

# 基于双分支残差结构的岩石图像超分辨研究

田茹梦, 朱联祥

西安石油大学, 陕西 西安

收稿日期: 2022年1月15日; 录用日期: 2022年2月11日; 发布日期: 2022年2月21日

## 摘要

岩石显微图像包含着油气藏分布信息, 清晰度高的岩石显微图像包含更多的岩石物理特性。受限于图像采集设备及自然拍摄环境, 岩石图像往往存在分辨率低, 图像细节不清晰等问题。本文将基于深度学习的卷积神经网络模型应用到岩石图像的超分辨率处理。在对现有网络模型改进的基础上, 提出了具有双分支残差结构的超分辨率重建模型, 并将改进的模型应用于岩石图像数据集, 实现了图像超分辨率效果的改善。实验结果表明本文提出的算法在岩石图像重建的峰值信噪比(PSNR)指标方面有所改善。

## 关键词

岩石图像, 超分辨率重建, 深度学习

# Research on Rock Image Super-Resolution Based on DBResNet

Rumeng Tian, Lianxiang Zhu

Xi'an Shiyou University, Xi'an Shaanxi

Received: Jan. 15<sup>th</sup>, 2022; accepted: Feb. 11<sup>th</sup>, 2022; published: Feb. 21<sup>st</sup>, 2022

## Abstract

Rock micro-image contains the information on oil and gas reservoir distribution. The higher rock micro-image resolution is, the more rock physical properties information it contains. Limited by image acquisition equipment and natural shooting environment, rock micro-image often has problems of low resolution and unclear image details. In this paper, deep learning-based convolutional neural network is researched for rock image super-resolution processing. A super-resolution reconstruction model with double branch residual network structure (DBResNet) is proposed after a modification of the current network model. With the application of proposed method to rock image, better super-resolution performance is obtained. The results show that the proposed algo-

rithm improves the peak signal-to-noise ratio (PSNR) of rock image reconstruction.

## Keywords

Rock Image, Super-Resolution Reconstruction, Deep Learning

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

岩石图像包含了岩石的特性信息, 这些特性信息能够准确反映出油气藏分布情况, 为学者研究油气开发和地质勘探提供有价值的参考[1]。显微镜下分析岩石薄片, 并对其进行分析是地质研究工作中非常重要的一部分[1], 因此清晰、分辨率高的岩石图像对油气田开发以及地质勘探工作十分重要。但是受限于图像采集设备以及自然拍摄环境, 经常无法获得清晰的岩石图像。在此需求的驱动下, 将超分辨率重建技术应用于岩石显微图像, 就显得尤为重要。

对于图像超分辨率的研究, 最早使用的是基于插值的方法。但是基于插值的方法在重建图像细节以及图像纹理信息方面的能力有限[2]。第二种是基于学习的方法。基于学习的方法分为浅层学习和深度学习, 旨在学习低分辨率图像和高分辨率图像对之间的映射[3]。但是基于浅层学习的方法依赖于从临近嵌入到稀疏编码的各种技术[4]。基于深度学习的方法以卷积神经网络为主, 自从 Dong 等人提出的 SRCNN [5]将卷积神经网络应用到图像超分辨处理, 并且取得了显著的效果后, 基于深度神经网络的图像超分辨率研究就引发了广泛的关注。2016年 J. Kim 等人提出的 VDSR [6]将网络深度推进到 20 层, 并实现了重建效果的显著提升。SRResNet 是 SRGAN [7]中的生成网络, 采用了 He 等人提出的残差结构, 将残差网络引入超分辨研究领域。Bee Lim [4] [8]等人指出残差结构中的 BN 层是为了处理更高级别的图像处理任务。对于图像超分辨的研究, BN 层非但不利于图像重建, 而且会大大提高计算量[4]。Bee Lim 等人提出的 EDSR 将 SRResNet 的 BN 层去掉, 减少了大量参数。与此同时, 加深的网络深度可提取更多、更深层次的特征, 在计算量和图像超分辨性能等方面都取得了明显的提升。但是上述模型中都是采用单分支的残差结构, 残差结构中使用单一的卷积核。这不利于提取更丰富的图像特征, 存在重建图像细节不清晰的问题。

针对上述问题, 本文以 EDSR 为基础模型, 并对网络结构进行改进, 即使用双分支的残差结构来代替原有的单分支残差结构, 并使用不同尺寸的卷积核, 提取更丰富的图像特征, 为图像重建提供更多信息, 提升重建质量。本文的其余部分组织如下: 第 1 部分在对 EDSR 算法进行介绍与分析基础上, 提出了具有双分支残差结构(DBResNet)的超分辨处理算法模型。第 2 部分将所提出的算法应用于岩石图像数据集, 通过与 EDSR 算法的对比, 验证了所提算法在超分辨性能上的改进。最后在第 3 部分对全文做了总结。

## 2. DBResNet 超分辨处理算法

具有双分支残差结构的超分辨处理模型是以 EDSR 为基础, 通过使用双分支的残差结构来代替其单分支的残差结构而提出的新的超分辨处理算法。因此, 在介绍新算法之前, 有必要对 EDSR 算法的整体架构、组成模块、工作流程、损失函数, 以及相关的评价指标作简单介绍。

## 2.1. EDSR 的整体架构

EDSR 是对 SRResNet 的改进, 通过去掉其批归一化层来大大减少计算开销, 使得在相同计算资源下可以加深网络的深度, 提取更多特征, 从而提升模型的性能。模型的整体架构如图 1 所示, 同样给出的还有其局部模块的解释说明。

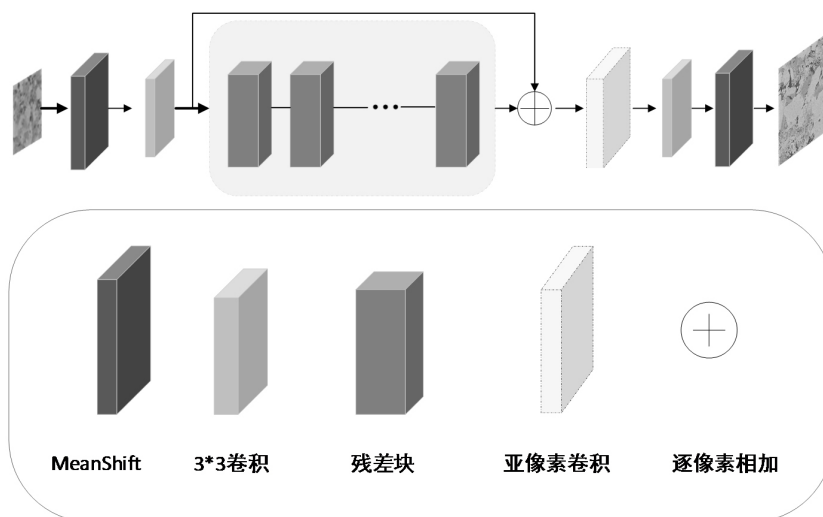


Figure 1. Overall architecture of network model  
图 1. 网络模型整体架构

按照功能, EDSR 模型可分为特征提取和上采样重建两个步骤。特征提取又可分为浅层特征提取模块和残差模块。其第一个卷积层进行浅层的特征提取, 卷积核大小为  $3 \times 3$ , 卷积步长为 1, 后续连接 16 个残差模块, 且最后一级残差块的输出和残差结构的输入相加, 从而形成全局残差。在整个卷积的过程中, 图像尺寸保持不变。特征提取模块的输出进入上采样模块。模型的上采样部分使用的是亚像素卷积的方法。

## 2.2. EDSR 的特征提取

EDSR 模型主体由 16 个残差块线性堆叠构成, 每个残差块结构如图 2 所示。残差网络可以解决深层次网络结构中因网络堆叠带来的梯度消失或退化的问题[4]。每个残差结构的右边支路首先经过一个  $3 \times 3$  的卷积, 卷积步长为 1, 然后经过 ReLU 线性激活单元, 接着再经过一个  $3 \times 3$  卷积, 最后卷积的输出与参与残差结构的输入进行逐像素相加操作, 结果即为此残差结构的输出。通过多个残差结构的级联处理, 可以得到原输入图像不同深度的特征信息。

## 2.3. EDSR 的上采样重建

特征提取模块的输出进入上采样模块, 从而实现原图像的高分辨率重建。目前现有的上采样方法主要分为两大类: 基于插值的方法和基于学习的方法[2]。基于插值的方法是指通过传统的图像插值操作将图像放大, 原理相对简单, 但是会引入噪声放大的副作用[2]。基于学习的上采样方法有转置卷积法和亚像素卷积法, 这两种方法也被应用到不同的网络模型。例如转置卷积被应用到 SRDenseNet [9], 亚像素卷积被应用到 RDN [10]和 SRGAN [7]等模型。由于亚像素卷积比转置卷积有更大的感受野, 能够提供更多的邻域信息, 更有利于重建图像细节[2], 因此被 EDSR 所采用。

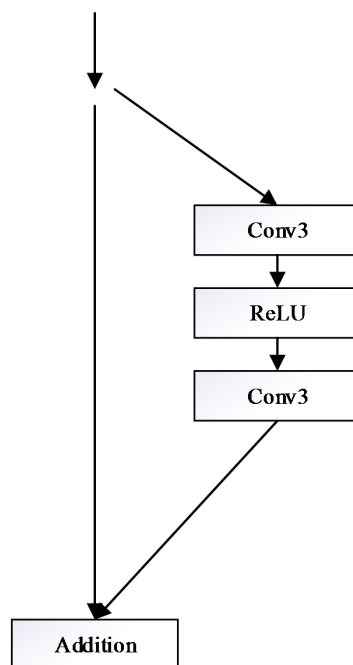


Figure 2. Residual structure of EDSR  
图 2. EDSR 的残差结构

亚像素卷积是通过卷积操作得到多个通道的特征图，然后对这些特征图进行重新排列，并在多个通道之间重组，从而放大图像。图 3 以放大因子 2 为例，展示了亚像素卷积的工作过程。其中输入图像为  $W \times H$ ，经过卷积处理得到 4 个尺寸为  $W/2 \times H/2$  的特征图，然后对 4 个特征图对应位置进行排列重组，得到  $2W \times 2H$  的图像，完成图像的上采样。

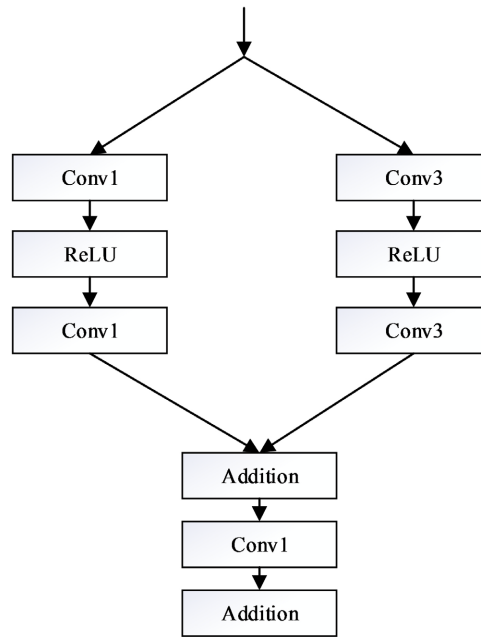


Figure 3. Subpixel convolution  
图 3. 亚像素卷积

#### 2.4. DBResNet 超分辨处理对 EDSR 的改进

岩石图像中岩石的特性不仅需要通过轮廓、颜色等简单特征来描述，更需提取其它丰富的特征[11]。大小不同的卷积核执行不同的卷积运算，不同的卷积运算获得图像的不同信息[12]。使用卷积核不同的双分支卷积运算，可以提供更多的激活函数，增加网络的非线性表达能力。两个支路的卷积核提取的特征信息采用相加进行特征的融合，从而获得图像更丰富的特征信息[12]。此外，尽管残差结构可以在一定程度上解决深层次网络的梯度消失或退化问题，但随着深度的增加，网络的参数数量及所需的训练时间仍会急剧增大，甚至带来无法收敛的问题。采用双分支结构的残差块，理论上可以用更少的网络深度获得相同或相近的性能。

本文对 EDSR 中的残差单元进行改进，改进后的残差块变为一个如图 4 所示的双分支结构，其中左分支采用  $1 \times 1$  的卷积核，右分支采用  $3 \times 3$  的卷积核。 $1 \times 1$  的卷积核提取图像的细节信息， $3 \times 3$  的卷积核提取其轮廓信息。



**Figure 4.** Double branch residual structure  
**图 4.** 双分支残差结构

## 2.5. 损失函数

在图像超分辨率的研究中, 常用的损失函数有  $L_2$  均方误差损失(MSELoss)和  $L_1$  平均绝对误差(MAELoss)损失。由于  $MSE$  的取平方运算会放大较大误差和较小误差之间的差距, 而  $MAE$  对样本的异常值显得更加健壮, 因此本文的算法采用  $L_1$  损失。(1)式给出  $L_1$  损失的数学表达, 其中  $r$  为放大因子,  $W$  和  $H$  分别表示图像的高和宽,  $f_H$  表示高分辨率图像,  $f_S$  表示通过模型重建的超分辨率图像。

$$l_{MSE}^{SR} = \frac{1}{r^2WH} \sum_{i=1}^{rW} \sum_{y=1}^{rH} |f_H(x, y) - f_S(x, y)| \quad (1)$$

## 2.6. 评价指标

本文从两个方面对重建效果进行评价: 峰值信噪比(PSNR, Peak Signal-to-Noise Ratio)和主观感受。峰值信噪比(PSNR)的数学表达见式(2):

$$PSNR = 10 \cdot \log_{10} \left( \frac{MAX_I^2}{MSE} \right) \quad (2)$$

其中  $MSE$  表示两张图像的均方误差,  $MAX_I$  表示重建图像的最大像素值。 $MSE$  的数学表达见式(3):

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \|I(i, j) - \hat{I}(i, j)\|^2 \quad (3)$$

峰值信噪比(PSNR, Peak Signal-to-Noise Ratio)已经成为图像超分辨研究领域应用最广泛的评价指标之一。

## 3. 实验结果及过程分析

### 3.1. 数据集

实验使用新南威尔士大学 Wang 等人制作的岩石数据集。该数据集包含三种不同类型的岩石显微图

像: 碳酸盐(carbonate)、砂岩(sandstone), 以及二者的混合(shuffled)。本实验选取其中的碳酸盐图像数据集(carbonate2D)。carbonate2D 一共 1000 张, 包含了 800 张训练图像, 100 张验证图像和 100 张测试图像。

### 3.2. 实验过程

首先通过随机水平翻转、旋转 90 度等操作进行数据的增强。继而进行将所有的图像减去其均值的预处理。在网络的学习阶段, 设置的初始学习率为 0.0001, batch\_size 为 16, 并用 L1 损失代替 L2 损失, 优化采用 ADAM 优化器。网络性能的提升可通过增加参数数量来实现[4], 因此在卷积神经网络中, 可以通过堆叠更多的层数, 或者增大通道数, 来实现网络性能的提升。但当层数增加到一定程度时, 会使得训练极不稳定。因此本文采用的是增加通道数的方法, 实验中的通道数设置为 64。

实验使用 Python3.8 环境下的 Pytorch 框架。计算平台为 Intel 志强, 处理器为双路 E5-2660v3, 一共有 20 核 40 线程, 主频基准为 2.6 GHz, 最高为 3.3 GHz。内存 21 GB, 频率为 2133 Hz。

### 3.3. 实验结果及分析

#### 3.3.1. 相同网络深度时基于双分支残差的模型与原始 EDSR 模型的性能对比

采用与原始 EDSR 模型相同的层数设置, 即残差块的个数为 16, 通道数设置为 64。图 5 给出了两种算法在放大系数为 2 时的 PSNR 随迭代次数的变化曲线, 图 6 为 Bicubic, EDSR 以及本文提出的算法的重建效果视觉对比。从图 5 可以看出, 双分支残差结构的 PSNR 值整体上高于原始的 EDSR, 平均提高了 0.064 dB。

双分支的结构优点是可以融合更多的图像信息。从视觉效果来看, 图 6(d)的图像比 Bicubic 和 EDSR 的重建效果都要好。重建图像的岩石颗粒更加明显, 更接近图 6(a)所示的真实高分辨率图像。

图 7 和图 8 分别是在放大因子为 4 时的 PSNR 随迭代次数的变化曲线及重建效果视觉对比。从图 7 可以看出, 双分支残差结构的 PSNR 值比 EDSR 平均提高了大约 0.055 dB。图 8 可以看出本文提出的算法对比 Bicubic 有较明显的提升, 对比 EDSR 重建的线条更加清晰, 视觉效果上有所改善。实验结果表明, 对于放大因子为 4 和 2 的超分辨率重建结果, 改进后的基于 DBResNet 结构的模型, 优于原始的网络模型。

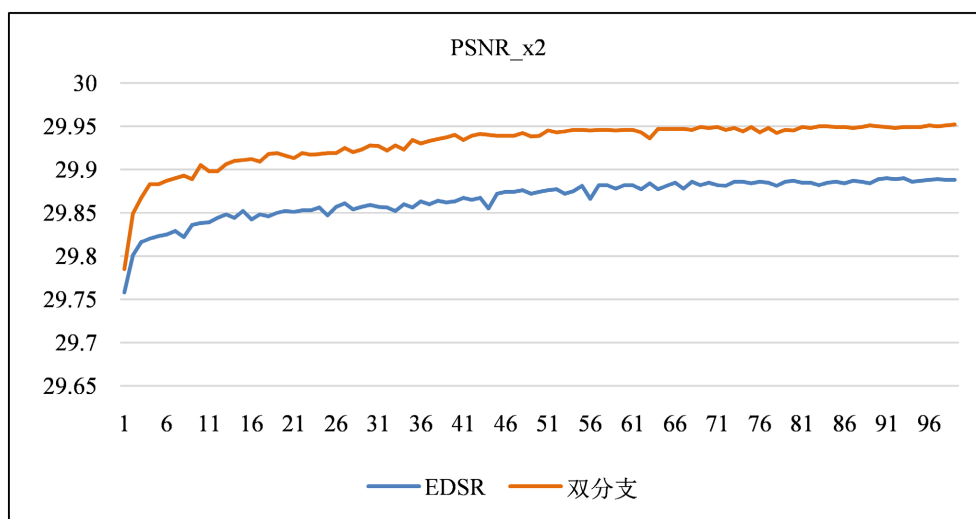
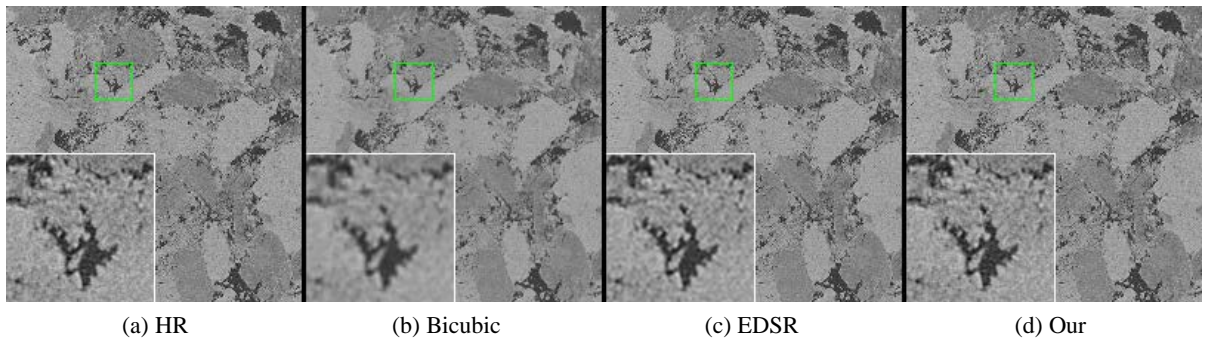
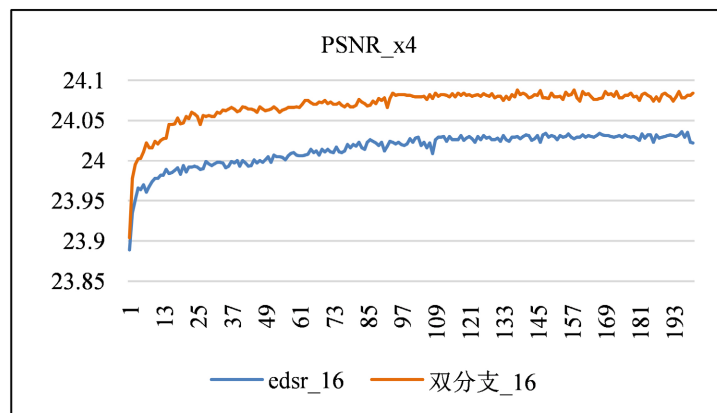


Figure 5. Comparison of PSNR values between EDSR and dual-branch structure when the amplification factor was 2

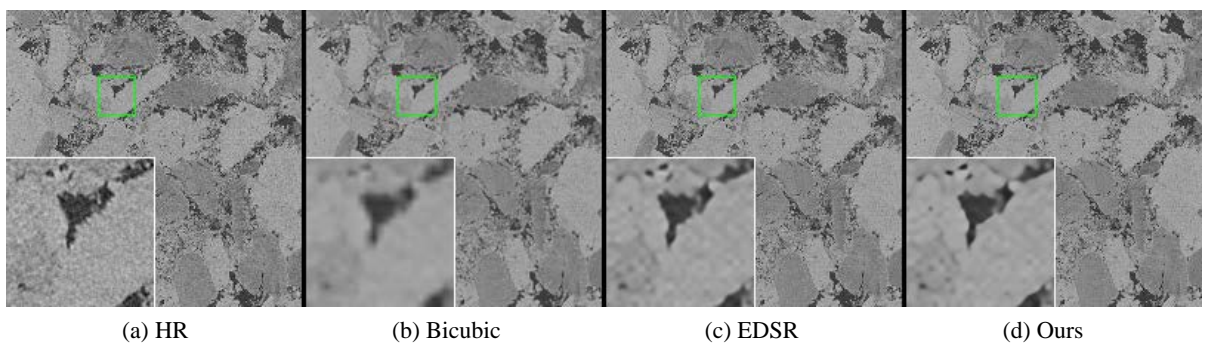
图 5. 放大因子为 2 时 EDSR 和双分支结构模型的 PSNR 值对比



**Figure 6.** A  $\times 2$  scale magnification of different models in the carbonate2D dataset (0801 $\times 2$ )  
**图 6.** 碳酸盐数据集(0801 $\times 2$ )不同模型放大 2 倍效果图



**Figure 7.** Comparison of PSNR values between EDSR and dual-branch structure when the amplification factor was 4  
**图 7.** 放大因子为 4 时 EDSR 和双分支结构模型的 PSNR 值对比



**Figure 8.** A  $\times 4$  scale magnification of different models in the carbonate2D dataset (0810 $\times 4$ )  
**图 8.** 碳酸盐数据集(0810 $\times 4$ )不同模型放大 4 倍效果图

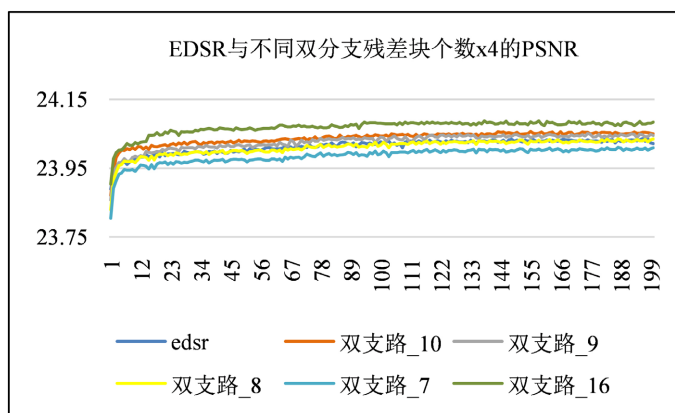
### 3.3.2. 不同深度 DBResNet 的重建性能

由于在理论上, 采用双分支残差结构的网络模型可以用更少的网络层数来获得与原始 EDSR 算法相同的性能, 本实验对不同网络深度下的双分支残差模型在岩石图像超分辨处理上的性能进行了实验仿真。表 1 给出了在放大因子为 4 时, 原始 EDSR 模型及不同网络深度的双分支模型的数量及所需的训练时间。可以看出, 对于同样 16 层的网络深度, 双分支结构的参数量比原始 EDSR 增加了 47.7%, 训练时间增加了 60%。由于双分支提取了图像更丰富的信息, 因此可适当减少残差块数量。本文试验了深度为 10,

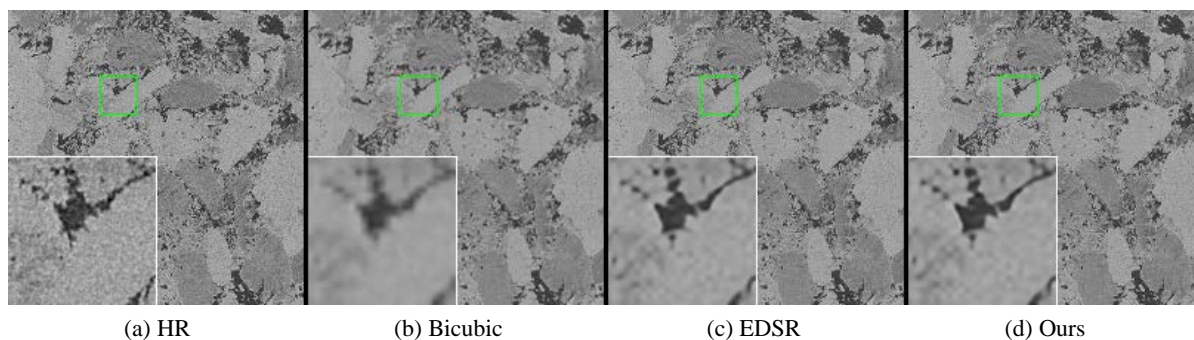
9, 8, 7 几种情况的双支路残差模型在岩石图像超分辨处理上的性能。图 9 是其 PSNR 性能曲线, 可以看出, 深度为 8, 9, 10 的双分支模型的性能很是相近, 层数为 7 时则开始明显下降。在层数为 8 时, 双分支模型的 PSNR 曲线与 EDSR 有很大重合, 基本上达到了原 EDSR 在网络深度为 16 时的性能。图 10 所示为网络深度为 8 的双分支模型在岩石图像超分辨重建效果上与其他方法的对比。从视觉效果来看, 双分支结构获得了与 EDSR 相类似的性能, 且在局部表现来看似乎更优。其参数数量相较于 EDSR 则减少了大约 15%, 训练时间缩短了 54%。

**Table 1.** Comparison between parameters and training time of different model at  $\times 4$   
**表 1.** 不同模型的参数量及训练时间对比( $\times 4$ )

模型	参数量	训练时间
edsr	1,517,595	5 h
双支路残_Res = 10	1,526,930	3 h
双支路残差_Res = 9	1,407,835	2 h 38 min
双支路残差_Res = 8	<b>1,288,731</b>	2 h 18 min
双支路残_Res = 16	2,241,563	8 h



**Figure 9.** PSNR curves of different residual blocks  
**图 9.** 不同残差块的 PSNR 曲线



**Figure 10.** A  $\times 4$  scale magnification of different models in the carbonate2D dataset (0808 $\times 4$ )  
**图 10.** 碳酸盐数据集(0808 $\times 4$ )不同模型放大 4 倍效果图



## 4. 总结

本文在对 EDSR 算法进行剖析与研究的基础上, 通过对原有的残差结构的改进, 即使用双分支的残差结构来替换原有的单分支结构, 提出了基于 DBResNet 的图像超分辨率处理改进算法, 并将其应用到岩石图像数据集。实验结果表明, 改进后的模型不管是 PSNR 还是视觉效果, 都取得了一定改善。

此外还证明, 在保证相同超分辨率重建效果的前提下, 使用改进后的残差结构可显著降低网络深度, 从而大大减少参数数量, 加快网络训练的学习时间。如要获得与深度为 16 的 EDSR 相当的超分辨率处理效果, 所提出的 DBResNet 结构只需要深度为 8, 此时的网络参数量减少了 15%, 相应的训练时间也明显缩短。

影响超分辨率图像重建效果的因素不光有网络模型, 还与所采用的损失函数定义有关。本文采用 MAE 损失函数, 对于 PSNR 值可能有较好的结果, 但对于重建的视觉效果则不一定具有优势。因此在以后的实验中, 可以考虑其他类型的损失函数定义, 并引入与人类认知较为符合的感知损失项, 以期取得更好的超分辨率效果。

## 参考文献

- [1] 程国建, 张福临. 基于 SinGAN 的岩石薄片图像超分辨率重建[J]. 西安石油大学学报(自然科学版), 2021, 36(2): 116-121.
- [2] 张德, 林青宇, 郭茂祖. 单幅图像超分辨率重建的深度学习综述[J]. 计算机工程与应用, 57(22): 1-16.
- [3] 李彬, 王平, 赵思逸. 基于双重注意力机制的图像超分辨率重建算法[J]. 图学学报, 2021, 42(2): 206-215.
- [4] Lim, B., Son, S., Kim, H., Nah, S. and Lee, K.M. (2017) Enhanced Deep Residual Networks for Single Image Super-Resolution. 2017 *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Honolulu, 21-26 July 2017, 1132-1140. <https://doi.org/10.1109/CVPRW.2017.151>
- [5] Dong, C., Loy, C.C., He, K.M. and Tang, X.O. (2014) Learning a Deep Convolutional Network for Image Super-Resolution. *European Conference on Computer Vision (ECCV)*, Zurich, September 2014, 184-199. [https://doi.org/10.1007/978-3-319-10593-2\\_13](https://doi.org/10.1007/978-3-319-10593-2_13)
- [6] Kim, J., Lee, J.K. and Lee, K.M. (2015) Accurate Image Super-Resolution Using Very Deep Convolutional Networks. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, July 2016, 1646-1654.
- [7] Ledig, C., Theis, L., Huszar, F., Caballero, J., Aitken, A., Tejani, A., Totz, J., Wang, Z.H. and Shi, W.Z. (2017) Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 105-114. <https://doi.org/10.1109/CVPR.2017.19>
- [8] Dong, C., Loy, C.C. and Tang, X.O. (2016) Accelerating the Super-Resolution Convolutional Neural Network. *Proceedings of the European Conference on Computer Vision (ECCV)*, Amsterdam, October 2016, 391-407. [https://doi.org/10.1007/978-3-319-46475-6\\_25](https://doi.org/10.1007/978-3-319-46475-6_25)
- [9] Tong, T., Li, G., Liu, X.J. and Gao, Q.Q. (2017) Image Super-Resolution Using Dense Skip Connections. 2017 *IEEE International Conference on Computer Vision (ICCV)*, Venice, 22-29 October 2017, 4809-4817. <https://doi.org/10.1109/ICCV.2017.514>
- [10] Zhang, Y.L., Tian, Y.P., Kong, Y., Zhong, B.N. and Fu, Y. (2018) Residual Dense Network for Image Super-Resolution. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Lake City, 18-23 June 2018, 2472-2481. <https://doi.org/10.1109/CVPR.2018.00262>
- [11] 余晓露, 叶恺, 杜崇娇, 宫晗凝, 马中良. 基于卷积神经网络的碳酸盐岩生物化石显微图像识别[J]. 石油实验地质, 2021, 43(5): 880-885+895.
- [12] 李现国, 冯欣欣, 李建雄. 多尺度残差网络的单幅图像超分辨率重建[J]. 计算机工程与应用, 2021, 57(7): 215-221.