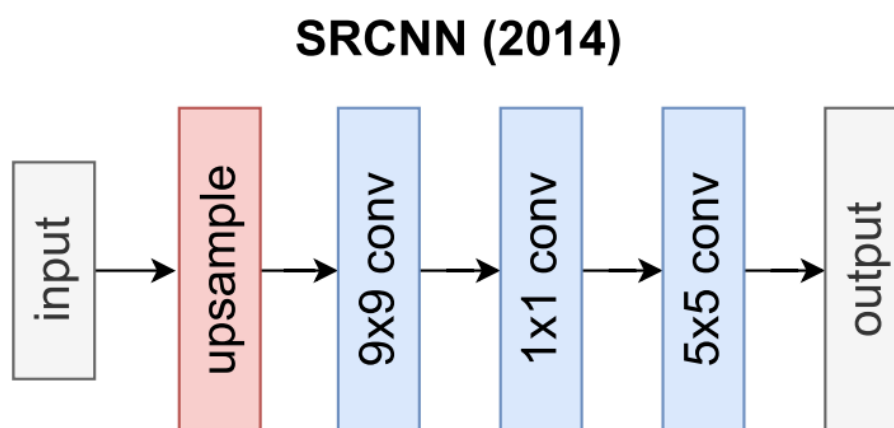


基本要求：

- 训练集T91-train,测试集Set5-test，不能更改。
- 需要分析网络架构不同所引起的性能变化，并在提交报告中对比分析。（可以是模型不同，也可以只是层数不同）
- 报告内容必须包含数据处理部分、模型部分、训练部分和测试部分。
- 超分辨任务b可以只选择放大倍数为4倍（纵横各4倍，图像大16倍）。
- 老师使用的模型：
 - SRCNN:

- 结构：

-



- 损失：nn.MSELoss
- 优化器：optim.Adam
- 训练批次：200轮
- 评价指标：Peak Signal-to-Noise Ratio (PSNR)

- ```
1 def calc_psnr(img1, img2):
2 return 10. * torch.log10(1. / torch.mean((img1 - img2)
 ** 2))
```

- 可以比较的模型：（完成意味着代码写完了，可以跑起来，但是效果需要针对训练集进行微调，延续老师代码的损失、优化器和学习率）
  - 初始模型SRCNN
    - 输入为16\*1\*76\*76
    - 输出为16\*1\*76\*76
    - 预处理不同，其对数据处理的规则如下
      - 将图片降分辨率：降低四倍
      - 再将图片resize原图片同样大小
      - 组成同样大小的降分辨率图像和原图像
      - 随机裁剪16个76\*76大小的方块
  - 来源于ResNet的模型SRResNet
    - 输入为16\*3\*19\*19

- 输出为 $16*3*76*76$
- 预处理不同，其对数据处理的规则如下
  - 对图片进行随机裁剪16个 $76*76$ 大小的方块
  - 将图片降分辨率：降低四倍变为 $16*3*19*19$ 的降分辨率图像
- best epoch: 166, psnr: 24.29
- 来源于DenseNet的模型SRDenseNet
- 基于递归的方法
  - DRCN
  - DRRN
  - CARN
  - VDSR(完成，但是在第三个epoch，损失就不再降低了，效果也不如SRCNN，所以需要针对训练集进行微调)
    - 输入为 $16*1*76*76$
    - 输出为 $16*1*76*76$
    - 预处理与SRCNN相同
- 基于像素的方法：
- 生成式对抗性网络（GANs）：
  - SRGAN
  - ESRGAN
- 基于光流的方法
- 可以使用的对比评价指标
  - 峰值信噪比（PSNR）：峰值信噪比（PSNR）是评价SISR重建质量最广泛使用的技术之一。它表示SR图像 $\hat{y}$ 与实际图像 $y$ 之间的最大像素值 $L$ 与均方误差（MSE）之比。
  - 结构相似度指数（SSIM）：SSIM和PSNR一样，是一种流行的评价方法，侧重于图像之间结构特征的差异。它通过比较亮度、对比度和结构来独立地捕获结构上的相似性。SSIM估计一个图像 $y$ 的亮度 $\mu_y$ 为强度的平均值，而它估计对比度 $\sigma_y$ 为其标准差。
  - 平均意见评分（MOS）：MOS是一种主观的测量方法，利用人类的感知质量来评估生成的SR图像。人类观众会看到SR图像，并要求他们进行质量评分，然后映射到数值，然后取平均值。通常，这些范围从1（坏）到5（好），但可能有不同的[15]。虽然这种方法是对人类感知的直接评估，但与客观指标相比，进行它更耗时和麻烦。此外，由于这个度量标准的高度主观性，它很容易受到偏见的影响。

后面是学习一些综述

## 超分辨率评价指标：

### 图像质量评估（IQA）

许多特性与优秀的图像质量有关，如锐度，对比度，或没有噪声。因此，对SR模型的公平评价具有挑战性。本节展示了属于图像质量评估（IQA）范畴的不同评估方法。广义上说，IQA指的是任何基于对人类观众的感知评估的度量，即应用SR方法后图像的真实程度。IQA可以是主观的（例如，人类评分者）或客观的（例如，正式的指标）。

1)平均意见得分 (MOS, Mean Opinion Score) : 数字图像最终是为人类观看的。因此, 评估图像的最合适的方法是主观评价[12], [13]。一种常用的主观IQA方法是平均意见评分 (MOS) 。人类观众给有质量分数的图像打分, 通常是1 (差) 到5 (好) 。MOS是所有评分的算术平均值。尽管具有可靠性, 但调动人力资源是耗时和麻烦的, 特别是对于大型数据集。

2)峰值信噪比 (PSNR, Peak Signal-to-Noise Ratio) : 由于近年来产生的大量图像和主观测量的弱点, 客观评估质量具有无可争辩的重要性。一种流行的目标质量测量方法是峰值信噪比 (PSNR) 。它是可能的最大像素值 $L$  (8位表示为255) 与参考图像的均方误差 (MSE) 之间的比率。给定近似值  $\hat{y}$  和地面真实值 $y$ , PSNR是一个使用分贝尺度[dB]的对数量:

$$\text{PSNR}(\mathbf{y}, \hat{\mathbf{y}}) = 10 \cdot \log_{10} \frac{L^2}{\frac{1}{N_y} \sum_{p \in \Omega_y} [\mathbf{y}_p - \hat{\mathbf{y}}_p]^2} \quad (1)$$

虽然它被广泛用作SR模型的评价标准, 但在真实场景中往往导致平庸的结果。它关注像素水平的差异, 而不是哺乳动物的视觉感知, 后者更吸引结构[14]。随后, 它与主观感知质量的相关性较差。像素的轻微变化 (例如, 移动) 可能会导致一个显著的PSNR降低, 而人类几乎不知道这种差异。因此, 新的指标关注于图像中更多的结构性特征。

## SRnet

---

### 简单的网络

---

简单的网络是一种主要应用卷积链的体系结构。它们很容易理解, 并且由于它们的大小, 通常只使用最少的计算资源。大多数这些体系结构都可以在基于dl的SR的早期找到, 因为它们的性能低于最先进的水平。此外, DL的“越深越好”的范式并不能很好地适用于简单的网络, 因为正在消失/爆炸的梯度[98]。

图6显示了不同的简单网络设计。

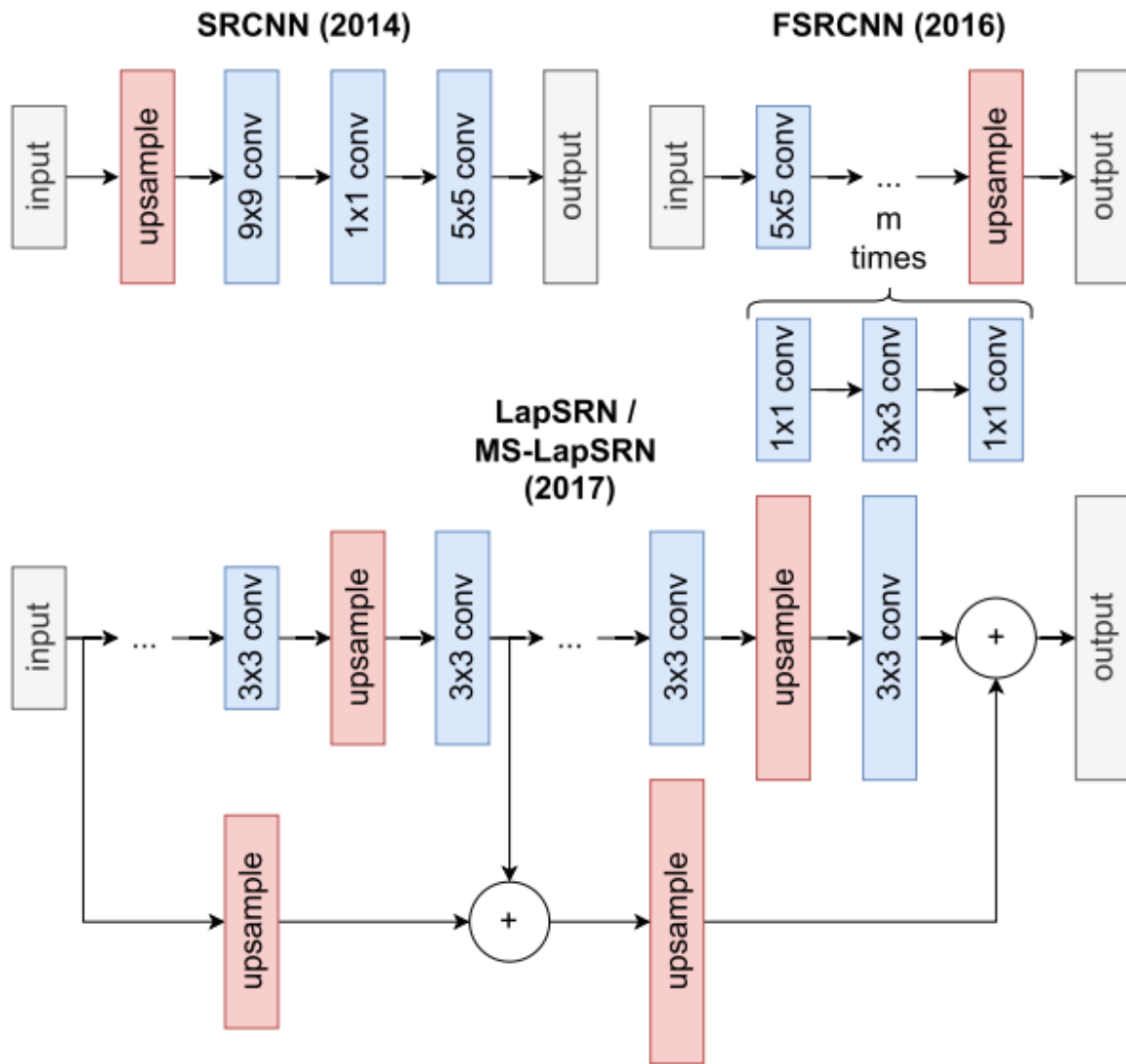


Fig. 6. Architecture designs of SRCNN [39], FSRCNN [33], ESPCN [66] and LapSRN/MS-LapSRN [44]. They are grouped together in this work as simple network designs, because they use only convolutional operations and upsampling methods.

第一个引入SR数据集的CNN是由Dong等人[39]提出的SRCNN（2014）。它使用双边缘预上采样来匹配地面真实空间大小（见第7.1节）。随后，它由三个卷积层组成，这遵循了图像恢复中流行的策略：补丁提取、非线性映射和重建。SRCNN的作者声称，应用更多的图层会损害性能，这与DL范式“越深越好”的[99]相矛盾。

如下所示，这个观察结果是错误的，需要更高级的构建块才能正确工作，例如，像VDSR [98]中那样的残差连接。

在他们的后续论文中，作者探索了各种加速SRCNN的方法，导致FSRCNN（2016）[33]利用了三个主要技巧：

首先，他们减少了卷积层内核的大小。

其次，他们使用了一个1x1的卷积层来增强和减少在使用3x3卷积的特征处理之前和之后的通道维度。

第三，他们采用了带有换位卷积的后上采样，这是提高速度的主要原因（见第7.1节）。

令人惊讶的是，它们在获得更快的同时优于SRCNN。

一年后，LapSRN (2017) [44]被提出，其关键贡献是一个拉普拉斯金字塔结构[100]，可以实现逐步上采样（见第7.1节）。它以粗分辨率的特征图作为输入，并预测高频残差，逐步细化每个金字塔层的SR重建。为此，在一个前馈通道中预测多尺度图像是可行的，从而促进了资源感知的应用程序。

简单的网络架构设计主要出现在基于dl的SR的早期，因为由于它们的大小，它们学习复杂结构的能力有限。最近，研究人员关注的是更有深度的网络，无论是残差网络，还是基于循环网络的合成深度。下面的部分将介绍这两种可能性。

## 残差网络

残差的网络使用跳过连接来跳过图层。增加跳过连接的主要原因有两个：为了避免梯度的消失和降低精度饱和问题[99]。对于SR来说，引入跳过连接开启了深度构建模型的世界。其主要优点是深度架构用大的接受域替代卷积，这对于捕获重要特征至关重要。SRCNN的作者指出，“越深越好”的范式并不支持SR。相比之下，Kim等人用VDSR (2015) [98]驳斥了这一说法，并表明非常深的网络可以显著改善SR，如图7所示。

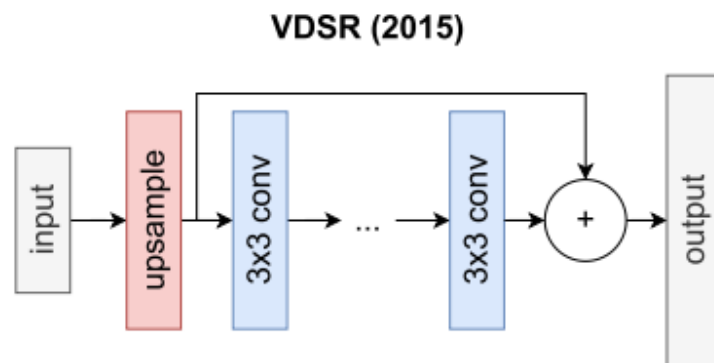


Fig. 7. Architecture design of VDSR [98]. It applies a deep convolutional network as feature extractor but also uses a residual from the input to the final prediction. Therefore, the feature extractor can concentrate to add high-frequency details onto the interpolated image.

他们使用了来自其他DL方法的两种见解：首先，他们应用了一个著名的架构VGG-19 [24]作为特征提取块。其次，他们使用了从插值层到最后一层的剩余连接。因此，VGG-19特征提取块在插值中添加了高频细节，导致目标分布呈正态分布，极大地降低了学习难度，如图8所示。

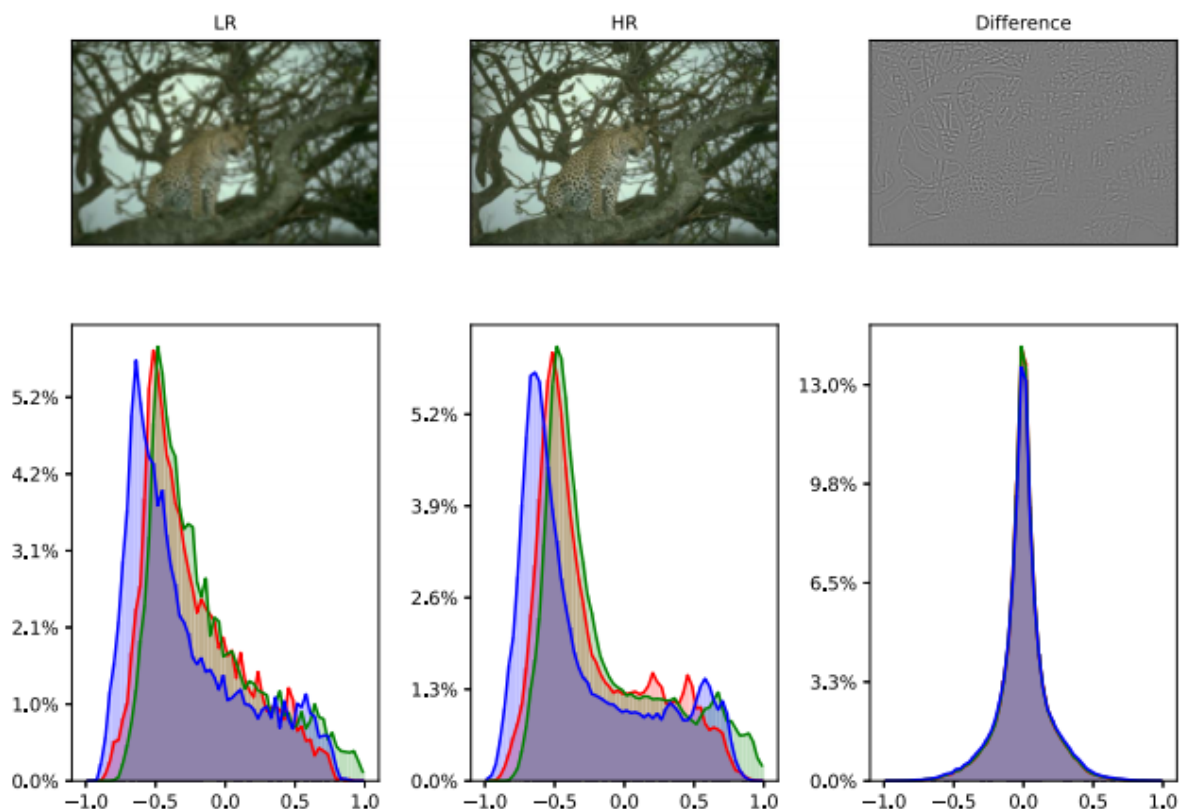


Fig. 8. Pixel value distribution of an example image from BSDS100 [2]. It can be observed that the distribution of the LR and HR image are skewed. In contrary, the distribution of the difference between LR and HR looks normally distributed, which is easier learned by an SR model.

此外，由于插值中高频的稀疏表示，它减少了消失/爆炸梯度。这一优点产生了跟踪残余网络的趋势，从而增加了使用的残差的数量。

一个例子是RED-Net (2016) [101]，它将U-Net [102]架构适应SR。它结合了一个下采样编码器和一个上采样解码器网络，它对给定的输出进行下采样以提取特征，然后将特征映射上采样到目标空间大小。在此过程中，RED-Net利用在整个下采样过程中获得的剩余信息扩展了上采样操作，从而减少了消失的梯度。因此，它在几个缩放因子上都优于SRCNN。

另一个例子采用了SRGAN (2016) [13]的出版物中，即由多个残差单元组成的RenseNet[104]，它将残差信息发送给所有后来出现的卷积操作。此外，作者还比较了这些架构应用于像素和对抗性损失（见第3节）。SRResNet由多个堆叠的剩余单元组成，允许高级特征提取通过大量的求和操作访问低级特征信息。因此，它通过提供一个简单的反向传播路径来简化优化。与SRResNet一样，SRDenseNet [40]应用密集的残差块，它利用更多的残余连接来允许直接路径到更早的层。相比之下，SRResNet的性能大大优于SRCNN、DRCN和ESPCN。对SRDenseNet的一个扩展是在2018年提出的剩余密集网络 [42]，它在密集块上包含了一个额外的剩余连接。

密集残差拉普拉斯网络 (DRLN) [6]是SRDenseNet的扩展，是一种基于后上采样、通道注意的残差网络，并取得了最先进的竞争结果。每个密集块后面都有一个基于拉普拉斯金字塔注意的模块，它学习特征映射之间的层间和层内依赖关系。它在每个DRLM中逐步加权子频带特征，类似于HAN（连接不同深度的各种特征图）。

残差块的另一种变体是信息蒸馏网络 (IDN) [43]。它使用剩余连接将特征映射的一部分积累到以后的层。给定六个卷积层，它将特征映射在中间分成两个部分。然后，其中一部分被最后三层进一步处理，并添加到输入部分和另一部分的连接中。简而言之，利用剩余连接的网络是最先进的。它们有效地传播信息的能力有助于对抗消失/爆炸的梯度，从而产生出色的性能。有时，剩余块的使用会与其他体系结构相结合，例如基于循环的网络。



## 基于递归的网络

人工深度可以通过重复来完成，其中接受野对于获取重要信息至关重要，通过重复相同的操作而扩大。此外，递归性减少了参数的数量，这有助于对抗小型设备的过拟合和内存消耗。它是通过不引入新的参数而多次应用卷积层来实现的。

Kim等人[105]通过DRCN（2015）引入了第一个基于循环的SR网络。它使用相同的卷积层多达16次，随后的重建层考虑所有递归输出进行最终估计。然而，他们观察到，他们的深度递归网络很难训练，但通过跳过连接和递归监督来缓解它，本质上是辅助训练。图9显示了DRCN。

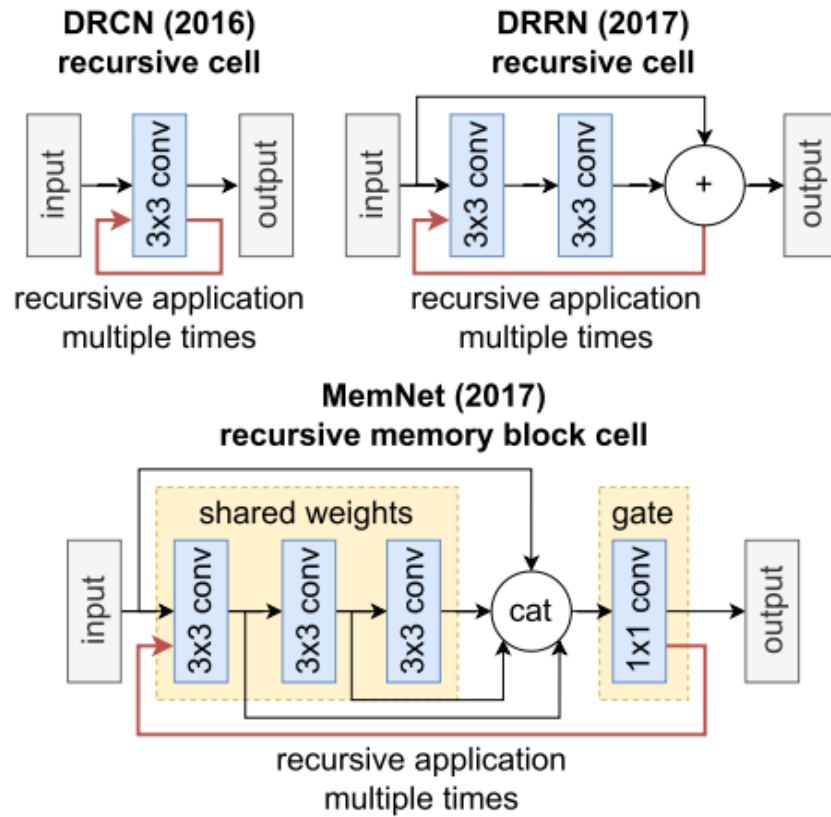


Fig. 9. Recurrent-based architecture designs of DRCN [105] (upper left), DRRN [65] (upper right), and MemNet [92] (bottom). The DRCN applies the same convolutional layer multiple times. The DRRN extends this idea to recursive application of a residual block. The MemNet introduces a gate unit to distill most important features from all intermediate convolution results in a recursive manner.

结合了DRCN [105]和VDSR [98]的核心思想，DRRN（2017）[65]在一个递归块结构中使用了几个堆叠的残差单元。此外，它使用的参数分别比VDSR和DRCN少6倍和14倍，同时获得了更好的结果。与DRCN相反，DRRN在剩余单元之间共享权值集，而不是在所有递归应用的卷积层中共享一个权值。通过强调多路径，DRRN比DRCN（总共52个）训练得更稳定，递归程度更深。

受DRCN的启发，Tai等人介绍了MemNet（2017）[92]。主要贡献是由递归单元和门单元组成的内存块，以挖掘持久内存。该递归单元被应用了多次，类似于DRCN。输出被连接并发送到一个栅极单元，这是一个简单的1x1卷积层。自适应门单元控制先前信息的量和保留的当前状态。图9显示MemNet。引入门对序列到序列任务（如LSTM [106]）的影响是开创性的，但深入了解其对SR任务的影响标志着一个开放的研究问题。

受DRRN的启发，DSRN（2018）[97]的作者探索了一种具有多路径网络的双态设计。它介绍了两种状态，一种在HR操作，另一种在LR空间，它们共同利用LR和HR信号。通过延迟反馈[107]，信号在lr到hr和hr到lr这两个空间之间反复交换。从lr到hr，它使用了一个转置的卷积层来进行上采样。HR-to-LR是通过分层卷积来执行的。最终的近似值使用了在人力资源空间中所做的所有估计值的平均值。因此，它

应用了迭代上下上采样的扩展公式（见第7.1节）。这两种状态使用的参数多于DRRN，但小于DRCN。然而，适当地开发双态设计在未来需要进行更多的探索。

超分辨率反馈网络（SRFBN，2019）[64]也在使用反馈[108]。最基本的贡献是反馈块（FB）作为一个实际的循环细胞。FB使用多个具有密集跳跃连接的迭代上下采样来产生高水平的判别特征。SRFBN为每次迭代生成一个SR图像，并且FB块接收前一次迭代的输出。它尝试在每次迭代中为单个退化任务生成相同的SR图像。对于更复杂的案例，它通过课程学习通过每次迭代返回更好、更好质量的图像（见第6.1节）。与其他框架相比，SRFBN已经显示出了显著的改进，但在未来还需要更多的研究。

Liu等人提出了NLRN（2018）[109]，它提供了一个非局部模块来产生自相似性的特征相关性。图像中的每个位置测量其邻近区域的每个位置的特征相关性。NLRN利用特征相关消息之间的相邻循环阶段。事实上，NLRN的表现也略好于DRCN、DRCN和MemNet。

然而，最近对SR中rnn的主要研究是针对MISR进行的，如视频SR [110]或元学习[111]相关任务。一般来说，基于循环的网络在保存参数方面很有趣，但其主要缺点是通过重复应用相同的操作来实现其计算开销。此外，由于时间依赖性，它们不能并行化。替代方案是轻量级架构，接下来将介绍它们。

## 轻量级网络

---

到目前为止，我们已经引入了能够提高SR图像质量的模型，以及一些尝试去做同样的事情，但计算量较少的模型。例如，FSRCNN [33]利用更小的内核大小、后采样和 $1\times 1$ 卷积层来增强/减少通道维度，从而比SRCNN [39]更快（见7.2节）。本工作中的另一个例子是基于循环的网络，它减少了第7.4节中所述的冗余参数。这些精益递归网络的缺点是，参数的减少是以增加操作和推理时间为代价的，这是现实世界场景的一个基本方面。例如，移动设备上的SR受到电池容量的限制，这取决于所需的计算功率。因此，轻量级体系结构明确地同时关注执行速度和内存使用情况。补充材料，在线提供，包括参数比较，和执行速度的公平比较是受欢迎的需求。

MDSR（2017）[38]使用多路径方法来学习具有共享参数的多个缩放因子。它有三个不相同的路径作为预处理步骤和三个路径作为上采样。对于给定的比例因子 $s \in \{2, 3, 4\}$ ，MDSR在三条路径之间选择确定性的。大尺度的路径比低尺度因子的路径建立得更深。在预处理和上采样步骤之间是一个由多个剩余块组成的共享模块。该特征提取块经过训练，常用于所有的缩放因子，如图10所示。



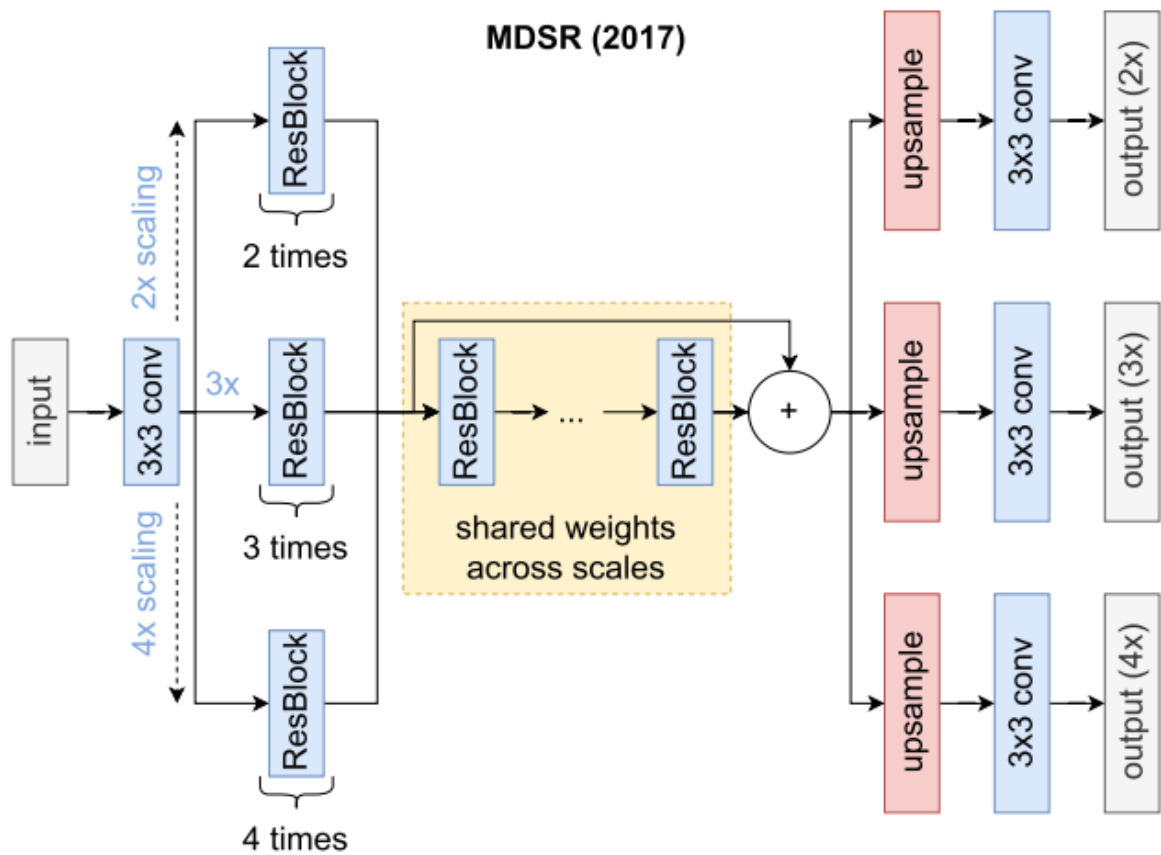


Fig. 10. Architecture design of MDSR [38]. It utilizes multi-path learning to select between multiple scaling factor paths (denoted by 2x, 3x, 4x on the bottom right of the blue boxes). Also, all paths share an intermediate feature extraction block to save parameters.

其主要优点是，一个模型就足以在多个尺度上进行训练，从而节省了参数和内存。相比之下，其他SR模型必须在不同的尺度上独立训练，并独立保存用于多尺度应用。然而，添加一个新的缩放因子需要从头开始进行训练。其他轻量级体系结构也采用了这一想法，以实现参数高效的多尺度训练，如CARN/CARNM (2018) [112]。

此外，它还在残差网络[103]上实现了一种级联机制。CARN由多个级联块（见图11）和它们之间的1x1卷积组成。

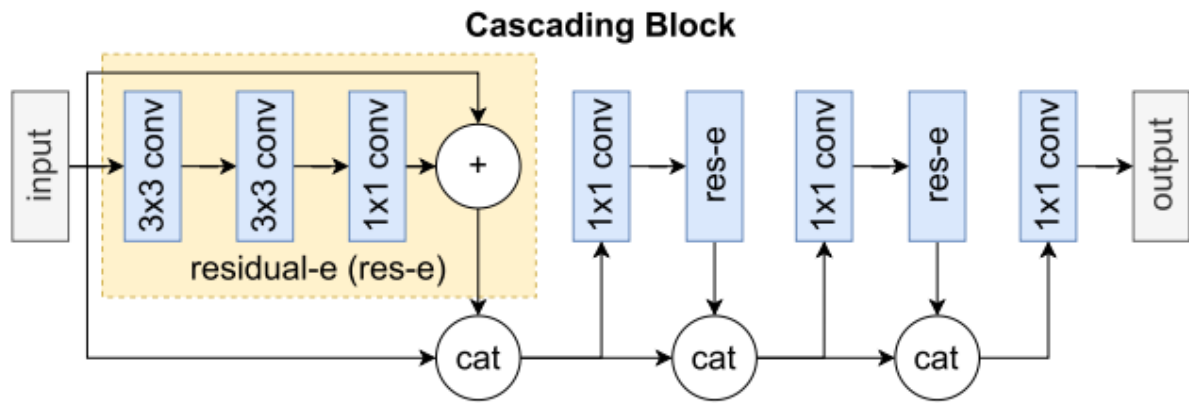


Fig. 11. Design of cascading blocks [112]. It is a densely connected block, which consists of Residual-E blocks [112] and 1x1 convolutions that take as input the Residual-E block's output and the residual connections from the layers before. The Residual-E block itself is built like a residual unit but performs group convolution, which is efficient in inference time, depending on the group size.

级联块的输出将被发送到所有后续的1x1卷积中，就像在级联块本身中一样。因此，局部级联几乎与全局级联完全相同。它允许多层次的表示和稳定的训练，如残余网络。最终，它在三条路径中进行选择，通过类似于MDSR的高效亚像素层，将特征映射上采样到2倍、3倍或4倍的缩放因子。受MobileNet [113]的启发，CARN还在每个残差块组件中使用分组卷积。这允许配置模型的效率，因为选择不同的组大小和由此产生的性能是在一种权衡关系中。具有组卷积的残差块根据组卷积的大小，最多可减少计算14倍。他们测试了CARN的一种变体，它设置了组的大小，从而使计算减少最大化，并将其称为CARN-Mobile (CARN-M)。此外，他们通过允许在每个级联块中的残余块的权重共享（与非共享块相比减少了3倍），进一步减少了CARN-M的参数。

受IDN和IMDB [115]的启发，RFDN (2020) [114]通过使用RFDB块重新考虑了IMDB体系结构，如图12所示。RFDB块由特征蒸馏连接组成，它们将1x1的卷积级联到最后一层。此外，它使用浅层残差块（SRBs），其中只有一个3x3的卷积，来进一步处理给定的输入。最后一层是一个1x1的卷积层，它结合了所有的中间结果。最后，它应用了专门为轻量级模型设计的增强型空间注意力[114]。RFDN架构包括后续的RFDB块，并使用具有最终亚像素层的后上采样框架。

XLSR (2021) [116]是一个非常具有硬件感知能力和量化友好性的网络。它应用多路径来减轻卷积操作的负担，并使用1x1卷积来按像素级进行组合。每个卷积层都有一个较小的滤波器尺寸（8、16、27）。在组合之后，它将分割特征贴图，并再次应用多条路径。XLSR的一个核心方面是末端激活层，它利用了量化的好处。量化是有用的，因为它可以通过使用更多的微型位表示[117]来保存参数。不幸的是，许多移动设备都支持8位数据。因此，对在浮32或浮16中表现良好的SR模型应用uint8量化不起作用。裁剪的ReLU（限制为最大值1）作为最后一个激活层而不是典型的ReLU可以消除这个问题。然而，作者建议通过进一步的实验来寻找其他的最大值。

一般来说，有很多想法可以让SR模型轻量级有待发现。它们包括对现有架构的简化、量化和修剪。此外，利用SR的资源有限的设备和应用是一个日益感兴趣的领域。

## 小波变换网络

不同的图像表示可以带来一些好处，比如提高计算速度。小波理论为表示和存储多分辨率图像提供了稳定的数学基础，描述了上下文和纹理信息[119]。离散小波变换（DWT）将一幅图像分解为一系列小波系数。在SR中最常见的小波是Haar小波，通过二维快速小波变换计算出来。通过对每个输出系数进行迭代重复分解，计算出小波系数。它捕获四个子波段的图像细节：平均（LL）、垂直（HL）、水平（LH）和对角线（HH）信息。DWSR (2017) [120]是第一个使用小波预测的网络之一。它使用了一个简单的网络架构来细化在一个预上采样框架中的LR和HR图像小波分解之间的差异。首先，计算放大（采

用双边插值) LR图像的小波系数。然后对小波系数进行卷积层处理。然后, 加入初始计算的小波系数, 并采用残差连接进行波面处理。因此, 卷积层学习了系数的额外细节。最后, 采用二维-DWT的反过程得到SR图像, 如图13所示。

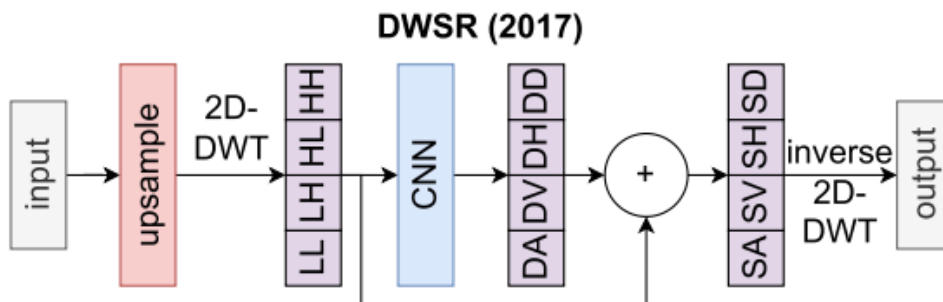


Fig. 13. Visualization of DWSR [120]. It calculates wavelet coefficients and uses a CNN to calculate the difference in the average (DA), vertical (DV), horizontal (DH), and diagonal (DD) sub-bands. Next, it adds the initially calculated coefficients via residual connection to the difference to derive the coefficients of the predicted SR image (SA, SV, SH, and SD), which is obtained by using the reverse 2D-DWT.

利用WIDN (2019) [121]提出了另一种方法, 它使用平稳小波变换代替DWT来实现性能更好。与DWSR使用小波-srnet (2017) [122]的同时, 我们提出了一个更复杂的模型。为了从LR图像中生成特征映射, 它提供了一个由残余块组成的嵌入网络。然后, 它多次应用小波变换, 并利用多个小波预测网络。最后, 它应用反向过程, 并使用转置卷积进行上采样。该系数用于小波损失函数, 而SR图像用于传统的纹理和MSE损失函数。因此, 他们的网络适用于不同放大倍数的不同输入分辨率, 并对MS-SR的未知高斯模糊、姿态和遮挡显示出鲁棒性。多级小波CNNs的思想也可以在以后的出版物中找到, 即MWCNN (2018) [123]。

下面的工作应用了一种混合的方法, 通过混合小波变换与其他著名的SR方法。即Zhang等人提出了一个基于小波的SRGAN (2019) 框架[124], 它融合了SRGAN和小波分解的优点。生成器使用嵌入网络将输入处理到特征映射中, 类似于小波-srnet。接下来, 它使用一个小波预测网络来细化系数, 类似于DWSR。2020年, Xue等人[125]将小波与包含通道注意和空间注意模块的残差注意块(混合注意, 见下文第5节)相结合, 并将其称为网络称为WRAN。在过去的几年里, 小波的应用也应用于视频SR [126]。

一般来说, 小波变换可以有效地表示图像。因此, 使用这种策略的SR模型通常会降低总体模型的大小和计算成本, 同时达到与最先进的架构相似的性能。然而, 这一研究领域还需要更多的探索。例如, 由于高频子带和低高频子带的分布, 合适的归一化技术存在显著差异, 或者由于高频子带的稀疏表示可能不合适, 因此可以替代卷积操作。

## 无监督超分辨率

监督SR的惊人性能归因于它们主要从许多LR-HR图像对学习自然图像的能力, 大多是已知的退化映射, 这在实践中通常是未知的。因此, 经过监督训练的SR模型对于切实可行的用例有时是不可靠的。例如, 当训练数据集生成LR图像(保留高频), 然后SR模型在该数据集上进行的训练不太适用于使用抗锯齿生成的真实LR图像(平滑图像)。使用抗混叠方法生成的LR图像(平滑图像)。此外, 一些专门的应用领域缺乏LR-HR图像对数据集。因此, 人们对无监督SR越来越感兴趣。我们简要地研究了 this 领域, 为了进一步阅读基于流的方法(退化核的密度估计), 我们参考Liu等人[127]的调查。

## 弱监督方法

---

弱监督方法使用未配对的LR和HR图像，如WESPE（2018）[128]。WESPE由两个发生器和两个鉴别器组成。第一个生成器获取一个LR图像并对其进行超级解析。第一发生器的输出构成一个SR图像，但也与电视损失[59]正则化。第二个生成器接受对第一个生成器的预测，并执行逆映射。第二个生成器的结果通过内容丢失[12]与原始输入的LR图像进行优化。这两个鉴别器取第一个生成器的SR图像，并被训练来区分预测和原始HR图像。第一鉴别器根据图像颜色将输入分类为SR或HR图像。第二个鉴别器使用图像纹理[61]来进行分类。

一个类似的方法是一个被称为CinCGAN（2018）[52]的周期中循环SR框架，基于CycleGAN [129]。它总共使用了四个发生器和两个鉴别器。第一个生成器取一个有噪声的LR图像，并将其映射到干净的版本。第一个鉴别器被训练来区分来自数据集的干净LR图像和预测的干净图像。第二台发电机训练逆函数。因此，它从预测的干净版本中生成有噪声图像，从而关闭了一个半周期周期的第一个周期。第三个生成器特别有趣，因为它是实际的SR模型，它将LR图像上采样到HR。第二个鉴别器被训练来区分预测的和数据集的HR图像。最后一个生成器将预测的HR图像映射到有噪声的LR图像，从而关闭CycleGAN的第二个周期。除了其有希望的结果和类似的方法[130]外，它还需要进一步的研究来降低学习难度和计算成本。

## 零次学习

---

零次学习或一次学习与对物体的训练和对从未观察到的完全不同的物体的测试有关。理想情况下，如果将“斑马看起来像条纹马”转换为[132]马，那么用马训练的分类器应该识别斑马。关于SR的零次学习的第一个出版物是ZSSR（2017）[87]。我们的目标是只训练手头的一张图像，一张独一无二的图像。ZSSR对LR图像进行下采样，并训练CNN以反转退化映射。训练后的CNN最终直接用于LR图像。令人惊讶的是，该方法的效果优于SRCNN，与VDSR比较接近。

在此基础上，提出了一种基于深度信息[134]的退化仿真网络（DSN，2020）[133]，以避免预定义的退化内核。它使用双循环训练来同时学习未知的退化核和SR图像的重建。MZSR [135]将ZSSR设置与元学习合并，并使用一个外部数据集来学习不同的模糊内核，这被称为元学习领域中的任务分布。然后将SR模型在类似于ZSSR的降采样图像上进行训练，并从元测试阶段返回模糊核。这种方法的好处是，它使SR模型更快地学习特定信息，比纯ZSSR性能更好。SR的零镜头学习标志着进一步研究的一个令人兴奋的领域，因为它非常实用，特别是对于特定于应用程序的数据集很少或不存在的应用程序。

## 深度图像先验

---

Ulyanov等人[131]提出了深度图像先验（DIP），这与在大数据集上训练CNN的传统范式相矛盾。它使用一个CNN来预测降采样时的LR图像，给定一些随机的噪声，而不是一个实际的图像。因此，它遵循了ZSSR的策略，只使用LR图像。然而，它将输入固定为随机噪声，并对预测采用固定的降采样方法进行修正。此外，它还优化了降采样预测与LR图像之间的差异。然后，CNN在不使用固定降采样方法的情况下生成SR图像。因此，它利用噪声生成一个SR图像，而不是转换一个原始图像。ZSSR和DIP之间的差异见图14。

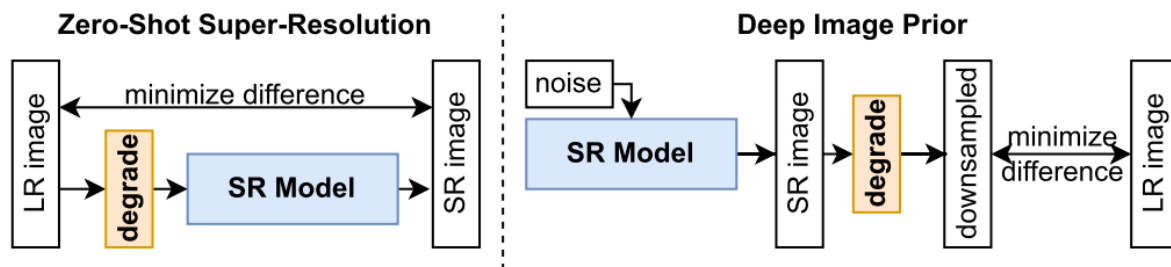


Fig. 14. Zero-Shot Super-Resolution (ZSSR) [87] and Deep Image Prior (DIP) [131]. ZSSR uses the LR image and downsamples it and the SR model learns to reverse the downsampling. For the final prediction of the SR image of the LR image, it is applied to the LR image directly. DIP uses fixed noise as input and predicts the SR image, which is downsampled to optimize the difference between the downsampled image and a given LR image. The final prediction uses the SR model to predict the SR image but skips the degradation mapping.

图14。零镜头超分辨率（ZSSR）[87]和深度图像先验（DIP）[131]。ZSSR使用LR图像进行降采样，SR模型学习反向降采样。对于LR图像的SR图像的最终预测，它直接应用于LR图像。DIP使用固定噪声作为输入，预测SR图像，并进行降采样，以优化降采样图像与给定LR图像之间的差异。最终的预测使用SR模型来预测SR图像，但跳过了退化映射。

令人惊讶的是，研究结果与LapSRN [136]很接近。不幸的是，它是一篇关于图像先验的理论出版物，而且正如作者自己所说的那样，这种方法太慢了，对大多数实际应用都不太有用。然而，它并不排除未来可以提高DIP关于更好的图像重建质量，特别是运行时的实用性的想法。