摘要

考古图像的虚拟修复，包括但不限于对绘画、壁画、器物彩绘、出土时有裂痕的物品、旧照片的修复，需要填补缺失部分的图像信息，在原理上和数字图像修复有共通性。先人在日常生活中所使用的的物质，构成了如今我们所见的文物，成为历史与文化的载体，有金石、陶瓷、绘画等介质。对于不同材质的文物，由于理化性质的不同，产生的缺损的形式也会有所不同，所以虚拟修复方案应有所区别。针对文保领域的需要，本文将探讨以壁画修复为研究对象的，基于深度学习的图像补全(Image Inpainting)的虚拟修复手段。并根据修复算法实施在文物图像修复上的效果，以“修复如初”为原则，提出对相关领域进行数字化修复时需要改进的方向。

关键词：深度学习，图像修复，文物修复，壁画

Abstract

目录【完成后填写】

1.问题回顾Introduction

1.1壁画常见的病害类型

受到自然环境的影响，壁画往往在长期的侵蚀破坏下产生不同种类的病害，壁画古代壁画修复时应充分考虑到其病害特征。病害的特征包括以下的种类，例如支撑体与壁画的地仗层部分脱离产生空鼓、地仗层完全脱离产生大面积脱落、底色或颜料层产生的网状的龟裂、小尺寸的粉化脱落[5]、以及动物损害等。【图[http://kns.cnki.net/KXReader/Detail?TIMESTAMP=637211142310401250&DBCODE=CJFQ&TABLEName=CJFDLAST2019&FileName=YCBH201902027&RESULT=1&SIGN=yhzYSY%2frENlltyd%2baxWkepjkB%2fw%3d](http://kns.cnki.net/KXReader/Detail?TIMESTAMP=637211142310401250&DBCODE=CJFQ&TABLEName=CJFDLAST2019&FileName=YCBH201902027&RESULT=1&SIGN=yhzYSY/rENlltyd+axWkepjkB/w=)】对于不同种类的病害，损坏区域的形状及面积不尽相同，修复手段也有所区别。本文中将选取大面积脱落以及小面积粉化脱落两种类型的病害来进行虚拟修复的研究。

1.2壁画修复工作的历史

最早的文物图像修复起源于文艺复兴时期的欧洲，当时的艺术家根据自己的经验和理解，对前人的画作进行复原。例如1565年，皮埃特罗·卡鲁奈模仿米开朗基罗的画风修复了其在梵蒂冈西斯廷教堂的湿性壁画[1]。但传统的图像修复依赖修复者的个人技巧，其修复周期长，且一旦出现偏差，将对文物本身造成不可挽回的毁坏。

随着现代计算机科学的发展，计算机视觉领域对图像处理作出了很多工作，而相关的成果也被广泛运用到实际工作中来。例如，王展等人[2]利用Criminisi算法对四川新津观音寺的明代壁画修复展开了研究；Wolfgang Baat[3]等人则利用CDD(Curvature Driven Diffusion,曲率驱动扩散)这一基于全变分(Total Variation,TV)模型的图像去噪方法来对奥地利维也纳发现的尼德哈特(Neidhart)壁画进行了虚拟修复。然而，以上算法在实际应用中都存在一定的问题和局限性。Criminisi算法的缺点有权重的可靠性不佳，以及patch搜索错误匹配率高等问题。CDD模型存在着边缘的视觉效果不自然、耗时长、迭代复杂等缺陷。

1.3基于深度学习的图像修复方法

近年来基于深度学习的图像修复(Image inpainting)手段日臻成熟。图像修复的途径主要分为三种，即基于序列模型(Sequential-based)，基于对抗生成网络(GAN-based)，和基于卷积神经网络(CNN-based)[4]。深度学习方法解决图像修复问题的算法研究成果丰富，对于来自网络上各大开源的图像数据集的修复处理已有广泛的实践。较之于传统的数学方法，深度学习修复效果更为出色。

古代壁画的虚拟修复，与图像修复在原理上相通。现有的修复方法可以作为参考和借鉴，应用到壁画，乃至于各种文物的图案修复工作中来。需要注意的是，目前深度学习的研究者在其修复样例中经常采用的掩膜覆盖方式一般为矩形，这与大面积缺失的形式类似，但是同小面积不规则缺失的缺损形式仍有一定的区别。这种情形下，我们需要向研究对象添加不规则的噪点，然后进行修补。在考古学的实际工作中，壁画常出现大面积的图像缺失。深度学习方法的修复往往是基于在大量的训练模型基础之上形成的幻觉，来对缺失区域进行图案上的预测。因此，当缺损区域较大时，深度学习方法修复的结果可能会与实际样式有一定的出入。



在风格上，古代和现代的作品之间也存在明显差异，而互联网上开源的数据集的组成多为现代风格图像或者自然风景等，对于古代艺术作品的图像收录较少，运用现有的较成熟的数据集，如CelebA和ParisStreetView，修复结果可能不尽如人意。因此，对于虚拟修复古代壁画的深度学习方法，仍需要进行探究，以期获得在视觉上最为合理的修复结果。

基于考古行业的需要，我们应采取不需要掩膜即可训练神经网络的方法，这样在修复时可对任意形状的掩膜进行修复，对于没有计算机科学专业背景的文保工作者可以降低其工作的难度。本文中，我们利用对抗生成网络(Generative Adversarial Network,GAN)进行语义图像修复(Semantic Inpainting)。Pathak[6]等提出的ContextEncoder(CE)是语义图像修复方面的开创性方法，是一种基于编码器-解码器结构的对抗生成网络的方法。这种方法给定了缺失区域的掩膜，以此训练神经网络进行上下文的编码，进而完成对缺损区域的预测。这种方法的缺点是，在缺损形状随机的情况下易造成修复出的图像模糊，它只在训练过程中而并未在推断过程中利用掩膜的结构。对于不同形状的掩膜也需要单独训练。Nazeri[7]等提出的基于对抗边缘学习的方法，从对缺失部分的边缘轮廓计算入手，在第一级网络中生成修复对象的完整的边缘信息，这将成为下一级修复网络的先验信息。随后轮廓与原图像一同作为神经网络的输入进行修补。Yeh[8]等采用的基于DCGAN(Deep Convolutional GAN)的语义图像修复方法，可以得到更为自然的修复结果。可以说，现今深度学习领域的日新月异的成果，很好地迎合文保工作者的工作需要。

此外，由于进行的虚拟修复并不破坏文物本身，其目的是为了将文物的原貌展现与世人，故可以对图像进行图像降噪、色彩复原等操作。跨平台的Opencv库向我们提供了Color-enhance插件，对于图像的色彩优化不失为一个良好的解决方案。本文将采用基于DCGAN的深度学习方法和OpenCV相结合图像修复方法的组合方法来完成对古代壁画的虚拟修复。

【文章结构】

第一章 绪论

第二章 相关工作

第三章 介绍自己的方法

第四章 实验

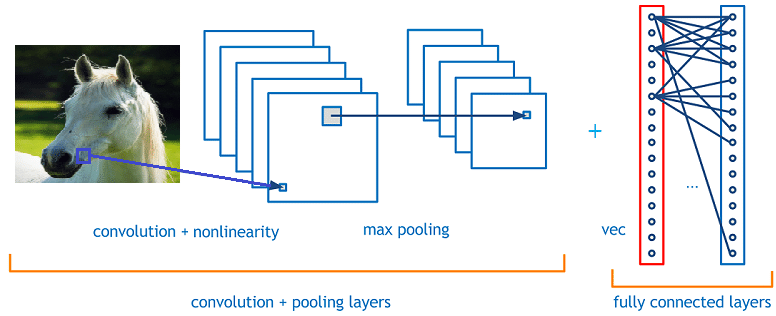
第五章 结论和展望

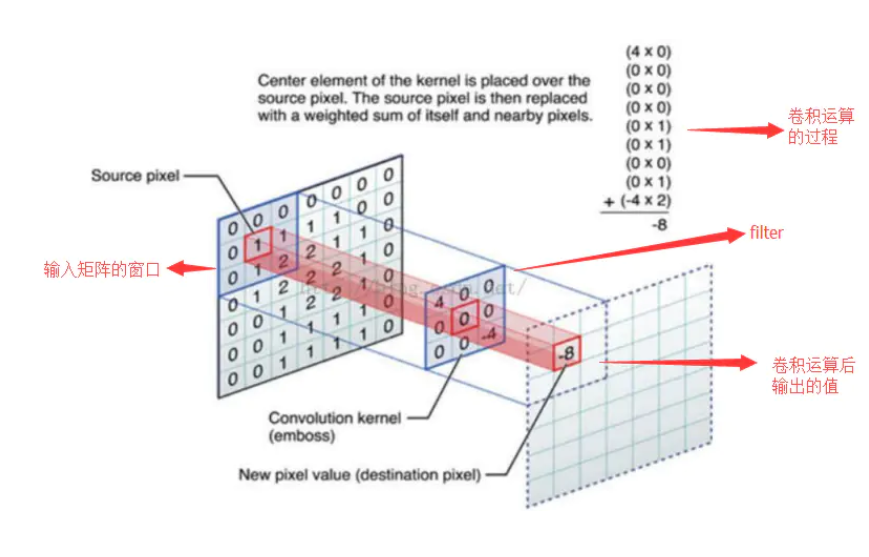
1. 方法综述Related Work

2.1卷积神经网络

卷积神经网络(Convolutional Neural Network,CNN)是深度学习的常用的算法之一。如图【】所示的是一个典型的CNN网络架构。卷积神经网络包括卷积层、池化层和全连接层组成。这是一种可以有效地提取图像特征的方法。

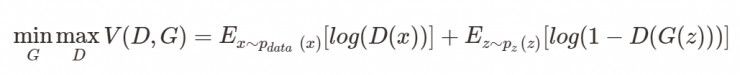
在卷积层中会使用正方形的卷积核，来遍历图片上的每一个像素点。卷积核覆盖的区域，又称为“感受野”，其中每一个点都有权重。如果使用3\*3大小的卷积核，则将此区域内的九个点做加权求和，然后加上偏移量，得到一个点的像素值。池化层作用是保留图片主要的信息，去除冗余信息，来压缩图像。出于特征的不变性，即使去掉了一部分冗余，仍可以表达出图片的特征，保留对于图像处理任务关键的特征，称为特征降维。池化层常用Max pooling方法，简而言之，是将图片分为若干个同等大小的块，每块里取一个最大的节点，来代表所在的块，与其他块取出的节点组合到一起构成输出，可以大大地压缩图像的大小。全连接层通过对特征值进行卷积处理，将特征整合到一起，来对数据进行分类，减少关键特征因位置不同对网络精确性的影响，可以增强神经网络的鲁棒性。





2.2生成对抗网络

生成对抗网络(Generative Adversarial Network,GAN)是一个无监督的学习过程，主要包括：负责生成真实图像的生成模型G，和负责对抗G，即判别G产生的图像的判别模型D,被其提出者Goodfellow[9]等比作两个人的博弈。G不断地从先验分布中取样，并生成着图像来交给D判断，试图骗过D的判断使之认为图片为真实，D则试图判断出G提交的图片的异常。在这样的博弈过程中生成模型和判别模型的“进化”，使得最终整个生成对抗网络能够得到感知上最为真实的图片生成结果。原始论文[9]中将D和G的优化过程看作一个最大最小优化问题,对其这给出了这样的目标函数：

(1)

Pdata(x)中的x是来自真实数据分布的样本，z是潜在空间的随机编码，D(x)是判别网络输出的图片是否为真的概率，取值大于0，小于1。

对于G的优化，(1)式取最小值：

IMG_256

优化生成器时应该以蒙骗过判别器D为目标，那么应当追求D(G(z))尽可能的大，所以log(1-D(G(z)))尽可能小。

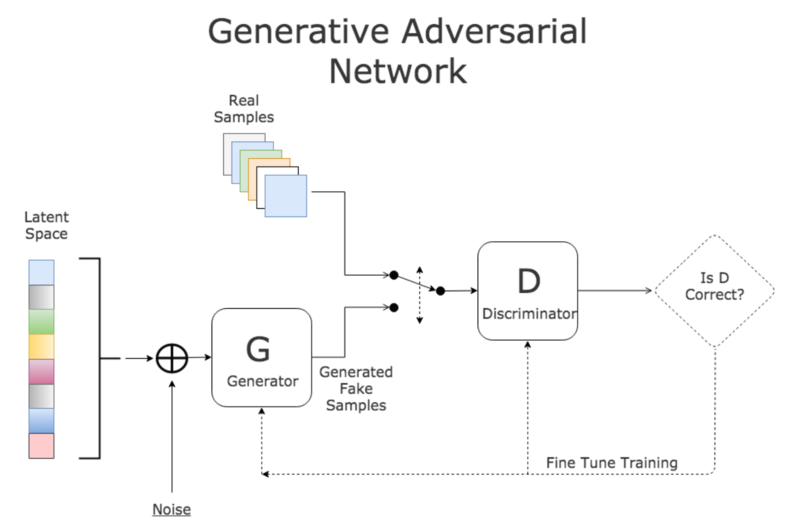
对于D的优化，(1)式取最大值：

IMG_256

与G的优化方向相反，D要尽可能地判断出G制造的假样本，则D(G(z))应尽可能小，对于来自于真实样本的输入x，则应该让D(x)尽可能大。

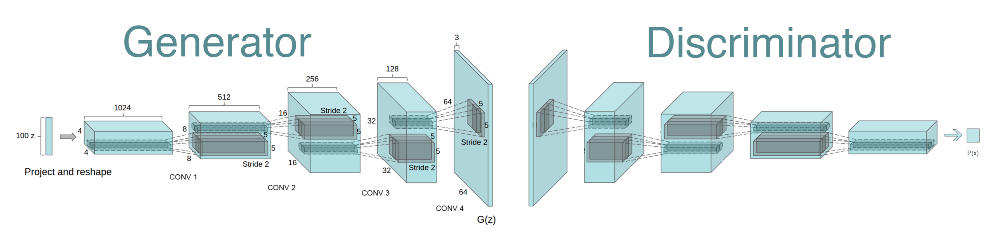
生成器G和判别器D在这样的过程中不断进步，最终要达到的目标是G生成的样本的分布pG与真实样本的分布pdata一致，即pG=pdata。

即D(G(z)) = 0.5。



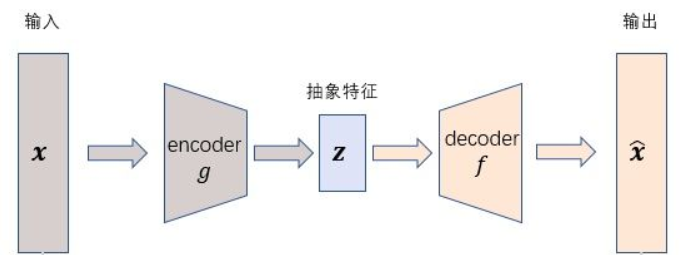
本文中使用的深层卷积生成对抗网络(Deep Convolutional GAN,DCGAN)，无论是生成器G，还是判别器D，都引入了卷积神经网络的思想。其中，多层感知器(Multilayer Perceptron)的将多个数据集映射到单一的输出数据集的功能由卷积神经网络来实现。相比于GAN或者CNN网络，DCGAN的优势在于整合了卷积神经网络的数据分析能力和生成对抗网络的数据生成能力。

无论是普通的GAN还是DCGAN，其训练的过程仍保持一致：第一步，不停地训练判别器D，使公式(1)取值最大；第二步，控制变量，令D不变，不停地训练生成器G，并使公式(1)取值最小，等价于D(G(z))取值最大。以上两步就是G和D之间的博弈，是生成对抗网络的基本思想，这样的过程需要不断地重复，直到判别器和生成器之间的博弈达到D(G(z)) = 0.5的纳什均衡。



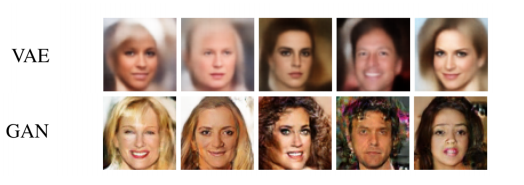
2.3自编码器

自编码器（Auto Encoder）是一种无监督学习的神经网络，由编码器(Encoder)和解码器(Decoder)两部分网络组成，引导着编码和解码两个过程。自编码器被用于数据的降维，及对数据特征的提取。在这一体系中，编码器将输入x做映射，到特征空间z，即编码过程。解码器再将z作为输入，将特征再度做映射，回到原始空间得到x’，网络的优化目标则是使得x与x’的分布尽可能地接近。



【这个图是网上找的】

变分自编码器(Variational Auto Encoder,VAE)[10]是自编码器(Auto Encoder,AE)的一种，利用了变分和贝叶斯公式。VAE作为一种常用的无监督网络模型，常与GAN作为比较。以目前流行的人脸图片修复研究为例，通过对比图片【】的视觉效果，可以看出目前VAE进行的图像修复时的输出相较于GAN的输出，存在着较为明显的输出图像模糊的问题。因此，本文中选择GAN而非VAE方法来探究对古代壁画的修复。



图x：来自VAE和GAN的人脸修复结果范例

2.4OpenCV

OpenCV是一个开源的计算机视觉库，适用于图像处理、物体识别、人机交互等方面的问题。其突出优势在于跨平台，且轻量、经过了大量的优化，为诸如Python,Java,Matlab,Ruby等语言提供使用接口，使得图像处理更易于进行，帮助其实现高效率的计算。OpenCV中提供有多种强大的的图像去噪函数，例如高斯模糊去噪、均值去噪、非局部均值去噪等。在诸如文献[11,12,13]的深度学习研究中，均采用增加高斯白噪声来模拟实际存在的复杂的来源的环境噪声。对于小面积且密集的粉化脱落类型缺失的古壁画，其特征与高斯白噪声类似。可以添加高斯白噪声来模拟其病害，并通过高斯模糊去噪等方法完成对相关类型的病害的修复。OpenCV也可以用于图片的对比度增强。对于灰度图片的对比度增强，常采用的方法是直方图均衡化；而彩色图像的彩色增强也可以用类似的办法，即RGB图像分通道进行直方图均衡化。进行在经过生成对抗网络的生成之后，可以对输出壁画图案再经过OpenCV进行诸如降噪和彩色增强处理，来增进修复的视觉效果。

1. 结合OpenCV的基于DCGAN的图像修复

3.1相关的网络模型设计

基于生成对抗网络的基本原理，本文中的网络结构中需要有一个生成器G和一个判别器D。神经网络需要经过以原始图片为数据集的训练过程，来得到一个先验分布pdata。在一般的生成对抗网络中，模拟人类的感知功能的组件是多层感知器，这是一种前馈的神经网络。它通过限制隐含层节点的方法，使得输出侧节点少于输入侧，对数据进行降维，对特征进行提取。这样的功能在深层卷积生成对抗网络中使用卷积神经网络作为替代。同一个特征映射层的卷积神经网络的神经元间存在着权值共享的结构，使CNN网络可以并行学习，这可以很大程度上降低运算的复杂度。

生成器将从噪声中提取出样本点，并将采样得到的向量zˆ映射到图像空间X上。既然一系列复杂的学习的最终目标是G的生成产品可以混淆判别器D的判断，那么生成的图片在分布上要尽可能地接近真实图片的分布。如果生成器G的进化程度足够高，那么在其编码流上应该出现的图像都应符合真实数据的分布pdata，所以问题等同于寻找最接近于原分布的向量zˆ。而这一问题被Goodfellow等人在论文中[9]归结为一个最大似然估计问题，并在一系列复杂的数学推导之后提出目标函数(1)，在本文的第二部分有给出。在Yeh等人的论文中[8],对最为接近真实数据样本的zˆ的代数衡量如下：



在这里我们参考Johnson的研究[15]，对zˆ作出如下的代数衡量：

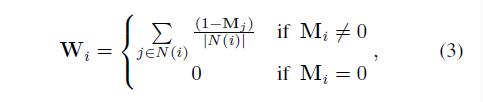
zˆ=arg min {αLc+βLp}

Lc是语境损失函数，用于限制G得到的从网络外部输入的已被损毁的图像和掩膜，来保证G可以生成符合原图片图案特征与内容的像素点。Lp是惩罚函数，代表着与真实图像的分布差别较大而被判别器D鉴别出时的罚分，计入到计算式中去。

3.2损失函数

3.2.1语境损失函数

对于一张有缺损的图片的修复，应充分地结合图片剩余部分的信息来推断缺损部分的内容，这样才能使生成器产生合理的图像输出。选取合适的像素点，则需要结合语境损失(Context Loss)函数来进行。类比人类语境的限制和补充所讲内容的意义，这里所提到的语境，指代的是取样过程中对选取像素点的权重限制：选取的像素点应当尽可能的与缺损区域的内容有相关性，所以缺损区域附近的图案显然更为重要，应当对这一区域的点更多地采样。我们需要量化每个像素点的重要性，因而规定赋予每个点权重的方法：

[](https://imgchr.com/i/32dnBT)

在这里M是图像二值掩膜，i是像素点的顺序，我们用Wi来表示第i个像素点的权重。|N(i)|代表着当前的像素点在采样窗口中的个数。对于惩罚函数的惩罚项的设置，在机器学习领域常用到l1和l2正则化(norm)。l1-norm指各元素的绝对值之和，它可以产生一个稀疏模型；l2-norm指各元素平方和的平方根，可以防止模型过拟合。参考计算机视觉领域的文献[14]可知基于l1正则化的回归算法(LASSO回归)具有更好地从稀疏模型中提取数据特征的能力，且相比l2正则化有更好的鲁棒性。因此在语境损失函数中，我们将原有的完好部分（也就是修复工作所基于的“语境”），与填补后的图像做l1范数差，得到以下的语境损失函数表达式

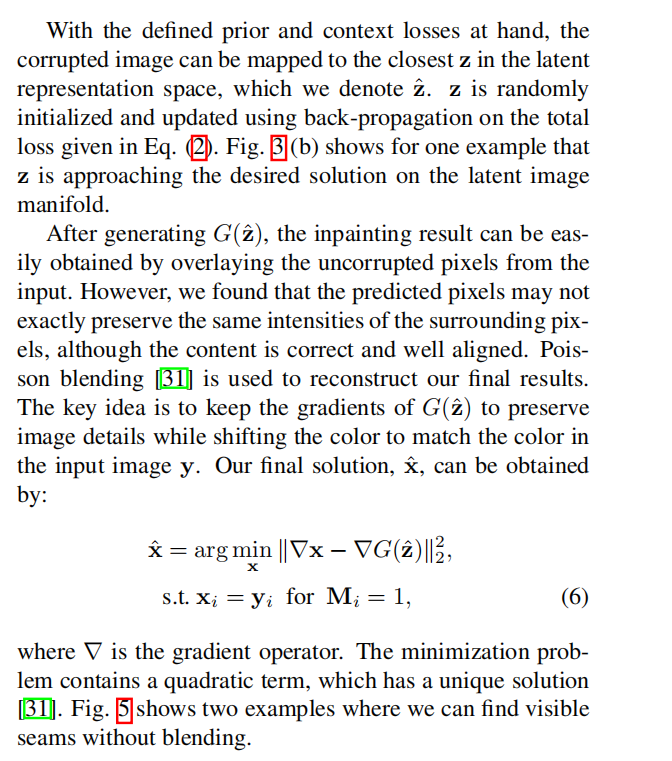
[IMG_257](https://imgchr.com/i/32wOyt)

3.2.2先验损失函数

为了减少在机器学习中出现的过拟合，可以在损失函数后添加一个先验损失函数，来增强模型的泛化能力，确保这个模型不只是针对特定的对象起到作用。先验分布可以理解为同类样本的统计概率，可以对模型的大致特征有一个基本的刻画。对于()式中的先验损失函数Lp ,应该作为判别器D的工作规范，起到鼓励生成器G生成与真实数据接近的图像，并惩罚与原图像差异大的图像的作用。所以，可以参照判别器D的目标函数()来定义先验损失函数，即

Lp=(1-D(G(z)))

3.3Inpainting



1. 实验Experiments
   1. 数据集和掩膜Datasets and Masks

4.1.1获取图像

4.1.2对图像进行预处理

4.1.3训练数据集

* 1. 视觉上的比较Visual Comparisons
  2. 定量比较Quantitative Comparisons

4.4讨论 Discussion

5.结论 Conclusion

引用文献

[1]詹长法.意大利现代的文物修复理论和修复史(下)[J].中国文物科学研究,2006(03):92-95.

[2]王展,王慧琴,吴萌,陈卿.新津观音寺明代壁画图像的计算机自动虚拟修复研究[J].文物保护与考古科学,2018,30(03):109-113.

[3]Baatz W,Fornasier M , et al. (2008). Inpainting of Ancient Austrian frescoes[J]. Proceedings of Bridges,2008:150-156.

[4]Elharrouss O , Almaadeed N , Al-Maadeed S , et al. Image inpainting: A review[J]. Neural Processing Letters, 2019.

[5]温利龙. 基于神经网络的古壁画破损修复与风格复原研究[D].云南大学,2019.

[6]Pathak D , Krahenbuhl P , Donahue J , et al. Context Encoders: Feature Learning by Inpainting[J]. 2016.

[7]Nazeri K , Ng E , Joseph T , et al. EdgeConnect: Generative Image Inpainting with Adversarial Edge Learning[J]. 2019.

[8]Yeh R A , Chen C , Lim T Y , et al. Semantic Image Inpainting with Deep Generative Models[J]. 2016.

[9]Goodfellow I, Pouget-Abadie J , et al. Generative Adversarial Nets[J]. ArXiv.2014

1. Kingma D P, Welling M.Auto-Encoding Variational Bayes[J]. stat, 2014, 1050: 10.
2. Dai, Jifeng, Qi, Haozhi, Xiong, Yuwen,等. Deformable Convolutional Networks[J].
3. Min J , Carlini L , Unser M , et al. Fast live cell imaging at nanometer scale using annihilating filter-based low-rank Hankel matrix approach[C].Wavelets and Sparsity XVI. International Society for Optics and Photonics, 2015.
4. Fan, Qian, Zhang, Lifeng. A novel patch matching algorithm for exemplar-based image inpainting[J]. Multimedia Tools & Applications.
5. 张振月. 基于范数正则化回归的人脸识别[D]. 2016.
6. Johnson, Justin, Alahi, Alexandre, Fei-Fei, Li. Perceptual Losses for Real-Time Style Transfer and Super-Resolution[J].