# **IRSRGAN: 红外图像超分辨率生成对抗网络**

Guannan He Second Author

Institution 1 Institution 2

Institution 1 address Institution 2 address

[firstauthor@i1.org](mailto:firstauthor@i1.org) secondauthor@i2.org

**Abstract**

*超分辨率生成对抗网络（SRGAN）[1]是一项开创性的工作，它能够在单幅图像的超分辨率过程中生成真实的纹理。针对红外图像分辨率偏低的问题，设计了一种改进的针对于红外图像的超分辨率生成对抗网络(Super-Resolution Using a Generative Adversarial Network，SRGAN)算法。为了进一步提高红外图像视觉质量，我们深入研究了ESRGAN的个关键组成部分——网络结构，下采样退化模型，并采用专业的红外图像数据集对模型进行训练，得到了红外图像SRGAN（IRSRGAN）。特别地，我们引入了没有批量归一化并且引入了全局特征提取的残差密集块（C*comprehensive*RDB）作为基本的网络构建单元。此外，我们借用相对论性GAN [2]的思想，让鉴别器预测相对真实度，而不是绝对值。最后，我们在数据增强阶段引入了随机退化模型，使模型能够更好的适应真实世界中的噪声环境。得益于这些改进，所提出的IRSRGAN获得了更真实和自然的纹理获得了持续更好的视觉质量。实验结果表明：所提方法重建的超分辨率红外图像质量在主观及客观评价中均要优于当前具有代表性的方法。该代码可在https://github.com/HeGuannan-duludulu/IRSRGAN上获得*

**关键词：**红外图像；超分辨率重建；生成式对抗网络；残差密集网络

**Abstract**

*The Super Resolution Generative Adversarial Network (SRGAN) [1] is a pioneering work that can generate real textures in the super resolution process of a single image. Abstract: To solve the problem of low infrared image resolution, an improved algorithm based on Super-Resolution Using a Generative Adversarial Network (SRGAN) is designed. In order to further improve the visual quality of infrared images, we deeply study the key components of ESRGAN -- network structure and downsampling degradation model, and use professional infrared image data set to train the model, and obtain infrared image SRGAN (IRSRGAN). In particular, we introduce residual-dense blocks (CRDBBS) without batch normalization and global feature extraction as the basic network-building unit. In addition, we borrow the idea of relativistic GAN [2] and let the discriminator predict relative truth rather than absolute value. Finally, we introduce the stochastic degradation model in the data enhancement stage, so that the model can better adapt to the noise environment in the real world. Thanks to these improvements, the proposed IRSRGAN achieves a more realistic and natural texture to achieve consistently better visual quality. The experimental results show that the quality of the super-resolution infrared image reconstructed by the proposed method is superior to the current representative methods in both subjective and objective evaluation. This code can be obtained at* [*https://github.com/HeGuannan-duludulu/IRSRGAN*](https://github.com/HeGuannan-duludulu/IRSRGAN)

***Keywords:*** *infrared image; super-resolution reconstruction; generative adversarial network; residual*

*dense network*

# **引言**

# 超分辨率（SR）是从给定的低分辨率（LR）图像中生成高分辨率（HR）图像的过程，该图像被用于各种应用程序，如监视。

# 这一任务的目标曾经是将生成的图像与原始图像之间的均方误差（MSE）的最小化。这导致了最大限度的峰值信号比（PSNR），这是SISR的标准测量。然而，面向PSNR的方法并不能产生感知良好的图像[2]。然后提出了面向感知的方法。超分辨率生成对抗网络（SRGAN）[2]同时使用感知损失[3,4]和生成对抗网络（GANs）[5]来生成位于自然图像流形中的图像。增强的超分辨率生成对抗网络（ESRGAN）[1]通过引入一种由残差-不残差密集块（RRDB）组成的架构，改进了SRGAN，而不需要批量归一化（BN）[6]层。此外，我们还采用了相对论平均GAN（RaGAN）[7]作为鉴别器，并在激活前使用了这些特征。我们的目标是进一步提高图像生成的感知质量。

# 相关工作

我们采用了深度神经网路才解决红外图像超分辨率重建的问题。早在2014年，董超等人就提出了利用SRCNN的方法从低分辨率到高分辨率之间学习映射，以端到端的方式来解决图像超分辨率重建的问题[]。在该领域取得了显著的性能提升。.后来，该领域出现了各种网络架构，如具有残差学习[5]的深层网络、拉普拉斯金字塔结构[6]、残差块[1]、递归学习[7,8]、密集连接网络[9]、深背投影[10]和残差密集网络[11]。.具体来说，Lim等[20]提出了EDSR模型，通过去除残差块中不必要的BN层，扩大模型尺寸，实现了显著的改进。Zhang等[11]提出在SR中使用有效的残差密集块，并进一步探索了具有通道注意[12]的更深层网络，实现了最先进的PSNR性能。除了监督学习外，还引入了[26]强化学习和无监督学习[27]等方法来解决一般的图像恢复问题。

已经提出了几种方法来稳定训练一个非常深的模型。例如，开发了残差路径来稳定训练和提高性能[18,5,12]。残差缩放首先由Szegedy等人[21]使用，也用于EDSR。对于一般的深度网络，He等人[28]提出了一种无BN的VGG型网络的鲁棒初始化方法。为了便于训练更深层次的网络，我们开发了一个紧凑而有效的残差密集块，这也有助于提高感知质量。

知觉驱动的方法也被提出来提高SR结果的视觉质量。基于更接近感知相似性[29,14]的思想，提出了感知损失[13]，通过最小化特征空间而不是像素空间中的误差来提高视觉质量。上下文损失[30]是被开发的，通过使用一个关注特征分布而不是仅仅比较外观的目标来生成具有自然图像统计的图像。Ledig等人[1]提出了SRGAN模型，该模型使用感知损失和对抗性损失来支持位于自然图像流形上的输出。Sajjadi等人开发了类似的方法，并进一步探讨了局部纹理匹配损失。基于这些工作，Wang等人[17]提出了空间特征变换，以有效地将语义先验合并到图像中，并改进恢复的纹理。

在整个文献中，照片现实主义通常是通过使用GAN [15]的对抗性训练来实现的。最近有很多工作专注于开发更有效的GAN框架。WGAN [31]提出了瓦瑟斯坦距离的近似，并通过权重裁剪正则化鉴别器。其他改进的正则化方法包括梯度裁剪[32]和光谱归一化[33]。相对论鉴别器[2]不仅可以增加生成数据真实的概率，而且同时降低真实数据真实的概率。在这项工作中，我们通过使用一个更有效的相对论平均GAN来增强SRGAN。

SR算法通常通过几种广泛使用的失真测量方法来评估，例如PSNR和SSIM。然而，这些指标从根本上不同意对人类观察者[1]的主观评价。非参考指标，包括Ma评分[23]和NIQE [24]，两者都用于计算PIRM-SR挑战[3]中的感知指数。在最近的一项研究中，Blau等人[22]发现，失真和感知质量是相互矛盾的。

**3 Proposed Methods**

我们的主要目的是提高对SR的整体感知质量。在本节中，我们首先描述我们提出的网络体系结构，然后讨论从鉴别器和感知损失的改进。最后，我们描述了平衡感知质量和PSNR的网络插值策略。

本文生成网络和判别网络如图 2 所示。其中红

色块表示残差密集网络，蓝色块表示渐进式上采样，

*k* 为卷积核尺寸，*n* 为卷积核数量，*s* 为步长。在生

成网络提取特征信息时采用残差密集网络，重建图

像时利用渐进式上采样的方法；而判别网络是利用

卷积、激活函数、全连接神经网络组合设计而成。

生成网络输入输出图像尺寸均为 128×128×3。

前半部分残差密集网络包含 1 个 Conv-Relu 的组

合、14 个 Conv-BN(Batch Normalization)-Relu 的

组合和 1 个 Conv-BN 的组合，每个组合中的 Conv

卷积核大小均为 7×7，步长为 1，16 个组合中卷积

核数量均为 64。后半部分是渐进式上采样、多重

卷积和 tanh 激活函数的组合，其中一共经历 7 个

Conv 卷积，卷积核大小均为 7×7，步长为 1，卷积

核数量分别 256，64，64，256，64，32，3。最后

经过 tanh 函数，生成高分辨率图像。

Figure 1: 我们采用了SRResNet [1]的基本架构，其中大多数计算是在LR特征空间中完成的。此外我们采用了全局特征块(CRDB) 提高了超分辨率重建的性能。

**3.1网络体系结构**

为了进一步提高SRGAN网络的超分辨率重建性能，我们主要对SRGAN的生成器网络进行了两个方面的修改： 1)去除所有BN层；2)替换为我们设计的全局特征块（CRDB），后者结合了多层次残余网络和密集连接（子这里写在这里写全局特征的那些东西），如图4所示。k 表示卷积核的尺寸，n 表示卷积核的数量，s 表示步长。整个残差密集网络的结构如图 3 所示。在训练过程中，我们通过反向传播算法对网络中的参数进行优化，从而让网络能够更好地适应数据集中的图像，提高图像超分辨率重建的准确性和效果。

我们保留了SRGAN的高级架构设计（见图3），并使用了一个新的基本块，即RRDB，如图4所示。基于观察到更多的层和连接总是可以提高性能[20,11,12]，所提出的RRDB采用了比SRGAN中的原始残差块更深、更复杂的结构。具体来说，如图4所示，所提出的RRDB具有残差中结构，其中残差学习在不同的水平上使用。在[36]中也提出了一种类似的网络结构，它也应用了一个多级剩余网络。然而，我们的RRDB与[36]的不同之处在于，我们在主路径上使用密集的数据块[34]作为[11]，其中网络容量从密集的连接中变得更高。

Figure 1: Example of a figure with caption. Captions are set in roman, 9 point. Use a Drawing area to make space for figures.

除了改进的体系结构，我们还利用了几种技术来促进训练一个非常深的网络： 1)残差缩放[21,20]，即在将一个常数添加到主路径之前，以防止不稳定；2)较小的初始化，因为我们根据经验发现，当初始参数方差变小时，残差结构更容易训练。更多的讨论可以在补充材料中找到。培训细节和拟议网络的有效性将在第2节中介绍。 4.

**3.2 残差密集网络**

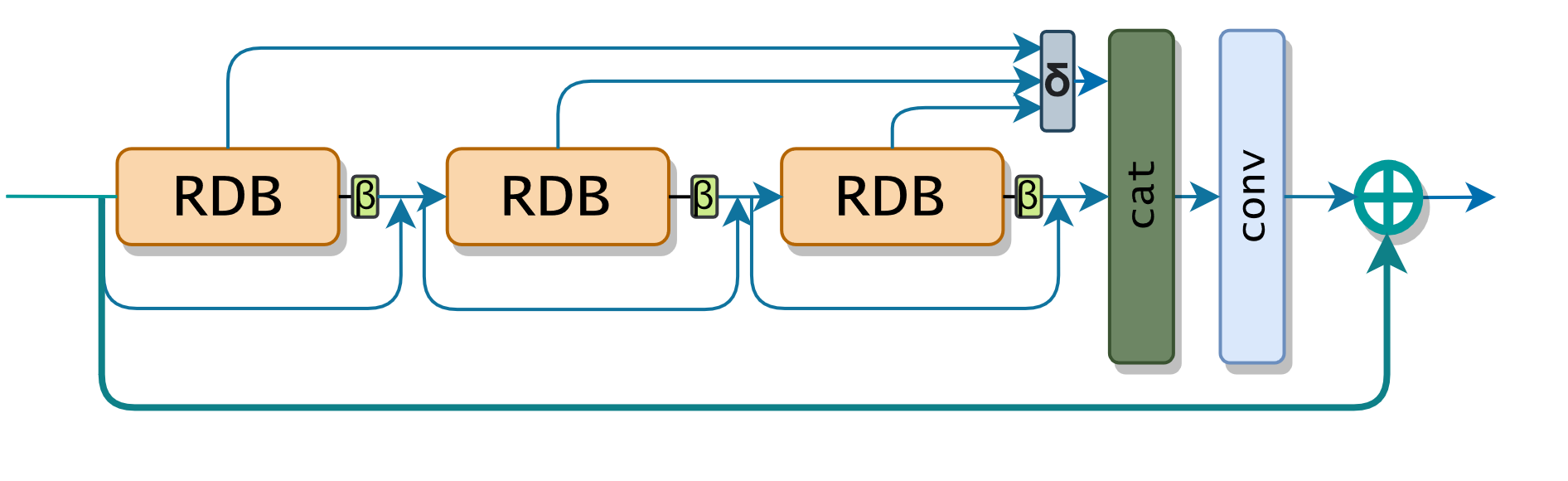
残差密集网络是一种结合了残差网络和密集网络的深度神经网络结构。它的优势在于能够同时提取全局特征和局部特征，通过将前一层的特征直接链接到当前层的所有层，形成连续的记忆机制，从而有效地解决了深度神经网络中的梯度消失问题。具体地说，残差网络负责提取全局特征，而密集卷积层则负责提取局部特征。这种结合方式能够让网络更好地学习到图像中的细节和特征信息。

每个残差密集块由 Conv-Relu 组合而成，其中 Conv 是卷积操作，Relu 是激活函数。我们去除了BN层，因为在[]中的工作和实验已经证明在图像超分辨率重建任务中，去除BN层可以提升系统的性能并且有效地减少计算复杂度。

**3.3 全局特征提取块**

为了使我们的IRSRGAN能够学习到不同层次的特征，我们设计了一种全局深度残差块(Comprehensive Residual Dense Block, CRDB)。其结构如图2所示。我们采用了RDB作为CRDB的构建基础。我们采用了CRDB块作为我们生成网络的基本块。CRDB增加了全局残差连接，这样网络就可以专注于学习不同维度的低分辨率图像和地面真实图像之间的差异，我们将在第四节详细的对比，采用这种技术使IRSRGAN在图像超分辨率重建任务上获得的增益。

图2:CRDB网络结构，其中为残差因子，为特征因子



**3.4相对论判别器**

除了改进了生成器的结构外，我们还增强了基于相对论性GAN [2]的鉴别器。与SRGAN中的标准鉴别器D不同，它估计一个输入图像x是真实和自然的概率，相对论鉴别器试图预测一个真实图像相对于一个真实的一个更真实的概率。标准的判别器可以被表示为，是sigmod函数，代表未变换的判别器输出值。那么采用了相对判别器的GAN网络(**Relativistic GAN)**可用如下公式所表示：

在公式(1)和(2)中，和分别代表真实数据和虚假数据，表示对数据取期望的操作。

生成器的损失函数和判别器的损失函数表示为如下公式：

其中，和代表输入LR图像。可以观察到，生成器的对抗性损失同时包含和。因此，我们的生成器受益于对抗性训练中生成的数据和真实数据，而在SRGAN中，只有生成的数据生效。在第四部分，我们将展示这种鉴别器的修改有助于学习更清晰的边缘和更详细的纹理。

**3.5随机退化**

## 在提供我们新的实用的SISR降解模型之前，提到以下关于双边缘和传统降解模型的事实是有用的：1。根据传统的退化模型，影响真实图像退化的三个关键因素是模糊、降采样和噪声等关键因素。 2.由于LR和HR图像都可能是有噪声和模糊的，因此不需要像传统的退化模型那样采用模糊/降采样/噪声添加管道来生成LR图像。 3.传统退化模型的模糊核空间应该因不同的尺度而变化，这使得在实践中很难确定非常大的尺度因素。 4.虽然双边缘退化很少适用于真实的LR图像，但它可以用于数据增强，确实是清晰、清晰的图像超分辨率的良好选择。受第一个事实的启发，提高退化模型实用性的一个直接方法是使这三个关键因素的退化空间尽可能大和现实。基于第二个事实，我们对三个关键因素采用随机洗牌策略，进一步扩展了退化空间。这样，LR图像也可以是HR图像的噪声、降采样和模糊版本。为了解决第三个事实，人们可以利用小因子对大尺度因子的分析计算。或者，根据第四个事实，对于一个大尺度因子，可以在尺度因子2退化之前应用双边（或双线性）降尺度。本文在不丧失一般性的前提下，设计了广泛使用的尺度因子2和4的退化模型。下面，我们将详细介绍以下几个方面的退化模型：模糊、降采样、噪声和随机洗牌策略。

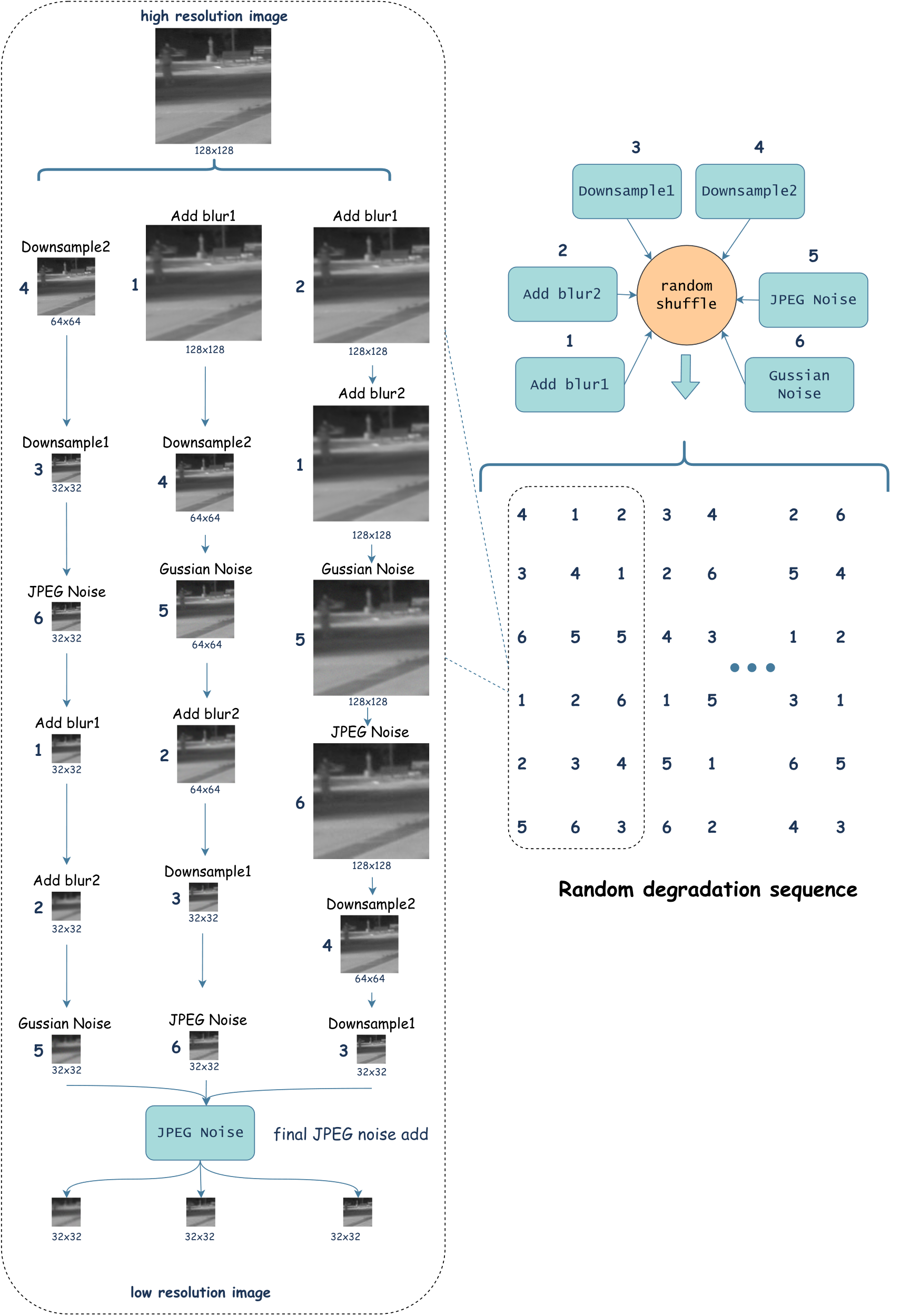


图 3: 提出的尺寸因子为4时的随机退化模型的示意图。

## 3.5.1.模糊噪声

模糊是一种常见的图像退化，会影响到各种应用中的图像质量。在本文中，我们提出了一种新的方法来建模模糊。

一方面，在传统的超分辨率图像恢复（SISR）退化模型[28,45]中，首先将高分辨率（HR）图像通过与模糊核的卷积进行模糊处理。在传统的SISR任务中，由于HR图像与LR图像的尺度差异较大，将HR图像直接下采样到LR尺度会导致混叠的问题，即多个HR像素映射到同一个LR像素的情况。这会导致信息的丢失和图像质量的降低。为了解决这个问题，通常采用模糊处理的方式，将HR图像与一个合适的模糊核进行卷积处理，使得相邻像素之间的信息被模糊化，减少混叠现象的出现。

在本文中，我们采用了HR模糊来防止混叠的问题，并保留更多的空间信息。具体来说，我们将HR图像与一个高斯核进行卷积处理，使得HR图像在进行下采样操作后，映射到LR图像时可以保留更多的细节信息，从而更好地进行超分辨率恢复任务。同时，我们也考虑到真实的LR图像可能已经存在模糊，因此在LR空间中建模这种模糊也是一种可行的方法。因此，我们从HR和LR两个角度对模糊进行了建模，从而更好地解决了混叠问题，并提高了图像质量。

为了进一步考虑高斯核在SISR任务中的应用，我们采用了3x3大小的各向同性的高斯核来生成模糊的图像。同时，在退化的过程中经过了两次模糊操作，大大地扩展了退化空间。这种退化模型不仅能够很好地模拟实际中的模糊，还能够提高图像的复杂度，更贴近实际的情况。本文提出了一种新的方法来建模模糊，从HR空间和LR空间两个角度对模糊进行了建模。采用了3x3大小的各向同性的高斯核来生成模糊的图像，并通过两次模糊操作来扩展退化空间。这种退化模型能够更好地模拟实际中的模糊，并提高图像的复杂度，更贴近实际情况。这种方法在SISR任务中有着广泛的应用前景。

## 3.5.2 高斯噪声

噪声在真实图像中普遍存在，因为它可能由不同的来源引起。我们的退化模型考虑到了现实世界中最广泛地高斯噪声。当无法知道现实世界中噪声的具体成分时，采用高斯噪声作为假设噪声是最稳妥的选择。我们采用了均值为0.01，标准差为0.01的高斯噪声。均值选择0.01，避免了因为均值过大或者过小导致叠加噪声后的图片过亮或者过暗。而标准差选择0.01是为了将噪声的强度控制在较小的水平，避免因为噪声强度和范围过强，导致图像的信息分量被噪声分量掩盖。本文添加高斯噪声的过程首先将输入的图片标准化，并在产生一个和图片大小相同尺寸的高斯噪声。将二者叠加，生成带有高斯噪声的图片。

## 3.5.3 JPEG噪声

JPEG是一种广泛应用的图像压缩标准，它可以降低带宽和存储空间的使用。然而，在高度压缩的情况下，它会引入8x8像素的阻塞伪影和噪声，这可能会给图像带来明显的质量损失。JPEG的压缩程度由质量因子决定，该因子的值在0到100之间，数值越高，质量越好，压缩比越小。

在我们的新的退化模型中，我们将JPEG质量因子的范围限制在70到95之间，并在该范围内随机选择一个值。这意味着我们会在不同的质量级别下进行JPEG压缩，从而生成不同程度的图像噪声和伪影。当质量因子小于90时，这种噪声和伪影可能会变得更加明显。然而，当质量因子大于90时，这种噪声和伪影就不太明显了。

为了更好地模拟真实的图像退化情况，我们采用了两个JPEG压缩步骤，分别可能性为0.75和1。特别是，我们将后者用作最终的降解步骤，以增加图像的失真和噪声，从而更好地模拟真实场景中的退化情况。这种方法可以使我们的模型更加真实地模拟真实图像，并为图像处理任务提供更具挑战性的数据集。

## 3.5.4 下采样

本文探讨的是超分辨率重建任务中的一种新颖方法，与一般方法不同的是，本文采用了下采样插值的随机选择方法。这种方法的创新之处在于，每次执行下采样时会随机选择双线性插值、双三次插值或区域插值中的一种方法。这种方法可以丰富退化样本的多样性，从而提高了重建的准确性。在本研究中，我们对三种插值方法进行了详细的研究和比较，包括双线性插值、双三次插值和最近邻插值。双线性插值是一种简单但有效的方法，它将每个像素的值设置为其周围四个像素的平均值。双三次插值则是一种更复杂的方法，它通过拟合一个二次函数来计算像素值。区域插值使用像素区域关系进行重采样，即将源图像中的每个像素映射到目标图像中最接近它的像素。该插值方法可以减少图像中的高频噪声，并且在保持图像细节方面表现良好。

为了使下采样插值方法更具有多样性，我们采用了随机选择的方法。在每次下采样时，从三种插值方法中随机选择一种。这样可以使得退化样本更加多样化，从而提高重建的准确性。

## 3.5.5 随机洗牌退化策略

传统的图像超分辨率技术使用简单的降采样和插值操作来产生低分辨率(LR)图像，然后使用插值或卷积核恢复高分辨率(HR)图像。虽然这些方法在有限的情况下能够取得一定的效果，但是当此类超分辨率重建模型面对真实世界中的模糊图像时，它们无法捕捉真实世界中不同因素对图像退化的影响，比如噪声、模糊、下采样、JPEG压缩等。

因此，为了更好地模拟真实世界中图像的退化情况，研究人员提出了各种不同的退化模型。例如，一些模型通过添加高斯噪声、运动模糊等方式模拟图像的退化。其他模型则采用混合降采样和模糊的方法来模拟退化过程。这些模型可以更准确地模拟真实世界中图像的退化情况，从而提高超分辨率重建的性能和效果。

为了进一步扩展这些模型的覆盖范围，本文使用了一种随机洗牌策略作为退化模型，来生成低分辨率的图像。

需要注意的是，这种随机洗牌策略并不是新的。事实上，在一些先前的研究中，类似的随机洗牌策略已经被提出用于图像分类和目标检测中，以增强高分辨率图像。但是，这种策略在图像超分辨率领域中的应用尚未得到广泛的研究和探索。通过将这种策略应用于图像超分辨率任务中，研究人员可以更加准确地模拟真实世界中图像的退化情况，并提高图像超分辨率的性能和效果。

通过采用随机洗牌策略，可以极大地扩展退化模型的应用范围。传统的退化模型虽然简单、方便，但很难完全覆盖真实低分辨率（LR）图像的退化情况。实际上，真实的LR图像可能是HR图像的含噪、模糊、下采样和JPEG压缩等版本。因此，我们提出了一种新的退化模型，可以通过模糊、降采样和不同阶段的噪声等操作来生成不同程度的LR图像退化。具体地，我们采用了一个随机洗牌序列，其中Ds表示尺度因子为2的降采样操作，在随机退化过程中总共会执行两次，每一次都会随机在双线性插值，双三次插值，区域插值中选择一种插值方法。值得注意的是，该降采样操作并非一定按照顺序连续执行，他们可以与其他的噪声或模糊操作组合进行，这可以显著扩展了图像的退化空间。

采用这种随机洗牌策略，我们不仅可以扩展现有的退化模型，还可以生成一系列全新的退化模型。通过对模糊和降采样操作进行不同排列，我们可以放大模糊退化空间，同时改变噪声特性，从而进一步扩大退化空间。例如，降采样可以降低噪声强度，使噪声（如处理后的高斯噪声和JPEG压缩噪声）不依赖于信号本身，也会受到降采样的操作的影响；这种噪声在真实图像中可能存在。

图1展示了我们提出的退化模型，我们可以通过改变不同的退化参数和操作，生成具有广泛退化范围的不同LR图像。在数据预处理的过程中，我们设置加载随机退化的过程的概率为0.25，而进行普通的双线性降采样的概率为0.75。我们提出的新的退化模型可以更好地反映真实世界中的图像退化情况，并可以为超分辨率重建提供更多的退化样本。总之，在面对具有复杂噪声的低分辨率图像时，我们引入随机退化操作的方法可以使图像超分辨率重建模型有更好的性能和表现。

**3 实验结果及分析**

## **3.1 数据集和质量评价指标**

本文的研究集中在解决红外图像超分辨率重建问题，并采用了FIR公司提供的开源红外图像数据集CVC-09和CVC-14。该数据集包含白天和黑夜的图像，覆盖了多种场景类型，如街道、行人、汽车、交通指示和房屋建筑等。这些图像均为单通道图像。为了增加训练数据集的多样性，本文在CVC-09和CVC-14数据集中随机选择了15,000张照片，并对每张图片进行了数据增强处理。首先将这些照片随机裁剪成大小为128 x 128的高分辨率真实图像，然后进行下采样，生成大小为32 x 32的低分辨率图像作为模型的训练数据集。设置的下采样因子为4。在测试时，本文还使用了未在CVC-09和CVC-14中出现的红外图像数据集的图像。这些图像进行相同的下采样处理后，得到的低分辨率图像作为测试数据集送入模型进行测试，生成高分辨率图像。

在超分辨率重建算法中，为了评估图像重建的质量，我们通常会采用主观评价和客观评价两种方法。主观评价是指通过实验人员的主观感受来评价图像的质量，这种方法需要进行一系列的受试者的实验，然后根据实验结果进行统计分析得出。而客观评价是通过设计量化指标来衡量图像质量的好坏，由于不需要受试者实验，因此更加的简单快捷。在图像质量评价的领域，客观评价的指标种类有很多，例如PSNR、SSIM、VIF、MS-SSIM等。

在本文中，我们选取了两种常用的客观评价指标： PSNR和SSIM, 来对我们提出的方法进行评估。PSNR是一种基于均方误差的指标，通过计算重建图像与原图像之间的差异来评估图像质量，其数值越高表示重建图像与原图像之间的相似度越高。而SSIM则是一种基于结构相似性的指标，其考虑了图像的亮度、对比度和结构等因素，可以更好地反映人眼对图像质量的感知。

同时，我们还对比了本文所提出的方法与一些经典和主流的方法，如Bic[]、SRCNN[]、ESPCNN[]、EDSR[]、SRGAN[]和ESRGAN[]，进行了实验评估。这些方法在超分辨率重建领域都有一定的影响力，可以作为我们的对比对象。通过实验结果的比较，我们可以更加全面地评估本文提出的方法的优劣，为超分辨率重建算法的研究和应用提供有益的参考。

Figure 1: Example of a figure with caption. Captions are set in roman, 9 point. Use a Drawing area to make space for figures.

## **3.2 训练细节和参数**

本文的IRSRGAN红外图像超分辨率重建算法在单张NVIDIA Tesla V100-16GB 显卡上实现训练。在训练生成网络时，对图像进行归一化处理，将输入图像的像素取值范围映射到[-1, 1]。对于每一批次，本文都为了训练一个高质量的超分辨率（SR）图像生成器，我们从一个包含 32 张高分辨率图像的数据集中随机选择输入图像。为了最小化生成的 SR 图像与真实高分辨率图像之间的像素损失，我们使用 L1 loss 进行训练。值得注意的是，L1 loss 可以很好地度量两个图像之间的平均差异，并且与 PSNR 之间存在一定的相关性。

我们使用 Adam 优化器进行优化，其中包括 =0.9 和 =0.99 的参数衰减率，以及学习率为 learning\_rate=0.0003。这些参数的含义和作用如下：beta1 和 beta2 分别控制历史梯度的衰减率，使得当前梯度更能够影响更新过程；学习率则控制每次迭代更新的步长大小。在我们的训练过程中，我们设置 epoch=150，以确保生成器可以充分学习数据集中的特征。

最后，我们使用 kaiming\_normal\_ 方法对模型的权重进行初始化。kaiming\_normal\_ 是一种常用的权重初始化方法，特别适用于使用 ReLU 激活函数的神经网络。在使用该方法时，权重的初始值会按照从均值为 0、标准差为 sqrt(2 / fan\_in) 的正态分布中采样生成，其中 fan\_in 表示当前层输入数据的维度。通过这种初始化方法，可以使得每个神经元的输出具有相同的方差，从而更好地保持信号的稳定性。

与传统的 Xavier 初始化方法不同，kaiming\_normal\_ 方法考虑到了 ReLU 激活函数的非线性特性。由于 ReLU 函数在输入为正时具有线性变换的性质，因此通过 kaiming\_normal\_ 初始化方法初始化的权重可以更好地适应 ReLU 激活函数，并使得神经网络的训练更加稳定和有效。

需要注意的是，kaiming\_normal\_ 方法仅适用于使用 ReLU 激活函数的神经网络，对于其他类型的激活函数可能不适用。此外，在使用该方法时，还需要注意权重的标准差不宜过大，以避免梯度爆炸的问题。

Figure 1: 训练时随着epoch变化的loss曲线

## **3.3 量化结果**

为了测试我们的红外图像超分辨率重建模型在实际应用中的表现，我们使用了TISR数据集进行评估。TISR数据集包含了多个真实场景下的低分辨率红外图像及其高分辨率版本，是一个常用的用于测试红外图像超分辨率算法的数据集。我们在TISR数据集上对我们的模型进行了测试，并使用了几个常见的评价指标来评估模型的性能，包括峰值信噪比（PSNR）、结构相似性指数（SSIM）等。通过实验结果，我们证明了我们的模型在TISR数据集上达到了优秀的性能表现。

001

Figure 2: Short captions should be centered.

本文方法及其他方法对测试图像集的评价情况如表 2 所示，从表 2 中

可以得出，本文所提出的 SRRDB 模型在 CVC-09

和 CVC-14 数据集测试所得的客观评价指标 PSNR

和 SSIM 值都要高于其他几种算法。

本文从测试图像集中选取包含道路、行人、车

辆的典型图像进一步说明本文所提方法的有效性，

实验结果如图 8~12 和表 3 所示。从实验结果目视

分析，传统的Bic方法所重建的图像整体比较模糊，

ESPCNN 方法所重建的图像存在噪声点多、视觉效果不佳等问题，而 SRCNN 重建后的图像在明亮程

度方面与原始高分辨率图像相差较大，且图像稍显

平滑。EDSR 方法在红外图像重建时出现很多矩阵

块，原因可能是本文所使用的红外图像是单通道

的，输入网络的图像 3 通道的值一样，此时的 EDSR

方法并不适用在单通道红外图像超分辨率重建任

务上。SRGAN 在重建质量上优于以上方法，但是

在细节纹理重建上还是略显不佳。本文 SRRDB 方

法实验结果优于其他方法主要原因有：①采用的红

外图像数据集是 CVC-09 和 CVC-14，所选数据集

为单通道、背景不复杂、图像内容均为街景；②在

生成对抗网络的基础上引入残差密集网络，残差密

集网络可以充分利用原始低分辨率图像的所有分

层特征，保证低层信息传到高层，相对其他方法而

言，对低分辨率图像信息利用更为充分，提高了信

息的利用率。在 RGB 图像上的实验结果表明，本

文方法在红外图像上的性能优于 RGB 图像，在

RGB 图像超分辨率重建上存在一定的局限性。

# **Formatting your paper**

All text must be

## **Margins and page numbering**

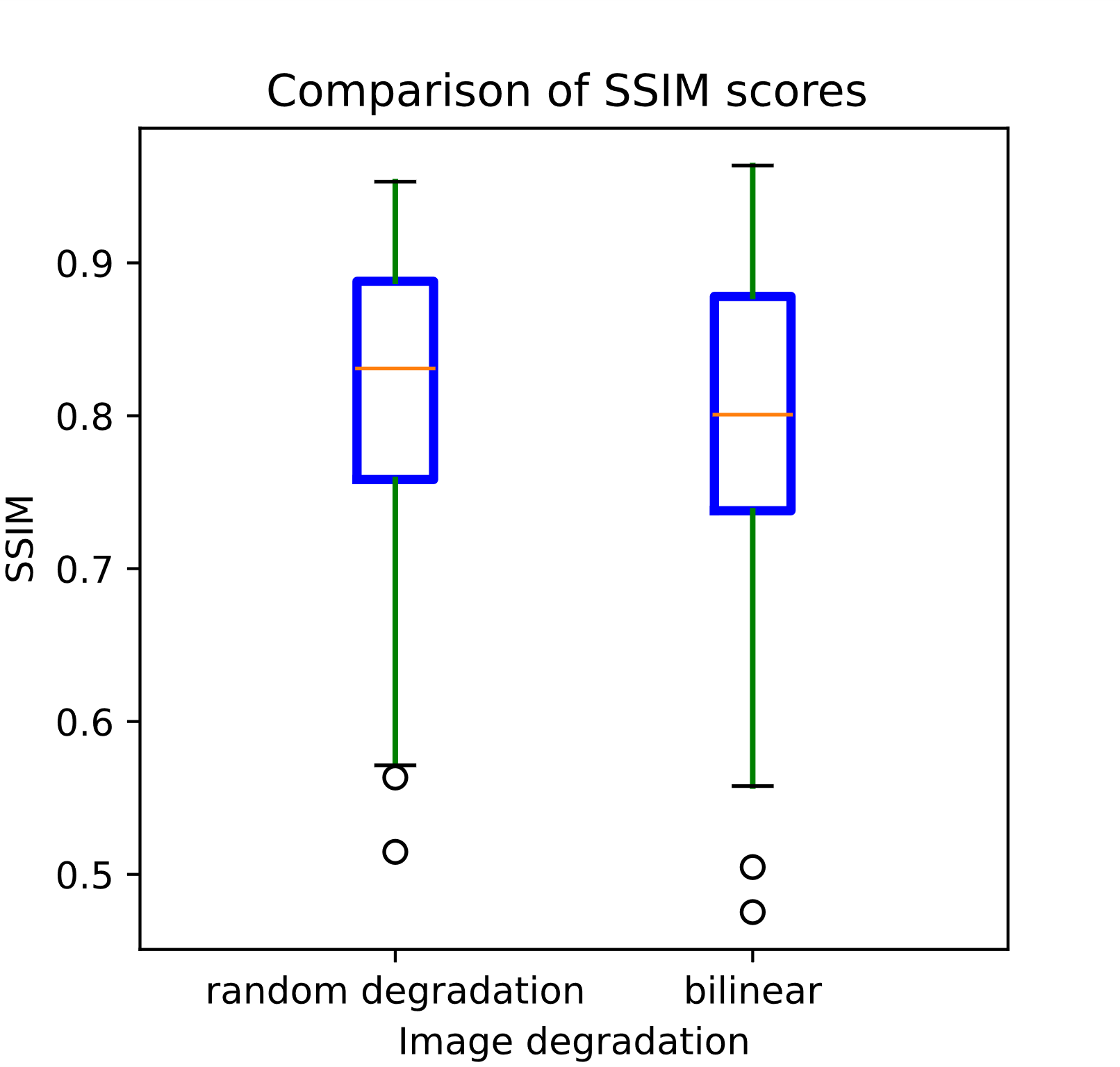
**4.4消融研究**

**要有一个大图片，表格**

随机退化对超分辨率重建的影响

本研究对比了在不加随机退化和加入随机退化两种情况下的图像超分辨率重建算法的表现。我们使用了TISR数据集，在其中随机抽取了100张图片并预处理作为低分辨率图像数据，并使用PSNR和SSIM两个指标来评价算法的表现。在数据集预处理方面，我们分别在训练时对训练集进行了不加随机退化和加入随机退化两种处理策略，然后使用训练出来的两种模型对TISR数据集进行了图像超分辨率重建。在评价指标计算方面，我们计算每个图像的PSNR和SSIM分数。

我们使用了箱线图来展示两种算法在100张图像上的PSNR和SSIM分数的分布情况。如图1所示，我们可以发现在使用训练时加入随机退化的模型的情况下，该模型比未在训练过程中加入随机退化的模型取得了更高的平均PSNR和SSIM分数，这说明加入随机退化可以提高超分辨率重建的表现，使得算法在面对具有随机模糊的红外图像时的表现更加稳定和出色。



我们还选择了一些典型的图像，对两种算法的效果进行了对比。

如图2所示，我们可以看到，两种模型的图像重建结果与原图相比都存在一定的失真，但是添加了随机退化的算法在对比度和清晰度上相比另一种算法更加优秀。其超分辨率重建出来的图像有更丰富的细节和纹理。

综上所述，本研究发现，在加入随机退化的情况下，基于SRGAN算法的图像超分辨率重建模型在面对模糊图像的重建方面的表现相对更好。

**2不采用预训练模型和采用预训练模型的对比（对比PSNR）**

在本研究中，我们对比了采用预训练生成器模型的GAN和未使用生成器模型的GAN在训练过程中的PSNR和SSIM值的变化。实验结果表明，采用预训练生成器模型的GAN在训练过程中的PSNR和SSIM值普遍高于未使用生成器模型的GAN。具体来说，预训练生成器模型的GAN在早期训练轮次中可能会表现稍差，但随着训练的进行，其PSNR和SSIM值逐渐上升并保持在较高水平，而未使用生成器模型的GAN在整个训练过程中表现相对较低。这表明预训练生成器模型能够为GAN提供更加有效的初始化，使得GAN在超分辨率重建任务中更容易学习到高质量的特征。

# **3全局特征引入**

在本研究中，我们进行黑白图像超分辨率重建的改进工作。在损失函数中，我们使用了vgg19的预训练网络模型。然而，我们发现，使用彩色图像训练的VGG预训练模型可能会对黑白图像重建任务带来一定的干扰，导致重建图像中出现绿色条纹伪影，因为彩色图像和黑白图像的特征分布不同，如果使用彩色图像训练的VGG网络，可能会将黑白图像中的一些纹理特征误认为是彩色信息，从而影响模型的性能和稳定性。。为了避免这个问题，我们对模型进行了一系列修改和优化，采用了全局特征引入的方法来规避使用彩色图像训练的VGG预训练模型所带来的干扰。

具体来说，我们使用了全局特征提取块来构建超分辨率图像重建网络，该技术可以有效地适应黑白图像重建任务的特点。这种方法不仅可以避免使用彩色图像训练的VGG预训练模型所带来的干扰问题，还可以充分利用黑白图像的特点来提高模型性能和稳定性。实验结果如图4所示，该实验结果表明，经过我们的改进和优化，新模型在黑白图像超分辨率重建任务中表现出了很好的性能，且重建图像中没有出现绿色条纹伪影，这表明我们所采取的改进方法是有效的。

。

## **5.结论**

通过本文的研究，我们成功提出了一种全新的红外图像超分辨率重建网络IRSRGAN，并在单帧红外图像超分辨率重建任务上取得了显著效果。与非 GAN 网络相比，本文所提出的方法在低分辨率重建到高分辨率图像过程中获得了更多的高频信息，使得放大因子为 4 的图像清晰且边缘尖锐。另外，我们也提出了一种随机退化模型，可以使得超分辨率重建网络在面对复杂的噪声环境下有着出色的性能表现。

然而，由于公开的高清红外图像数据集相对较少，可供训练的数据也不够丰富，且为单通道图像，制约了红外图像超分辨率质量的进一步提高。这也导致在本文中无法有效引入通道注意等机制，无法从多个通道关联中获取更多的信息。另外，由于经费限制，我们只能进行较少轮次的模型学习，使得模型性能并没有达到最优。

我们认为，在未来的研究中，可以进一步引入更多的先验知识，例如网络结构、上采样的方式、注意力机制以及传统的图像处理方法等，来进一步提高红外图像超分辨率重建的能力。同时，通过更多的实验和数据的积累，也可以探究多通道红外图像超分辨率重建的方法，从而进一步提高红外图像的质量和应用范围。

# **References**

1. FirstName Alpher, Frobnication. *IEEE TPAMI*, 12(1):234– 778, 2002.
2. FirstName Alpher and FirstName Fotheringham-Smythe. Frobnication revisited. *Journal of Foo*, 13(1):234–778, 2003.
3. FirstName Alpher, FirstName Fotheringham-Smythe, and FirstName Gamow. Can a machine frobnicate? *Journal of Foo*, 14(1):234–778, 2004.
4. FirstName Alpher and FirstName Gamow. Can a computer frobnicate? In *CVPR*, pages 234–778, 2005.
5. FirstName LastName. The frobnicatable foo filter, 2014. Face and Gesture submission ID 324. Supplied as supplemental material fg324.pdf.
6. FirstName LastName. Frobnication tutorial, 2014. Supplied as supplemental material tr.pdf.