基于 SRGAN 模型的监控图像增强应用研究

(临沂大学信息科学与工程学院，山东 临沂 )

张毅 沈可心 刘凯丽 林霞

**摘 要** 针对部分监控图像清晰度低所带来的辨识难度大，分析了SRGAN模型超分特性，提出了一种基于超分算法的web应用平台的解决办法。采用感知损失作为模型的训练指标，搭建生成器网络和判别器网络，对生成器网络和判别器网络进行交替训练，通过最小化重建损失学习如何生成视觉上接近于高分辨率图像的图像。结果表明：基于SRGAN模型实现的监控图像超分不仅在像素级别上重建图像，而且在视觉质量上进行优化，生成更加细腻和逼真的超分辨率图像。

**关键词：** SRGAN模型、监控图像、感知损失

**1 研究背景**

监控图像作为获取事件信息的重要手段，在公共安全、工业生产、交通运输等领域中发挥着关键作用。它们能够提供连续的视觉记录，帮助相关人员迅速、清晰地捕捉和分析目标特征信息，从而有效支持违法追责、维修调整等后续工作的进行。然而，监控图像的质量往往受限于诸多因素，如设备性能、环境遮挡和光线干扰等，这些不确定因素可能导致监控图像清晰度不高、模糊难以识别。监控图像低质量问题给信息的准确获取和分析带来了显著困难，影响了监控系统的有效性。

为解决监控图像质量受限的问题，采用SRGAN模型进行图像超分，通过web服务端进行部署。利用VGG的卷积神经网络，特别是其中间层输出的特征，来捕捉与人类感知相似性接近的图像细节。这种方法采用感知误差替代传统超分辨率算法中常用的逐像素均方误差，从而提高了监控摄像头捕捉到的低分辨率图像的质量，克服了超分辨率处理后图像缺乏纹理细节的问题，提高监控视频图像的清晰度和可识别性，使得相关人员能更有效地分析和利用这些图像，以便于更好地进行后续的监控和安全管理工作。

**2 SRGAN模型构建**

**2.1数据集简介**

模型训练采用苹果林地图像数据集作为基础训练集，图像包括苹果、果树、草地等类型，图像尺寸大小均为270像素×185像素，在生成式对抗网络中需要提供一组高分辨率图像和一组低分辨率图像作为目标数据和输入数据，传统方法是通过对高分辨率图像进行降采样，主动生成低分辨率的输入图像。但是在该过程中难免会出现一些误差导致低分辨率图像丧失原高分辨图像的信息。因此，通过采用同时提供低分辨率图像和高分辨率图像的方式，通过AI技术将原低分辨率图像提升清晰度并放大四倍作为高分辨率图像，生成200对低分辨率和高分辨率一一对应的训练数据集[1]。

如图1所示，低分辨率图像的规格为270像素×185像素，如图2所示，高分辨率图像的规格为1080像素×740像素。

图1 低分辨率图像 图2 高分辨率图像

**2.2 SRGAN超分模型生成器网络**

与传统的逐像素均方误差的方式不同，SRGAN模型采用adversarial-loss和content-loss组成的感知误差作为训练生成器的标准，由低层特征提取、 高层特征提取、转置卷积层以及CNN重建层组成生成器网络。如图3所示，SRGAN模型生成器网络可以用于实现低分辨率图像超分为高分辨率图像。

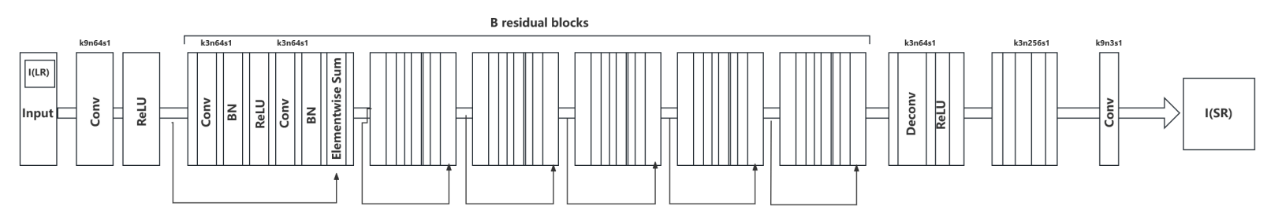


图3 生成器网络构建

生成器网络设计低层特征提取采用一个9×9的卷积核、64个特征图、步长为1的卷积层搭配一个PReLU激活函数的配置；其后设计一系列的残差块，每块由两个3×3的卷积层构成、特征映射数为64，步长为1的卷积层搭配批量归一化(BN)和PReLU激活函数[2]，为维持网络训练中的信息流，每个残差块还包含一个元素级相加操作；上采样层使用PixelShuffler卷积层，加强图像的空间分辨率的重构，最后通过一个9×9的卷积核、3个特征映射、步长为1的卷积层输出生成最后的高分辨率图像。

相关代码如下：

1 class ResidualBlock(nn.Module):

2 def \_\_init\_\_(self, in\_features):

3 super(ResidualBlock, self).\_\_init\_\_()

4 self.conv\_block = nn.Sequential(

nn.Conv2d(in\_features, in\_features, kernel\_size=3, stride=1, padding=1),

nn.BatchNorm2d(in\_features, 0.8),

nn.PReLU(),

nn.Conv2d(in\_features, in\_features, kernel\_size=3, stride=1, padding=1),

nn.BatchNorm2d(in\_features, 0.8),

)

5 def forward(self, x):

6 return x + self.conv\_block(x)

7 class Generator(nn.Module):

8 def \_\_init\_\_(self, scale\_factor=2, num\_residual\_blocks=16):

super(Generator, self).\_\_init\_\_()

self.conv1 = nn.Conv2d(3,64,kernel\_size=9,stride=1,padding=4,padding\_mode='reflect', bias=True)

self.prelu = nn.PReLU()

self.residual\_blocks=nn.Sequential(\*[ResidualBlock(64)for\_inrange(num\_residual\_blocks)])

self.conv2 = nn.Conv2d(64, 64, kernel\_size=3, stride=1, padding=1, bias=True)

self.bn2 = nn.BatchNorm2d(64)

self.upsample = nn.Sequential(

nn.Conv2d(64, 256, kernel\_size=3, stride=1, padding=1,padding\_mode='reflect', bias=True),

nn.PixelShuffle(scale\_factor),

nn.PReLU(),

nn.Conv2d(64, 256, kernel\_size=3, stride=1, padding=1,padding\_mode='reflect', bias=True),

nn.PixelShuffle(scale\_factor),

nn.PReLU(),

nn.Conv2d(64, 3, kernel\_size=9, stride=1, padding=4, bias=True)

)

**2.3 SRGAN超分模型判别器网络**

如图4所示，判别器网络包含多个卷积层，每层后添加归一化(BN)和LeakyReLU激活函数，初始卷积层使用3×3卷积核、64个特征图、步长为1，随后的卷积层逐渐增加特征图数量[3]，特征图经历一系列的卷积层和批量归一化处理后，被展平传递到密集层(Dense Layer),通过一个1024单元的密集层、LeakyReLU激活函数和1单元的密集层，输出一个二分类结果；最后通过Sigmoid激活函数，得到一个介于0到1之间的单一值，用于区分输入图像是真实图像(I(HR))还是生成图像(I(SR))[4]。

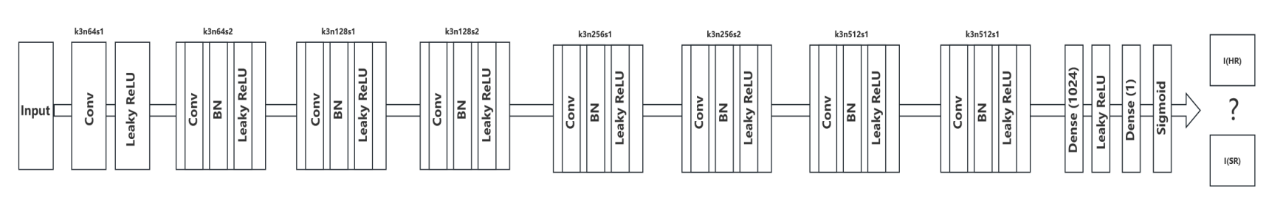


图4 判别器网络构建

相关代码如下：

class Discriminator(nn.Module):

def \_\_init\_\_(self):

super(Discriminator, self).\_\_init\_\_()

self.layer = nn.Sequential(

nn.Conv2d(3, 64, kernel\_size=3, stride=1, padding=1, bias=True),

nn.LeakyReLU(0.2),

nn.Conv2d(64, 64, kernel\_size=3, stride=2, padding=1, bias=True),

nn.BatchNorm2d(64),

nn.LeakyReLU(0.2),

nn.Conv2d(64, 128, kernel\_size=3, stride=1, padding=1, bias=True),

nn.BatchNorm2d(128),

nn.LeakyReLU(0.2),

nn.Conv2d(128, 128, kernel\_size=3, stride=2, padding=1, bias=True),

nn.BatchNorm2d(128),

nn.LeakyReLU(0.2),

nn.Conv2d(128, 256, kernel\_size=3, stride=1, padding=1, bias=True),

nn.BatchNorm2d(256),

nn.LeakyReLU(0.2),

nn.Conv2d(256, 256, kernel\_size=3, stride=2, padding=1, bias=True),

nn.BatchNorm2d(256),

nn.LeakyReLU(0.2),

nn.Conv2d(256, 512, kernel\_size=3, stride=1, padding=1, bias=True),

nn.BatchNorm2d(512),

nn.LeakyReLU(0.2),

nn.Conv2d(512, 512, kernel\_size=3, stride=2, padding=1, bias=True),

nn.BatchNorm2d(512),

nn.LeakyReLU(0.2),

nn.AdaptiveAvgPool2d(1),

nn.Flatten(),

nn.Linear(512, 1024),

nn.Dropout(),

nn.LeakyReLU(0.2),

nn.Linear(1024, 1),

nn.Sigmoid()

)

**2.4 SRGAN超分模型训练测试**

完成SRGAN模型训练后，指定路径图片进行测试，测试生成超分辨率图像，三种图像对比情况如图5所示。



图5 三种分辨率图像对比图

**3 超分平台设计**

**3.1 超分评价指标设计**

采用结构相似性（SSIM）和峰值降噪比（PSNR）两个指标作为衡量模型方案有效性的评价标准。

PSNR的值越高，图像重建的质量越高，计算方式如式(1)所示：

(1)

(2)

式中：MSE——图像I和图像J的均方误差[5]；

m×n——图像尺寸；

I(i,j)——图像I；

J(i,j)——图像J。

SSIM值的范围在0~1之间，SSIM值越大则两张图像的相似度越高，计算方式如式(3)所示：

(3)

其中，,,,

式中：——均值；

——方差；

——图像I的均值；

——图像J的均值；

——图像I的方差；

——图像J的方差；

——图像I和图像J的协方差。

指定路径试用数据集（DIV2K、Set14、Flickr、Set5），加载已训练好的SRGAN生成器模型，将低分辨率图像I(LR)转化为超分辨率图像I(SR),计算转换后的图像和原始高分辨率图像之间的PSNR和SSIM值，测试情况如下表1所示：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 表1 SRGAN模型各数据集SSIM与PSNR数值情况表 | | | | |
|  | DIV2K | Set14 | Flickr | Set5 |
| SSIM | 0.7042 | 0.6327 | 0.7067 | 0.6966 |
| PSNR(dB) | 23.53 | 22.78 | 23.24 | 23.23 |

对比表1测试结果，对于PSNR指标而言，训练后的SRGAN模型对于DIV2K、Set14、Flickr、Set5训练集测试结果分别为23.53dB、22.78dB、23.24dB、23.23dB，说明可以满足基本的人眼视觉的质量要求，也能生成符合模型且MSE要求最小化的图像；对于SSIM指标而言，训练后的SRGAN模型对于DIV2K、Set14、Flickr、Set5训练集测试结果分别为0.7042、0.6327、0.7067、0.6966，说明此方法能较好的保证图像在整体结构上的相似性。

**3.2 超分图像重建**

平台采用训练好的SRGAN模型作为超分页面的服务端技术支持，如图6所示，用户可以在用户端平台的超分页面上传所需超分的原始图像，平台会将原始图像传输到服务端进行保存，并通过训练好的SRGAN模型生成超分图像并保存，超分完成后，自动生成超分历史记录，用户可以在用户端的超分页面直接下载超分图像和原始图像，也可以在历史记录页面下载原始图像和超分图像，同时历史记录页面支持对过往超分记录进行分组设置。

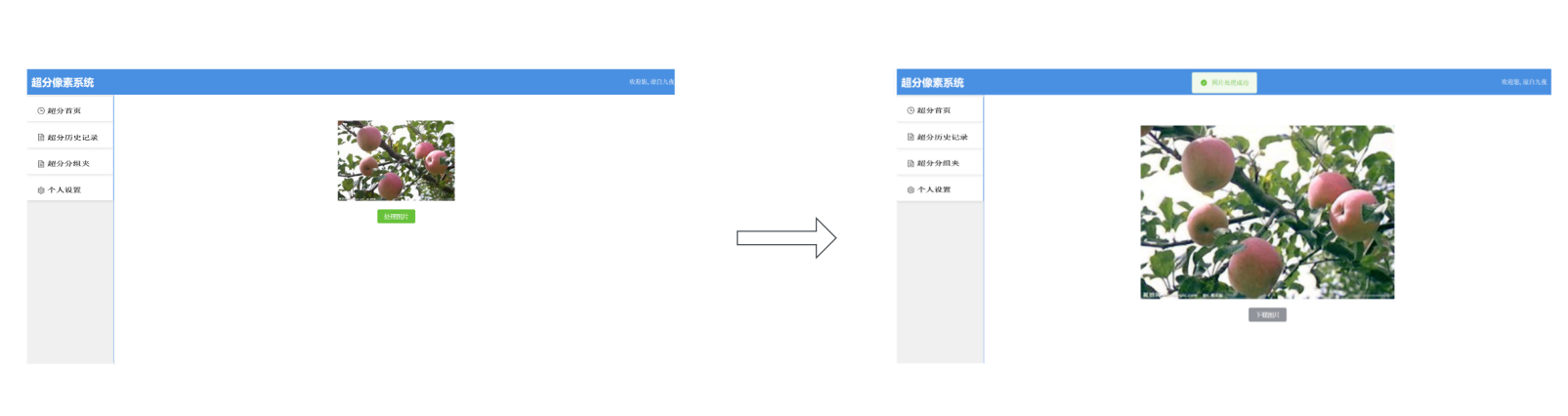


图6 超分图像重建过程

**4 结语**

监控视频作为关键的信息来源，在多个领域中扮演着至关重要的角色。它们提供了持续的视觉记录，帮助相关人员快速、清晰地捕捉和分析目标特征信息，从而有效支持后续工作的进行。然而，监控视频质量受到各种因素如设备性能、环境遮挡、光线干扰等的影响，这些因素可能导致视频图像清晰度不高、模糊难以识别，给信息的准确获取和分析带来了显著的挑战。

为应对这些挑战，开发了一款基于SRGAN模型的超分辨率图像处理平台。利用VGG的卷积神经网络，尤其是其中间层输出的特征，以捕捉与人类感知相近的图像细节。通过采用感知误差替代传统的逐像素均方误差，较为明显地提高了监控摄像头捕捉到的低分辨率图像的质量，并解决了超分辨率处理后图像缺乏纹理细节的问题。这一创新不仅提高了监控视频图像的清晰度和可识别性，而且使得相关人员能够更有效地分析和利用这些图像，为后续的监控和安全管理工作提供了强有力的支持。

通过这一研究，展现了先进图像处理技术在提升监控视频质量方面的巨大潜力。未来的工作将集中在进一步优化这些技术，以适应更多复杂多变的实际应用场景，不断提升监控系统的效能和可靠性。

参考文献

1. 何朝霞,朱嵘涛,徐俊英.基于生成对抗网络的田间杂草图像超分辨率重建[J].中国农机化学报,2023,44(09):154-160.DOI:10.13733/j.jcam.issn.2095-5553.2023.09.022

[2]郑柏慧,陈礼根,刘凤秋.一种改进的生成对抗网络在朱金木雕图像修复中的应用[J].宁波工程学院学报,2023,35(03):31-37.

[3]周紫微.基于SRGAN的工程造价票据超分辨率研究[J].智能城市,2023,9(10):105-107.DOI:10.19301/j.cnki.zncs.2023.10.033.

[4]侯开纪,叶海洋,张莆林等.基于SRGAN的图像超分辨率方法研究[J].现代计算机,2023,29(03):69-72.

[5]朱联祥,郑逸.自注意力SRGAN在岩石CT图像超分辨中的应用研究[J].西安石油大学学报(自然科学版),2022,37(02):131-137.