

基于级联卷积网络的红外图像超分辨率重建

甄有恒^{1,2}

(1. 四川大学计算机学院, 成都 610065; 2. 中国人民解放军 95876 部队, 张掖)

摘要:

针对基于卷积神经网络的红外图像超分辨率重建方法存在细节特征提取困难的问题,提出一种小样本量残差学习的两级卷积神经网络方法实现特征的分步重建。首先,利用编解码子网络快速提取边缘特征;然后,级联递归残差子网络用以进一步提取纹理特征;最后,采用全局跳跃连接实现端到端的超分辨率重建。在 Caffe 环境下的实验结果表明,该方法较好地提升拟合能力。

关键词:

卷积神经网络; 红外图像; 图像超分辨率

0 引言

红外成像技术利用目标热源信息差生成图像空间数值,能够有效克服可见光图像无法实现穿透检测的缺陷,已经在军事、工业、遥感、医学等领域有着广泛应用。然而,受红外传感器工艺限制,红外成像质量明显低于可见光成像,主要表现在成像分辨率低以及各类非均匀性噪声显著^[1]。因此,提升红外图像分辨率具有重要的应用价值。面对硬件性能研发困难的问题,算法设计更能节约研发成本。

图像超分辨率重建是图像退化的反过程,其本质是回归问题^[2]。从机器学习过程看,超分辨率重建方法可以分为人工传授、自主学习和半自主学习 3 种形式。人工传授基于传统算法模型设计,需要扎实的学科理论基础,但是设计过程需要具备丰富的图像处理先验知识;自主学习形式以深度学习为典型代表,依靠高维复杂的映射函数实现端到端的学习,但是往往需要高算力和海量数据。半自主学习形式融合人工传授和自主学习两种形式的优点,通常采用传统算法优化输入端,然后再自主映射学习,但是对模型设计能力要求高。

利用深度学习技术实现图像超分辨率重建的模型中,基于重构^[3]、残差^[4-5]、生成对抗^[6]等的卷积神经网络

方法处理可见光图像效果明显。受此启发,文献[7]首次将卷积神经网络用于红外图像增强处理,采用 4 层网络对图像进行端到端的重构。文献[8]对 MNIST 数据集进行伽马变换和高斯滤波方法实现红外数据仿真,利用多尺寸卷积核提取不同规模特征来实现增强对比度、突出红外弱小目标和抑制背景杂波。以上两种方法均采用小型网络结构,有利于低运算量条件下的部署,但是网络层数太浅导致模型对特征的非线性映射能力变得较差,无法提取更高频的红外图像特征信息。文献[9]在可见光图像亮度域完成特征提取、映射和重建过程的模型训练,通过迁移学习将训练模型用于初始化红外测试模型。

以上方法均取得了一定的效果,但是在高倍数重建任务中,网络模型缺少从低频特征向高频特征的信息补充,对细节特征的处理能力较低。因此,本文在 VDSR 网络模型基础上进行改进,提出一种级联重建网络模型 CCNSR (Cascaded Convolutional Network for Super-Resolution) 学习全局残差特征。该模型第一级网络利用卷积和反卷积操作进行编解码学习,第二级网络对编解码后的特征图进行递归残差学习,两个网络级联后,利用全局跳跃连接使得两级网络学习端到端的整体残差。实验验证了本文模型在 VDSR 模型基础

上的改进效果,同时,对其他模型进行性能分析。

1 本文方法

1.1 问题描述

设 I_{LR} 表示红外低分辨率图像, I_{HR} 表示高分辨率图像, I_{SR} 表示重建图像。重建模型描述如下:

$$I_{SR} = F(I_{LR}, \nu) \quad (1)$$

其中 I_{LR} 表示低分辨率图像, I_{SR} 表示超分辨率重建图像, $F()$ 表示重构映射函数, ν 表示 $F()$ 的参数集。

全局残差学习的主要优势在于:网络模型只需要学习输入输出之间的残差信息,充分利用残差特征的稀疏性用以实现训练过程的快速收敛。残差学习模型描述如下:

$$I_{SR} = R(I_{LR}, \omega) + I_{LR} \quad (2)$$

其中 $R()$ 表示残差映射函数, ω 表示残差映射参数集。

模型目标函数如下:

$$y = \arg \min \frac{1}{2N} \sum_{i=1}^N \|I_{HR(i)} - I_{SR(i)}\|^2 = \arg \min \frac{1}{2N} \sum_{i=1}^N \|I_{HR(i)} - I_{LR(i)} - R(I_{LR(i)}, \omega)\|^2 \quad (3)$$

其中 N 表示样本容量,通过对参数集 ω 的迭代训练,达到残差学习模型优化的目的。

1.2 网络结构

本文前馈卷积神经网络结构主要由编解码子网络和递归残差子网络级联组成。网络结构及流程见图 1。

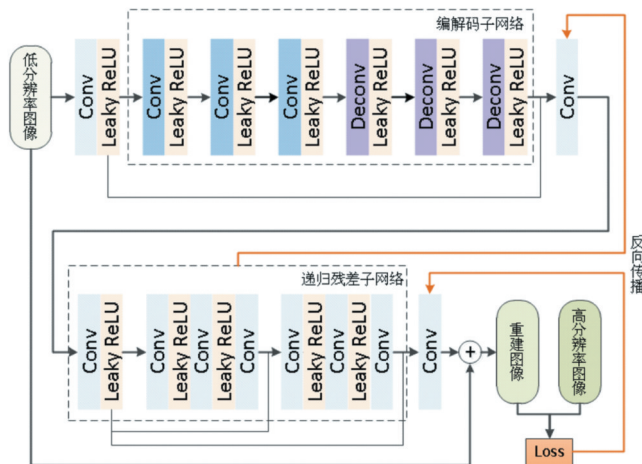


图1 CCNSR网络结构及流程图

编解码子网络的设计主要参照 UNet 模型^[10]。该模型在图像语义分割中有很好的表现,有利于快速提取图像轮廓特征。编解码结构设计中需要计算卷积操作后的特征图尺寸,其计算公式:

$$F_y = \frac{\lfloor F_x + 2P - K \rfloor}{S} + 1 \quad (4)$$

其中 K 表示卷积核尺寸, P 表示边缘填充数, S 表示卷积核步长; F_x 表示卷积前的特征图尺寸, F_y 表示卷积后的特征图尺寸, $\lfloor \cdot \rfloor$ 表示向下取整。反卷积计算公式如下:

$$F_y = (F_x - 1) \times S + K - 2P \quad (5)$$

在网络内部,卷积核和反卷积核参数均设为 $K=3$, $S=1$, $P=0$ 。对称设置确保了特征图先降采样后升采样的编解码功能。子网络首位两端跳跃连接确保了低语义局部残差特征的快速学习。

递归残差子网络受 DRRN 模型^[5]启发,在编解码子网络提取的初级残差特征基础上深度提取高频残差信息。该子网络内部全部采用卷积操作,卷积核统一设为 $K=3$, $S=1$, $P=1$,用以保证特征图在输入输出前后始终保持一致。递归跳跃连接实现低频残差特征逐步向高频残差特征做补充,同时丰富的连接数量有利于反向传播梯度更新。文献[11]解释了批量正则化层(Batch Normalization, BN)对重建任务较为敏感,因此本文去除 BN 层。同时,取消递归部分连续三层卷积中最后一层的传递函数,用以消除网络结构冗余。

另外,除了两级子网络特征融合层卷积核数量设为 1,其余层卷积核数量均设为 15。激活函数统一采用 Leaky ReLU 函数,特点是在 ReLU 激活函数基础上对负向输出添加激励因子 λ ,适度更新神经元抑制参数。

1.3 损失函数和优化

本文采用 L_2 损失函数用以最小化图像均方误差。定义如下:

$$loss = \min \frac{1}{2N} \sum_{i=1}^N \frac{1}{W \times H} \sum_{j=1}^W \sum_{k=1}^H \|I_{HR(i)}(j, k) - I_{SR(i)}(j, k)\|^2 \quad (6)$$

其中 W 、 H 分别表示单个样本的宽和高, $I_{(i)}(j, k)$ 表示图像某像素点的值。

参数优化采用目前主流的 Adam 算法^[12],其在梯度更新方向和更新幅度方面很好的做到了自适应性。更

新公式如下:

$$g^{(l)}(k) = \frac{\partial O^{(l)}(k, \omega)}{\partial \omega^{(l)}(k)} \quad (7)$$

$$p^{(l)}(k) = \frac{\beta_1 p^{(l)}(k-1) + (1-\beta_1) g^{(l)}(k)}{(1-\beta_1(k))} \quad (8)$$

$$q^{(l)}(k) = \frac{\beta_2 q^{(l)}(k-1) + (1-\beta_2) (g^{(l)}(k))^2}{(1-\beta_2(k))} \quad (9)$$

$$\omega^{(l)}(k+1) = \omega^{(l)}(k) - \alpha \frac{p^{(l)}(k)}{\sqrt{q^{(l)}(k)} + \varepsilon} \quad (10)$$

其中 l 表示网络层序号, k 表示迭代次数, α 表示学习率, $\beta_1(k)$ 、 $\beta_2(k)$ 表示学习率衰减参数。

在式(7)中, $O^{(l)}(k, \omega)$ 表示从输入层到第 l 层的映射函数, $\omega^{(l)}(k)$ 表示第 l 层参数, $g^{(l)}(k)$ 表示迭代计算到 k 次时 $O^{(l)}(k, \omega)$ 的梯度, 并且当 l 为输出层时, $O^{(l)}(k, \omega) = loss$; 式(8)表示带有动量的梯度下降迭代计算, 用于控制梯度更新的方向, 式(9)表示均方根梯度下降迭代计算, 用于控制梯度更新的幅度。式(8)、式(9)分母用于修正数值; 式(10)利用式(8)和式(9)调整后的新梯度进行权值参数更新。根号是对式(9)中各维度的梯度分别求平方根, ε 保持分母数值稳定。

2 实验与结果

2.1 数据集

鉴于可见光图像集训练红外图像超分辨率重建模型的方法较为成熟^[7-9], 本文采用 General-100^[13] 和 Urban-100^[13] 两个可见光图像数据集, 为红外图像提供丰富的细节特征。同时, 高倍数差能够为拟合训练提供更高频的残差特征, 因此设置高倍数训练条件。

构建训练集, 首先对数据集进行数据增强处理, 得到 1600 张图像, 并分割成 32×32 的图像块; 然后, 利用双三次插值(Bicubic)对原始高分辨率图像进行 8 倍降采样和升采样, 得到相同尺寸的低分辨率图像; 最后, 配对图像块后构成高-低分辨率训练集。测试集取自公开的 LTIR 红外数据集^[14]和 FLIR 热红外数据集^[15], 共随机挑选 24 张图像, 测试集组成方法同训练集。

2.2 评价指标

下面介绍两种常用的评价指标:

峰值信噪比(Peak Signal-to-Noise Ratio, PSNR)从全局衡量图像重建效果。峰值信噪比公式如下:

$$MSE = \frac{1}{W \times H} \|I_{HR} - I_{SR}\|^2 \quad (9)$$

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

其中 MSE 表示均方误差, P_{\max} 表示图像像素最大值, 对于 8 位图像取 $P_{\max} = 255$; PSNR 值越高, 图像重建质量越好。

结构相似性(Structural SIMilarity, SSIM)^[16] 基于人眼视觉感受, 从图像亮度、对比度和结构三个方面全面评价图像整体复原质量。SSIM 值处于 0 到 1 之间, 越接近 1 越好。结构相似性公式如下:

$$SSIM = \frac{(2\mu_{HR}\mu_{SR} + C_1)(2\sigma_{HR,SR} + C_2)}{(\mu_{HR}^2 + \mu_{SR}^2 + C_1)(\sigma_{HR}^2 + \sigma_{SR}^2 + C_2)} \quad (11)$$

其中 μ_{HR} 表示 I_{HR} 的平均像素强度; μ_{SR} 表示 I_{SR} 的平均像素强度; σ_{HR} 表示 I_{HR} 的像素标准差; σ_{SR} 表示 I_{SR} 的像素标准差; $\sigma_{HR,SR}$ 表示协方差; C_1 、 C_2 取常数用于稳定分母。

2.3 实验流程

本文采用 Caffe 深度学习框架下的 CPU 训练模式, 借助 MATLAB 和 MatConvNet 对实验结果进行分析。选取 TEN、VDSR 和 DRRN 三种典型超分辨率重建模型进行对照实验, 具体操作如下:

(1) 按照网络结构编写 *.prototxt 文件, 将训练 batch_size 设为 128, 测试 batch_size 设为 2, 数据集保存成 HDF5 数据格式。

(2) 编写 Caffe 网络训练配置 *_solver.prototxt 文件并初始化训练参数。卷积层权值参数初始化采用 MS-RA 方法, 初始学习率 $\alpha = 10^{-4}$, Leaky ReLU 激活函数负向激励因子 $\lambda = 0.2$ 。参数优化选择 Adam 方法, 其中学习率衰减参数 $\beta_1 = 0.9$ 、 $\beta_2 = 0.999$, 衰减系数 $\varepsilon = 10^{-8}$, 最大迭代训练次数 $Iteration = 50000$ 。

(3) 训练模型, 迭代训练 50000 次后读取 Blobs 数据块并保存权值参数。输入测试集, 使用 MatConvNet 框架读取训练好的权值参数并按照前馈流程生成重建图像, 计算评价指标。

(4) 对 TEN、VDSR 和 DRRN 模型分别按照(1)到(3)依次操作。调整 VDSR 和 DRRN 的卷积核数量为每层 15 个, 得到 VDSR(15)和 DRRN(15), 用以保证实

验相对公平。

2.4 结果分析

实验重点测试本文模型在 VDSR 基础上改进后的性能提升,并对实验中各个模型的网路结构属性进行了相关统计,见表 1。

表 1 各模型网路结构主要参数

模型	权重值参数量	层数	跳跃连接数	BN 层
TEN	63840	4	0	无
VDSR (15)	36720	20	1	无
DRRN (15)	36720	20	10	有
CCNSR (Ours)	26595	15	4	无

其中 TEN 是基于重构的红外图像超分辨率重建代表模型;VDSR 算法是本文方法的模型基础;DRRN 算法用于验证 BN 层对重建任务的影响。本文模型的网路权值参数量是 TEN 的 42%,是 VDSR (15) 和 DRRN(15)的 72%,网路结构参数量少。

实验分别对两组测试集进行测试,分别得到 4 倍和 8 倍放大倍数下的平均评价指标值,统计结果见表 2,其中横向最优值加粗显示。图 2、图 3 分别表示对应重建倍数的性能变化曲线。

由表 2 统计数据,结合图 2、图 3 曲线分析可以得出:①本文方法在参数量减少的训练条件下,拟合 PSNR 和 SSIM 比 VDSR 模型稍有提升,但是在泛化能力上表现不足,这与编解码子网路中的卷积和反卷积操作过多有关,使得边缘分割严重。②传统的双三次差值方法和基于重构的 TEN 算法对低倍数下的图像重建效果较好,但是随着重建倍数的上升,重构方法的性能下降明显;基于残差学习的方法在不同倍数的重建任务中的影响,因此残差学习可以有效克服对模型参数量的需求。③带有 BN 层的 DRRN 模型在本文实验条件下的重建能力不太突出,原因之一是训练数据量少导致模型欠拟合,并从侧面印证了添加 BN 层的操作不一定适用于图像重建任务。

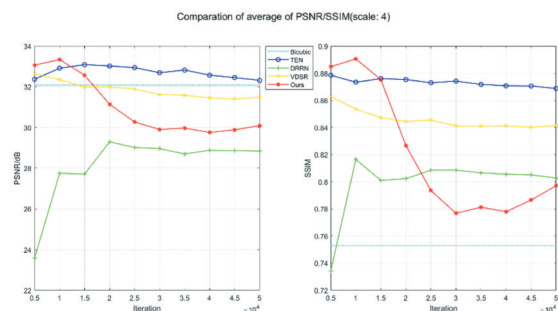


图 2 各模型对 4 倍重建的性能测试曲线

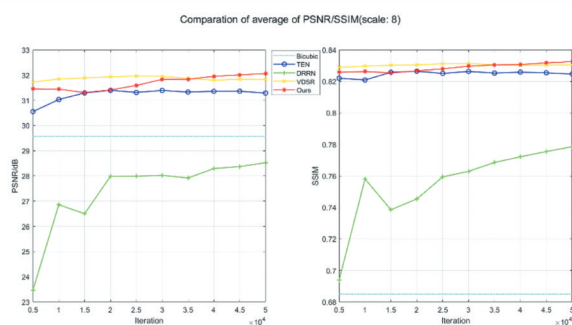


图 3 各模型对 8 倍重建的性能测试曲线

图 4 展示了 LTIR 测试集中 13 号图像的局部重建效果。可以看出,对于 8 倍重建测试,本文模型的重建效果与其他较优模型相比性能相当;在 4 倍重建测试中,本文模型重建能力不太理想,说明参数量对网路模型的表达能力影响明显。

3 结语

本文提出一种级联卷积神经网路的红外图像超分辨率重建模型。该模型利用典型模型的优势,通过将整体重建任务拆分为轮廓特征重建和纹理特征重建两部分,在 VDSR 模型的基础上对网路结构进行了改进,实现了网路结构的优化。实验在低数据量训练条件下进行,本文模型能够在参数量较少的情况下实现 PSNR 和 SSIM 指标以及拟合能力的提升,达到了相应的实验目的。

表 2 各模型分别在 4 倍、8 倍重建下迭代训练 50000 次的平均 PSNR/SSIM 值

测试集	重建倍数	Bicubic	TEN	VDSR	DRRN	CCNSR (Ours)
LTIR14	4x	32.54dB/0.755	33.04dB/0.877	32.58dB/0.858	29.02dB/0.812	31.18dB/0.817
	8x	30.25dB/0.692	31.91dB/0.835	32.65dB/0.843	28.90dB/0.793	32.95dB/0.846
FLIR10	4x	31.63dB/0.751	31.31dB/0.857	29.98dB/0.818	28.60dB/0.790	28.58dB/0.769
	8x	28.89dB/0.678	30.42dB/0.811	30.67dB/0.813	28.00dB/0.758	30.81dB/0.814

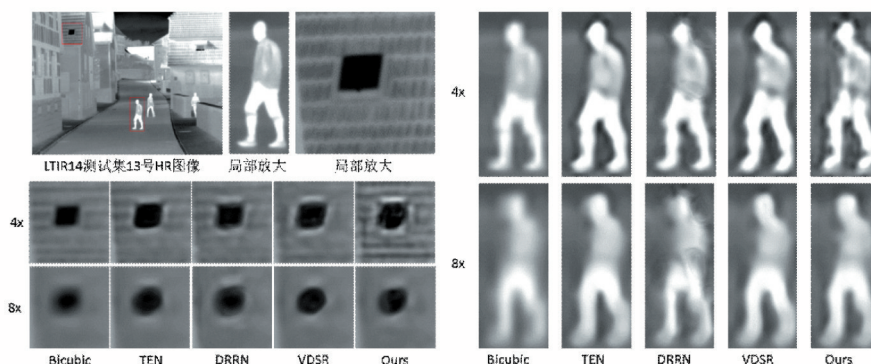


图4 重建效果展示

参考文献:

- [1]凡遵林,管乃洋,王之元,等. 红外图像质量的提升技术综述[J]. 红外技术,2019,41(10):941-946.
- [2]胡长胜,詹曙,吴从中. 基于深度特征学习的图像超分辨率重建[J]. 自动化学报,2017,43(5):814-821.
- [3]DONG C, LOY C C, HE K M, et al. Image super-resolution using deep convolutional networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(2):295-307.
- [4]KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016:1646-1654.
- [5]TAI Y, YANG J, LIU X. Image super-resolution via deep recursive residual network[C]//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017:2790-2798.
- [6]LEDIG C, THEIS L, HUSZÁR F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017:105-114.
- [7]CHOI Y, KIM N, HWANG S, et al. Thermal image enhancement using convolutional neural network[C]//Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway: IEEE Press, 2016:223-230.
- [8]FAN Z L, BI D Y, XIONG L, et al. Dim infrared image enhancement based on convolutional neural network[J]. Neurocomputing, 2018, 272:396-404.
- [9]LEE K, LEE J, LEE J, et al. Brightness-based convolutional neural network for thermal image enhancement[J]. IEEE Access, 2017, 5:26867-26879.
- [10]RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[OL]. [2020-06-22]. <http://arxiv.org/pdf/1505.04597.pdf>.
- [11]LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution[C]//Proceedings of the 2017 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. Piscataway: IEEE Press, 2017:1132-1140.
- [12]KINGMA D, BA J L. Adam: a method for stochastic optimization[OL]. [2019-05-24]. <http://arxiv.org/pdf/1412.6980.pdf>.
- [13]HUANG J B, SINGH A, AHUJA N, et al. Single image super-resolution from transformed self-exemplars[C]//Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2015:07-12, 5197-5206.
- [14]BERG A, AHLBERG J, FELSBERG M. A thermal object tracking benchmark[C]//Proceedings of the 2015 12th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). Piscataway: IEEE Press, 2015:1-6.
- [15]PORTMANN J, LYNEN S, CHLI M, et al. People detection and tracking from aerial thermal views[C]//Proceedings of the 2014 IEEE International Conference on Robotics and Automation (ICRA). Piscataway: IEEE Press, 2014:1794-1800.
- [16]WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity[J]. IEEE Transactions on Image Processing, 2004, 13(4):600-612.

作者简介:

甄有恒(1992-),男,宁夏银川人,硕士研究生,研究方向为机器智能

收稿日期:2021-01-05 修稿日期:2021-03-26

Infrared Image Super Resolution Based on Cascaded Convolutional Networks

ZHEN Youheng^{1,2}

(1. College of Computer Science, Sichuan University, Chengdu 610065;

2. No.95876 Force of PLA, Zhangye)

Abstract:

The methods of Infrared Image Super-Resolution based on convolutional neural networks still have the difficulty of detail feature extraction. For this issue, a two-cascaded convolutional neural networks with residual learning in a small amount of dataset is proposed for feature reconstruction step by step. Firstly, we use an encoder-decoder sub-network to fast extract the edge feature; Secondly, the texture feature is extract further by next recursion residual sub-network; Finally, a global skip-connection cause the super-resolution end to end. The experimental results show that the proposed algorithm has obvious performance compared with other models in fitting.

Keywords:

Convolutional Neural Networks; Infrared Image; Image Super-Resolution

(上接第 5 页)

Summary of Research on Traffic Sign Recognition Based on Deep Learning

YANG Lichuan

(College of Computer Science, Sichuan University, Chengdu 610065)

Abstract:

Traffic sign recognition plays an important role in assisted driving and unmanned driving. In this field, a variety of technical methods flourish, all of which are to process and recognize the obtained logo images. The development of deep learning algorithms such as convolutional neural networks and candidate regions provides new ideas for the recognition technology of traffic signs, and these works have an important role in the recognition of traffic signs. This paper summarizes these works, expounds the current status and results of domestic and foreign research on traffic sign recognition based on deep learning, analyzes the advantages of existing technologies and areas for improvement, and finally prospects the development of traffic sign recognition.

Keywords:

Deep Learning; Neural Network; Traffic Signs