

華南師範大學

《人工智能导论》课程项目

课 程 项 目 报 告

项 目 题 目：图像风格迁移

所 在 学 院：计算机学院

项 目 组 长：周浩瀚

小 组 成 员：陈赋全，李树春，王东伟

开 题 时 间：2020 年 4 月 6 日

目录

| | |
|--------------------------|----|
| 一. 引言..... | 1 |
| 二. 国内外研究现状..... | 1 |
| 2.1 国内研究现状..... | 2 |
| 2.2 国外研究现状..... | 2 |
| 三. 模型与算法..... | 3 |
| 3.1 提取和表示图像的内容特征..... | 4 |
| 3.2 提取和表示图像的风格特征..... | 4 |
| 3.3 VGG-19 卷积神经网络结构..... | 5 |
| 3.4 风格转移总损失函数..... | 6 |
| 四. 实验结果分析..... | 7 |
| 4.1 实验搭建环境..... | 7 |
| 4.2 实验过程和结果..... | 7 |
| 五. 结论..... | 10 |
| 六. 参考文献..... | 11 |

一、引言

我们希望通过计算机，可以学习到描述绘画风格的方法，并且，还希望从任意图像的集合中，学习该图像集合风格的描述，特别是该集合与其他集合相比特有风格，同时将对图像的特定风格按照人类预定的想法进行修改。这里的风格是一个概括，可以具体到画面的色彩、纹理的风格，或画面中某类特定对象物体的外观、状态风格等等。实现图像艺术风格化。

图像艺术风格化的研究意义在于，使得计算机能够模拟人类的艺术创作形式。如果计算机能够模拟出艺术家的风格，则计算机有了更强的学习能力和智能性，从而使得具有更强的艺术表现能力和心理震撼感力的非真实感图像能够被大量生成。再者，随着移动硬件设备的快速发展，和其具备的良好拍摄功能，图像的标准化、真实感感染已经不再是图像处理技术的唯一追求目标。人们对于图像处理也已经不再只关注平滑、复原、增强等处理形式，而是更期待为图像赋予个性化和艺术化的美好视觉体验。现代人对于图像的风格艺术化有很高的需求，因此产生图像艺术化的应用也越来越多。这些艺术化后的图像除了给读者和观众带来视觉上的享受之外，还创作出更多有艺术价值的作品。

二、国内外研究现状

图像风格迁移是当前图像处理领域内的研究热点和难点。图像风格迁移可用于艺术创作、影视特效等领域，尤其是现在移动互联网的快速发展，通过图像风格迁移创作梦幻图片，并分享在社交网站上，深受人们欢迎。

许多研究者已经对图像风格迁移进行了研究。

2.1 国内研究现状:

2010 年,王东等人提出了一种基于纹理传输的图像风格仿真算法。该算法在 YCbCr 颜色空间中,计算并匹配其亮度直方图的相似性,在纹理传输过程中优化误差。最后通过对传输结果图像与背景图像进行加权处理,来完成图像风格融合。该算法主要应用于烙画风格的仿真,但是绘制出的仿真结果图像中的人物轮廓断断续续。

2014 年,李阳等人提出了多变量高斯差分风格化方法。该方法采用双边滤波的方式,平滑处理输入图像;在平滑后的图像上,构造输入图像的 ETF 流;在获得 ETF 流梯度方向及切线方向后,根据这两个方向,沿着其梯度轴和流线轴进行高斯差分滤波处理;最终得到风格化的图像。该方法仅处理了主体边缘的粗糙问题,但对于大面积少边缘的情况处理欠佳。

2015 年,王洋等人提出了基于非线性滤波和纹理传输的图像风格化算法。该算法通过相对总变差和 L_0 梯度最小化滤波器,将源图像去纹理;通过相对总变差滤波器,提取样本图像的纹理;通过亮度重映射和线性叠加的方式,实现纹理传输。该算法仅限于纹理风格特征的迁移,并没有处理色彩风格迁移。

2018 年,许哲豪等人提出了基于生成对抗网络的图像风格迁移方法。通过训练 GANs 网络结构,采用生成器和判别器,实现两个图像 X 与 Y 之间的相互映射;根据 GANs 引入的循环一致性损失,图像 X 被映射到 Y 的同时被再次映射;计算图像 X 与映射回来的图像之间的损失;最后迭代优化损失值,得到风格迁移结果。该方法的缺点在于 GANs 网络难以训练,

经常会出现模型崩塌现象通过以上分析，当前的图像风格迁移方法大多数存在色彩特征和纹理特征缺失的问题。

2.2 国外研究现状：

2015 年，Gatys 提出了风格迁移的深度学习方法，他是对白噪声进行训练的。输入图片在深层卷积神经网络中不同中间层的输出，可以分别看作其风格特征和内容特征。

2016 年，Frigo 等人提出了一种基于贴片分割并重构的非监督图像风格化算法。运用四叉树，对源图像进行分割；搜索到风格图片中与其匹配的贴片，再使用双线性混合将缝线消除；最后完成整体的颜色迁移，得到风格化图片。该方法是将纹理和色彩信息分开处理来完成迁移，导致风格效果出现断层现象，并且处理速度比较慢。

2016 年，Johnson 利用 Radford 提出的生成器在经过网络训练后，可以大大提高风格迁移的速度，也就是快速风格迁移。

2017 年，Dumoulin 利用相同的网络结构，结合了归一化思想，提出了条件归一化，实现了不同风格的图像风格融合。

三、模型和算法

我们采用的方案是输入原始的内容图片和风格图片，用内容图片做为初始化对象，定义总损失函数，通过最小化总损失函数，不断优化目标图像内容，最终生成的目标图片，同时具有输入的内容和风格。我们优化的目标是目标图像，不需要训练新的网络，只需要使用预训练好的 VGG-19

模型。

我们的损失函数网络首先来自 Gatys 2015 年论文中提出的风格迁移损失函数。论文指出，输入图片在深层卷积神经网络中不同中间层的输出，可以分别看作其风格特征和内容特征，它取内容图片和生成图片内容特征的差作为内容损失函数，取风格图片和生成图片的风格特征的差作为风格损失函数，内容损失函数和风格损失函数加权综合，就作为风格迁移的总损失函数。

3.1 提取和表示图像的内容特征： 计算生成图像图像与内容图像的特征均方误差。使用二范数，建立目标图像与内容图像的特征表示之间的误差方程，如下图所示：

$$L_{\text{content}}(\vec{p}, \vec{x}, l) = \frac{1}{2} \sum_{i,j} (x_{i,j}^l - p_{i,j}^l)^2$$

其中 \vec{p} 表示源图像， \vec{x} 是不断更新的噪声图像，即最终的生成图像。 l 是每一个卷积层。 $x_{i,j}^l$ 和 $p_{i,j}^l$ 分别表示生成图像和源图像在第 l 层位置为 j 的第 i 个卷积核的激活值。

使用 torch 模型来实现这两个特征映射集合之间的平均方差的方法，通过 pytorch 里面的函数来计算。

3.2 提取和表示目标图像的风格特征：

我们使用一个特征空间来获得输入图像的风格表示，这个特征空间可以很好地捕获图像的纹理信息，此空间可以构建在网络任何层的滤波器响应之上。Gram 矩阵可以满足上述要求，Gram 矩阵是用来度量两两响应结

果的相关性。它包含了不同滤波器响应之间的特征相关性，它通过测量同一卷积输出层的不同通道之间的相关性来提供样式的度量。其中 G_{ij}^l 是层中的矢量化特征映射和之间的内积，即第 l 层的第 i 个矢量特征图和第 j 个矢量特征图的内积，用来计算特征之间相关性。在公式中， F_{ik}^l 和 F_{jk}^l 分别表示在第 l 层位置为 k 的第 i 个和第 j 个卷积核的激活值， F^l 是五个卷积层后的响应结果。该公式如下：

$$G_{ij}^l = \sum_k F_{ik}^l F_{jk}^l$$

通过计算多个卷积层之间的特征相关性，捕获其纹理信息。我们可以可视化这些网络层。这是通过使用白噪声图像的梯度下降来最小化原始图像的 Gram 矩阵与要生成的图像的 Gram 矩阵之间的均方距离来完成的。令 \vec{a} 和 \vec{x} 是原始图像和生成的图像，并且 A^l 和 G^l 是它们在层中的相应样式表示。那么第 l 层的损失就是：

$$E_l = \frac{1}{4N_l^2 M_l^2} \sum_{i,j} (G_{ij}^l - A_{ij}^l)^2$$

总的风格损失是：

$$\mathcal{L}_{style}(\vec{a}, \vec{x}) = \sum_{l=0}^L \omega_l E_l$$

其中 ω_l 是每层对总损耗的贡献的加权因子。

3.3 VGG-19 卷积神经网络结构

在我们的项目中，使用了 VGG-19 损失函数网络，取 VGG-19 第四层的卷积特征来计算内容损失，取一，二，三，四，五层的卷积特征来计算风

格损失。

内容损失函数使用的是内容图像和生成图像的第四层卷积特征来计算损失函数。

风格损失函数使用到了 GRAM 矩阵，GRAM 矩阵是用来衡量风格特征的一种方法，计算方法为，矩阵乘以其装置矩阵，我们分别提取图像一，二，三，四，五层的卷积特征，分别计算每层卷积特征的 GRAM 矩阵，然后计算风格图像和生成图像 GRAM 矩阵的损失函数，最后加总得到总风格损失。

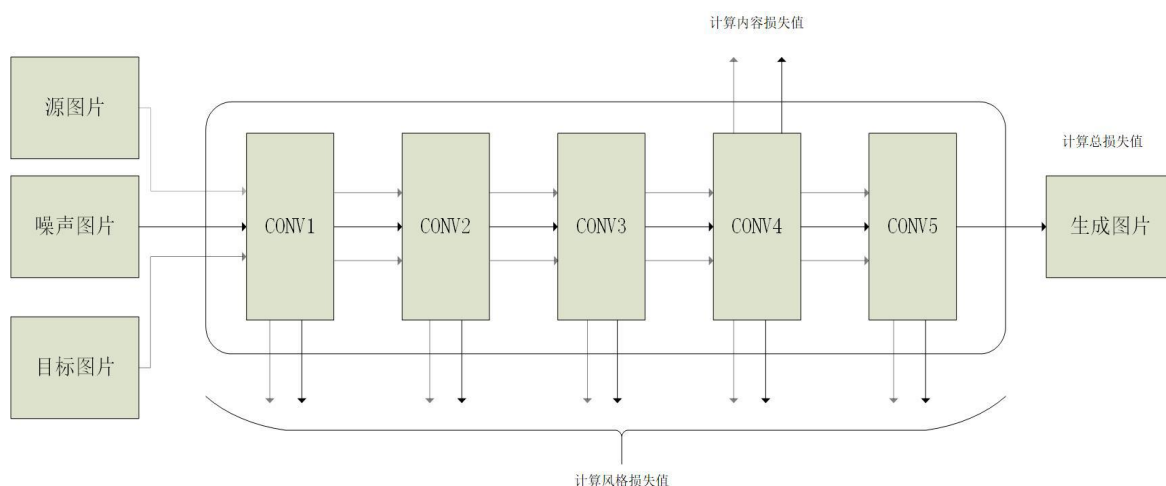
3.4 风格转移总损失函数

为了实现图像的风格迁移需要同步拟合源图像和目标图像的特征表示，从而融合生成一幅新的图像。因此，通过不断优化网络中每一层的内容特征图与风格特征图联立的损失函数，迭代至总损失值趋于全局最小，即完成风格迁移。构建的损失函数如下：

$$L_{total}(\vec{p}, \vec{a}, \vec{x}) = \alpha L_{content}(\vec{p}, \vec{x}) + \beta L_{style}(\vec{a}, \vec{x})$$

其中 α 和 β 分别是内容和样式重构的加权因子。

风格转移的实现流程如图所示：



输入网络的图片有三张，最终将通过网络训练输出图片。途中不同输入图片的训练和计算流向以用颜色区分。内容的损失函数采用的是 conv4 中的卷积层来计算，风格损失值取一，二，三，四，五层的卷积特征来计算风格损失。最后计算总损失值。

四、实验结果分析

4.1 实验搭建环境：

实验采用 python 语言编写，采用 pytorch 框架，使用 CPU 进行训练。

本项目使用一张校门口的图片作为内容图像，想使它展现出不同的艺术效果，选择具有代表性的《星空》和其他不同风格的画作作为风格图像，进行了多次实验。

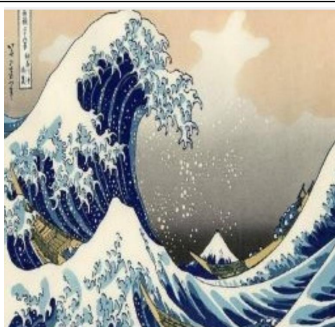
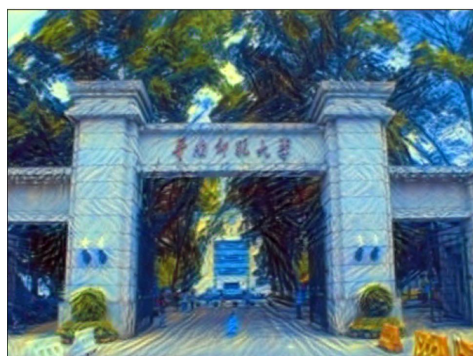
4.2 实验过程和结果：

以下是该实验使用古今中外不同的艺术图片风格迁移到学校校门口图片的生成图（都采用 jpg 格式的图片）：

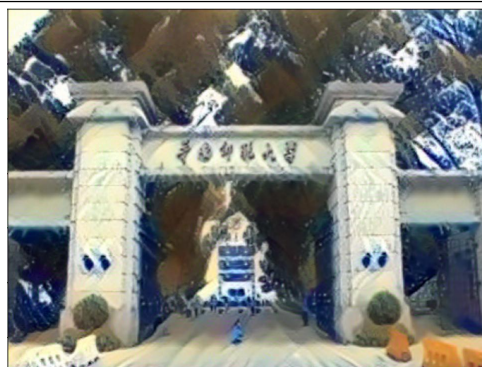
| 风格图片 | 内容图片 | 生成图片 |
|---|--|---|
|  张择端 清明上河图 |  |  |



梵高 星空



神奈川冲浪



齐白石 虾



徐悲鸿 马

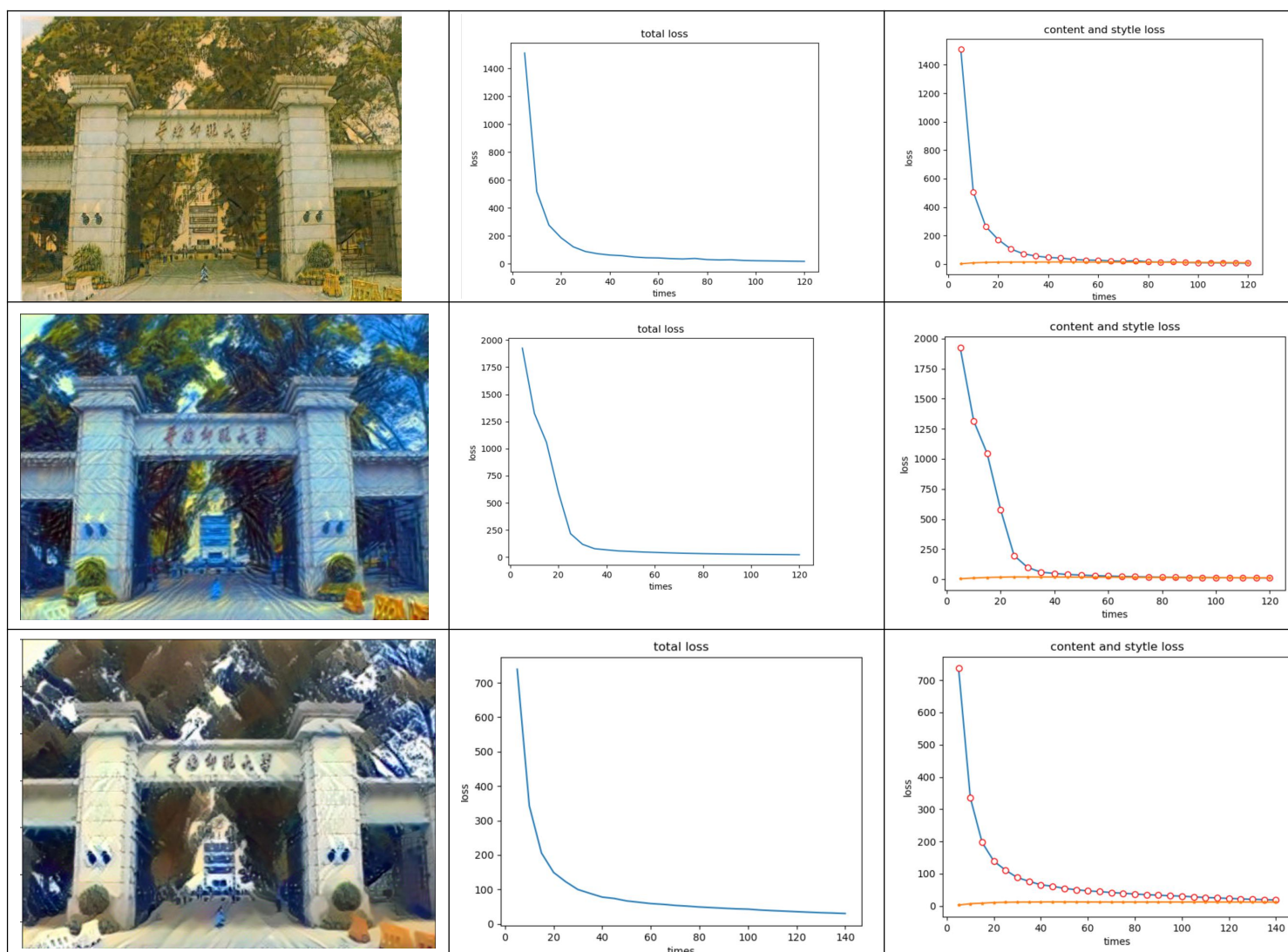


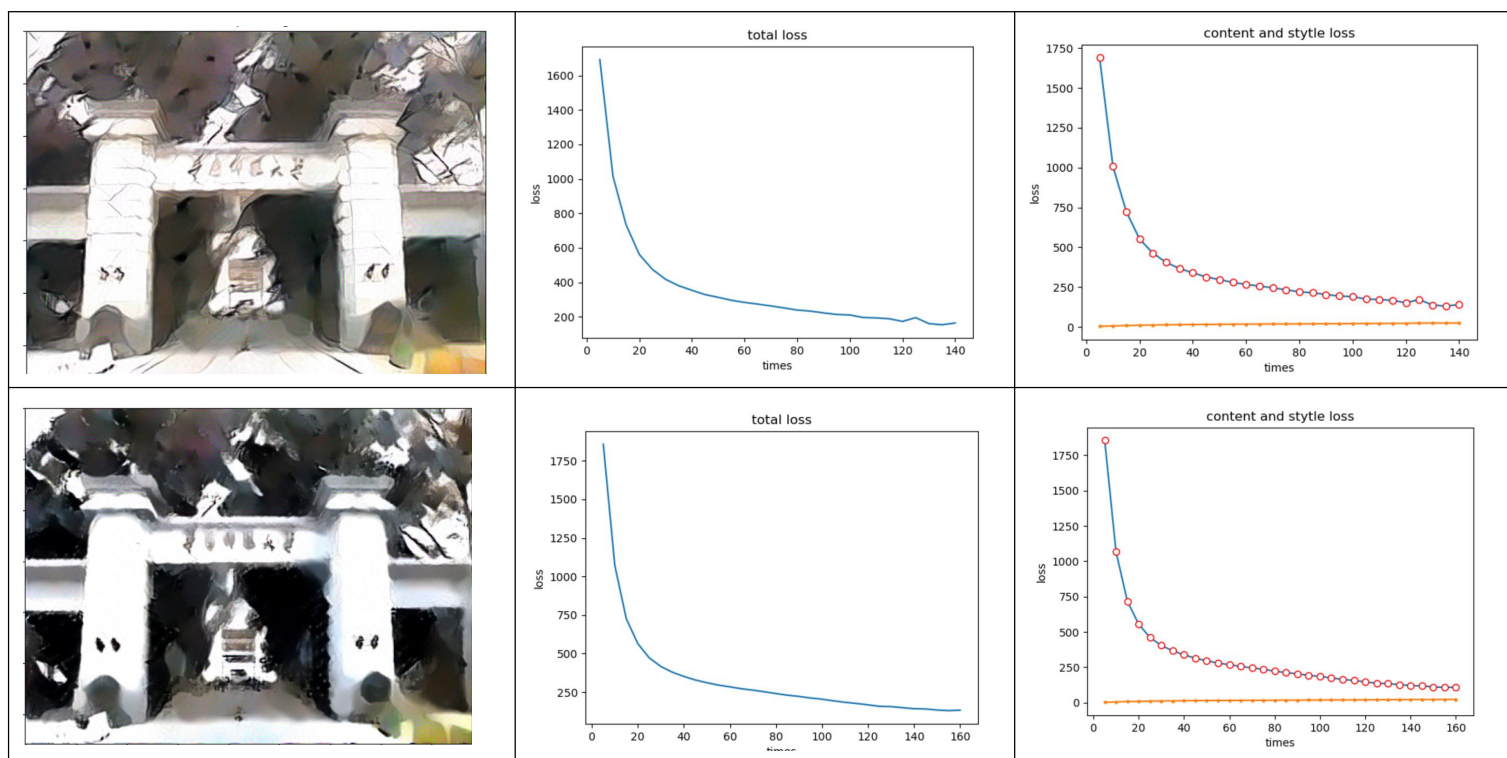
下面是使用各种风格训练过程中损失函数的变化情况，第二列是总损失函数，第三列中的蓝线表示风格损失，橙色的线表示内容损失。总的损失一开始急剧下降，后面趋缓，最后几乎看不出下降。而风格损失则是一

开始很大，后面急剧下降，内容损失很慢很慢地上升。这是因为我们一开始拷贝了内容图片来定义原始噪声图片，这使得一开始看不出内容损失，而风格的损失却很大。

而我们定义的风格损失和内容损失的超参数权重为 1000000 和 1。这也使得风格损失能够快速地下降。

从各个生成的图像和它们的损失函数可以发现，前面三组图片的迁移效果较好，而后面的两幅图片迁移效果较差，这说明迁移的效果如何得看具体的图片风格类型。





五、结论

项目设计了一种艺术风格迁移的算法，可以分离和重新组合图像内容与众多众所周知的艺术品的外观相结合。利用 VGG-19 网络学习提取特征，然后不断反馈，更新目标图像，最后得到要生成的图像。

由于我们小组没有使用 GPU 环境进行训练，只用 CPU 进行训练，也没有采用快速风格迁移的方案，所以训练时间会比较长，如果图片的像素低，训练次数少的话，5 分钟可以迁移出目标图片。但是如果图片像素高，训练次数多的话，迁移每张图片所需的时间会达到 10 分钟以上。

在实验过程中，我们还发现，使用不同的内容图片和风格图片的迁移效果是不一样的，迁移的效果如何得看具体的图片类型。当然，也可以通

过调整超参数或者增加训练次数来改变生成的图像效果。

六、参考文献

- [1] 王婷,李航,胡智. 一种 VGGNet 的图像风格迁移算法设计与实现[J]. 计算机应用与软件, 2019, 36(11):224-228.
- [2] Gatys L A, Ecker A S, Bethge M. A neural algorithm of artistic style [J]. arXiv preprint arXiv. 2015: 1508. 06576
- [3] 陈淑環 , 韦玉科 , 徐乐 , 等 . 基于深度学习的图像风格迁移研究综述 [J]. 计算机应用研究 , 2019, 36(8) : 2250-2255
- [4]王鹿,曾国辉,黄勃. 基于深度学习的风格迁移算法的研究与实现[J]. 智能计算机与应用, 2020, 10(02):57-60+65.
- [5]牟晋娟. 基于深度学习的图像风格迁移技术的研究[J]. 电子元器件与信息技术, 2019(04):82-85.
- [6] 卷积神经网络研究综述[J]. 周飞燕, 金林鹏, 董军. 计算机学报. 2017(06)
- [7]张月,刘彩云,熊杰. 基于 VGG-19 图像风格迁移算法的设计与分析[J]. 信息技术与信息化, 2020(01):70-72.