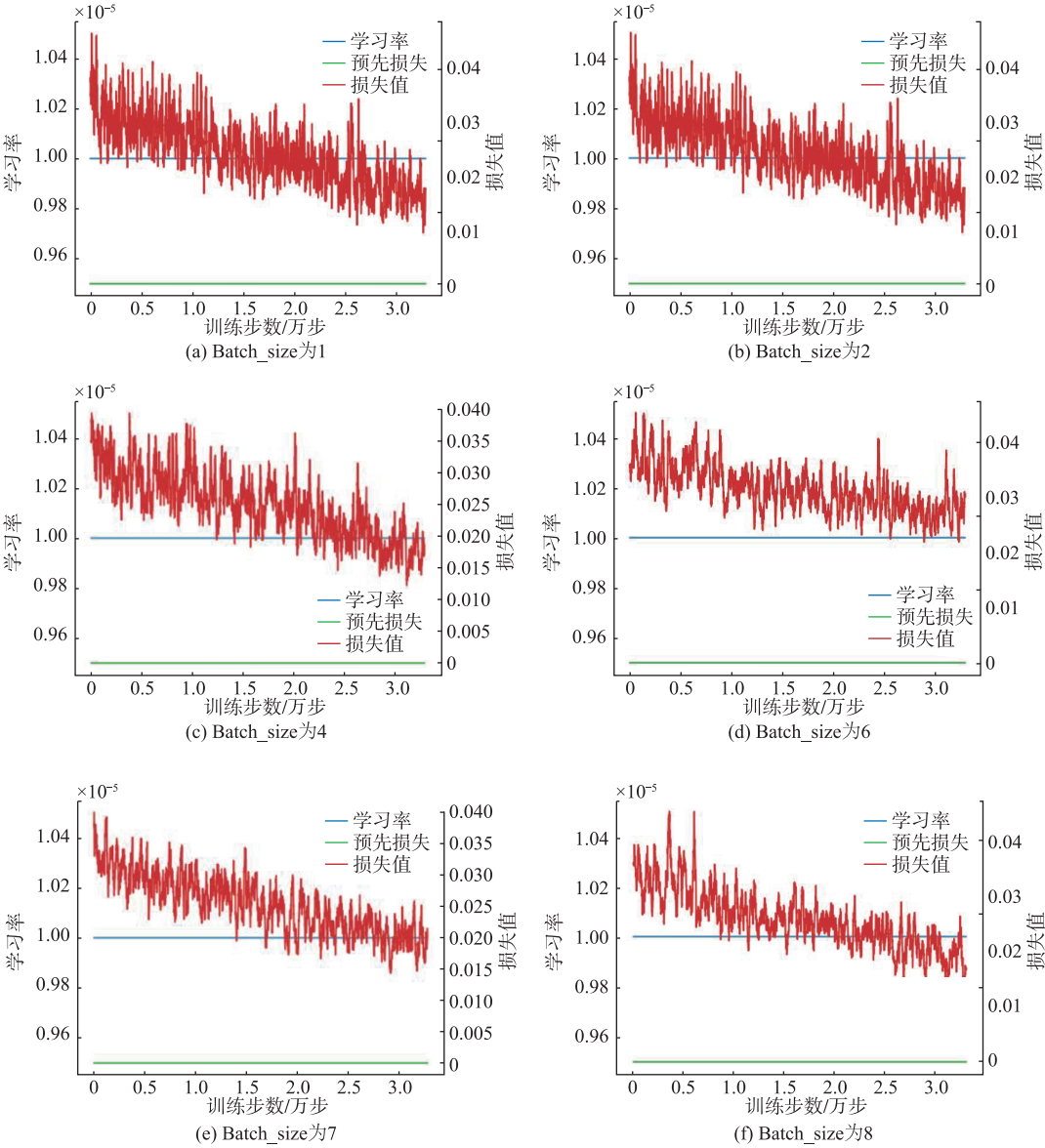


图 4 不同 batch_size 的模型训练损失曲线

Fig. 4 Model training loss curves of different batch_sizes

由图 4 可以看出, batch_size 为 8 时, 损失值波动相对较小, 损失曲线有更加明显的下降趋势, 说明模型训练取得了良好的效果。

4.2 图像相似性分析

文中通过对比 V1 模型生成图像与训练集图像的相似度, 评价模型拟合程度, 引入的图片相似性评价指标包括以下 4 个:

1) 结构相似性。结构相似性用来评估两幅图像之间的相似水平^[14], 对于图像 x 和 y , 其结构相似性的计算公式^[15]为

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

式中: μ_x 和 μ_y 为信号的平均值; σ_x 和 σ_y 为信号的标准差; σ_{xy} 为信号的协方差; C_1, C_2, C_3 皆为常数。

当结构相似性值在 0.9 以上时, 图像拥有极高的相似度; 当结构相似性值在 0.7 左右时, 图像之间有一定的相似性, 但存在肉眼可以察觉的区别。

2) 均方误差。均方误差用来评估修复后的图像 I 与原始图像 K 之间像素级别上的差异程度^[16], 其计算公式^[17]为

$$E(i, j) = \frac{1}{m \times n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2 \quad (3)$$

式中: m 为图像 I 的像素总数; n 为图像 K 的像素总数; $I(i, j)$ 和 $K(i, j)$ 分别为修复图像与原始图像在第 (i, j) 个像素处的亮度值。均方误差越小, 说明图像越相似^[18]。

3) 均方根误差。均方根误差也称为标准误差, 是均方误差的算术平方根^[19]。引入均方根误差与引入标准差 (均方差) 的原因是一致的, 即均方误差

的量纲与数据量纲不同,不能直观反映离散程度,故在均方误差上开平方根,得到均方根误差^[20] $R(i,j)$,其公式为

$$R(i,j) = \sqrt{E(i,j)}。$$

(4)

4) 峰值信噪比。峰值信噪比是评价噪声水平或图像压缩后失真情况的客观评价指标,其计算公式^[13]为

$$P(i,j) = 10\lg\left(\frac{I_{\max}^2}{E(i,j)}\right)。$$

(5)

式中: $E(i,j)$ 为两张图像的均方误差; I_{\max} 为图像中的像素最大值。峰值信噪比越大说明失真情况越

少,生成图像的质量越好。峰值信噪比大于 50 分贝时,说明压缩后的图像仅有非常小的误差;其数值为(30,50]分贝时,人眼很难察觉压缩后图像和原始图像的差异;其数值为(20,30]分贝时,人眼可以察觉出压缩前后图像的差异;其数值为(10,20]分贝时,可以用肉眼看出图像的原始结构,且直观上能看出压缩前后的图像存在较大差异。

文中使用 V1 模型在典型提示词下生成对应图案,用多个相似性评估指标对生成图案与训练集中对应分类下所有图案的相似度进行评估,结果见表 8。

表 8 不同元素图案相似度

Tab.8 Similarity of different element images

相似度评价指标	元素					平均值
	象、宝塔	象、人	孔雀	人、马	建筑	
结构相似性	0.643	0.673	0.701	0.752	0.869	0.729
均方误差	0.716	0.754	0.930	0.814	0.772	0.792
均方根误差	0.846	0.868	0.964	0.902	0.850	0.890
峰值信噪比	17.576	16.091	21.570	17.311	16.617	17.833

由表 8 可以看出,V1 模型用提示词生成的图案与训练集中的图案有一定相似度,同时又有足够的区分度,即模型拟合程度良好,训练取得了较好的效果。

5 结 语

文中在进行傣锦模型训练时,设置 epoch 为 100,batch_size 为 8,固定学习率为 1×10^{-5} ,此时模型的训练结果最佳。研究发现,epoch 低于 70 会导致模型训练失败,超过 100 后,训练意义降低。在算力允许范围内,batch_size 越大,损失值波动越小;数值下降越快,越容易取得良好的训练效果。学习率较小或较大,都会导致模型训练效果欠拟合,因此需选用合适的学习率。

文中将 AI 图像生成技术引入到传统纺织图案设计领域,得出 V1 模型具有良好的训练效果。该模型在训练时损失曲线波动性下降,最终收敛于较小的损失值。模型完成微调以后,输出图案在色彩和造型上具有鲜明的傣锦特征,且与对应分类的傣锦图案相似度较高,证明模型训练取得了较好的结果,也验证了 AI 图像生成经典纺织图案的可行性。

参考文献:

[1] 方婷玉,艾惠花,李启正,等. 傣族织锦支系地域的纹样特征及矢量化研究[J]. 丝绸, 2023, 60(2): 84-92.

FANG Tingyu, AI Huihua, LI Qizheng, et al. Study on the pattern characteristics and vectorization of Dai brocade

branches and regions[J]. Journal of Silk, 2023, 60(2): 84-92. (in Chinese)

[2] 桂宇,阴佳芮. 从傣锦织造技艺看女性主体构建的民族文化表征[J]. 服装学报,2023,8(4):350-356.

GUI Yu,YIN Jiarui. Viewing the national cultural representation dominatedly constructed by female from Dai brocade weaving technique [J]. Journal of Clothing Research,2023,8(4):350-356. (in Chinese)

[3] 刘倩. 自贸港背景下黎锦传统元素数字化传播机制研究[J]. 西部皮革, 2023, 45(8): 63-65.

LIU Qian. Research on the digital communication mechanism of Lijin traditional elements under the background of free trade port[J]. West Leather, 2023, 45(8): 63-65. (in Chinese)

[4] 肖红梅,王伟珍,房媛. 基于生成对抗网络的女上装图像属性编辑[J]. 服装学报,2024,9(1):42-47.

XIAO Hongmei, WANG Weizhen, FANG Yuan. Image attribute editing of women's tops based on generating adversarial networks [J]. Journal of Clothing Research, 2024,9(1):42-47. (in Chinese)

[5] RUIZ N, LI Y Z, JAMPANI V, et al. DreamBooth: fine tuning text-to-image diffusion models for subject-driven generation[C]//2023 IEEE/CVF conference on computer vision and pattern recognition. Vancouver: IEEE, 2023: 22500-22510.

[6] 汤琳,贺荣,蔡王丹,等. 傣族织锦的艺术特征及其应用分析[J]. 丝绸, 2021, 58(8): 113-119.

TANG Lin, HE Rong, CAI Wangdan, et al. Analysis on the artistic characteristics of Dai brocade and its applica-

- tion[J]. Journal of Silk, 2021, 58(8): 113-119. (in Chinese)
- [7] 赵梦蝶, 谢颖, 李安娜. 傣锦纹样特征与民族精神内涵关系研究[J]. 中国民族博览, 2023(1): 40-42.
ZHAO Mengdie, XIE Ying, LI Anna. Study on the relationship between Dai brocade pattern characteristics and national spirit connotation[J]. China National Exhibition, 2023(1): 40-42. (in Chinese)
- [8] 申玉红, 杨启祥, 周长军. 云南德宏傣族服饰中的数学文化[J]. 数学教育学报, 2013, 22(1): 66-69.
SHEN Yuhong, YANG Qixiang, ZHOU Changjun. Mathematical culture of Dehong Dai dress in Yunnan Province[J]. Journal of Mathematics Education, 2013, 22(1): 66-69. (in Chinese)
- [9] 陈晴, 孙宏, 朱淑芹, 等. 傣族纹样色织小提花面料的设计[J]. 纺织科技进展, 2017(5): 28-31.
CHEN Qing, SUN Hong, ZHU Shuqin, et al. Design of yarn dyed jacquard fabric with Dai nationality pattern[J]. Progress in Textile Science and Technology, 2017(5): 28-31. (in Chinese)
- [10] 骆璠, 王羿. 傣锦佛幡的图案纹样及其文化内涵浅析[J]. 山东纺织经济, 2013, 30(12): 62-63, 67.
LUO Yunman, WANG Yi. A brief analysis of the patterns and cultural connotation of Dai brocade[J]. Shandong Textile Economy, 2013, 30(12): 62-63, 67. (in Chinese)
- [11] 谭建荣, 岳小莉, 陆国栋. 图形相似的基本原理、方法及其在结构模式识别中的应用[J]. 计算机学报, 2002, 25(9): 959-967.
TAN Jianrong, YUE Xiaoli, LU Guodong. Basic principle, method of graphic similarity and its application to structure pattern recognition[J]. Chinese Journal of Computers, 2002, 25(9): 959-967. (in Chinese)
- [12] 李林红, 杨杰, 蒋严宣, 等. 基于改进 MobileNet v2 的服装图像分类算法[J]. 现代纺织技术, 2024, 32(4): 93-103.
LI Linhong, YANG Jie, JIANG Yanxuan, et al. Clothing image classification algorithm based on improved MobileNet v2[J]. Advanced Textile Technology, 2024, 32(4): 93-103. (in Chinese)
- [13] 赵昀睿, 许倩倩, 姜阳邦彦, 等. 迭代式的深度 PU 学习与类别先验估计框架[J]. 计算机学报, 2022, 45(12): 2667-2686.
ZHAO Yunrui, XU Qianqian, JIANG Yangbangyan, et al. An iterative framework for deep PU learning and class prior estimation[J]. Chinese Journal of Computers, 2022, 45(12): 2667-2686. (in Chinese)
- [14] 佟雨兵, 张其善, 祁云平. 基于 PSNR 与 SSIM 联合的图像质量评价模型[J]. 中国图象图形学报, 2006, 11(12): 1758-1763.
TONG Yubing, ZHANG Qishan, QI Yunping. Image quality assessing by combining PSNR with SSIM[J]. Journal of Image and Graphics, 2006, 11(12): 1758-1763. (in Chinese)
- [15] 张小利, 李雄飞, 李军. 融合图像质量评价指标的相关性分析及性能评估[J]. 自动化学报, 2014, 40(2): 306-315.
ZHANG Xiaoli, LI Xiongfei, LI Jun. Validation and correlation analysis of metrics for evaluating performance of image fusion[J]. Acta Automatica Sinica, 2014, 40(2): 306-315. (in Chinese)
- [16] 张维强, 宋国乡. 基于一种新的阈值函数的小波域信号去噪[J]. 西安电子科技大学学报, 2004, 31(2): 296-299, 303.
ZHANG Weiqiang, SONG Guoxiang. Signal de-noising in wavelet domain based on a new kind of thresholding function[J]. Journal of Xidian University, 2004, 31(2): 296-299, 303. (in Chinese)
- [17] 孟球, 杨旭. 目标跟踪算法综述[J]. 自动化学报, 2019, 45(7): 1244-1260.
MENG Lu, YANG Xu. A survey of object tracking algorithms[J]. Acta Automatica Sinica, 2019, 45(7): 1244-1260. (in Chinese)
- [18] 刘贵喜, 杨万海. 基于小波分解的图像融合方法及性能评价[J]. 自动化学报, 2002, 28(6): 927-934.
LIU Guixi, YANG Wanhai. A wavelet-decomposition-based image fusion scheme and its performance evaluation[J]. Acta Automatica Sinica, 2002, 28(6): 927-934. (in Chinese)
- [19] 钟建军, 宋健, 由长喜, 等. 基于信噪比评价的阈值优选小波去噪法[J]. 清华大学学报(自然科学版), 2014, 54(2): 259-263.
ZHONG Jianjun, SONG Jian, YOU Changxi, et al. Wavelet de-noising method with threshold selection rules based on SNR evaluations[J]. Journal of Tsinghua University (Science and Technology), 2014, 54(2): 259-263. (in Chinese)
- [20] 褚江, 陈强, 杨曦晨. 全参考图像质量评价综述[J]. 计算机应用研究, 2014, 31(1): 13-22.
CHU Jiang, CHEN Qiang, YANG Xichen. Review on full reference image quality assessment algorithms[J]. Application Research of Computers, 2014, 31(1): 13-22. (in Chinese)

(责任编辑:沈天琦)