

重庆理工大学学报(自然科学)

Journal of Chongqing University of Technology(Natural Science)

ISSN 1674-8425,CN 50-1205/T

《重庆理工大学学报(自然科学)》网络首发论文

题目: 训练集对遥感图像超分辨率下自动目标识别影响的研究
作者: 王艳, 李昂, 王晟全
收稿日期: 2020-5-12
网络首发日期: 2020-09-15
引用格式: 王艳, 李昂, 王晟全. 训练集对遥感图像超分辨率下自动目标识别影响的研究. 重庆理工大学学报(自然科学).
<https://kns.cnki.net/kcms/detail/50.1205.T.20200915.1054.004.html>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

训练集对遥感图像超分辨率下自动 目标识别影响的研究¹

王艳¹, 李昂^{1,2}, 王晟全¹

(1. 南京理工大学紫金学院 电子工程与光电技术学院 江苏 南京 210023

2. 南京邮电大学通信与信息工程学院, 江苏 南京 210003)

摘要：图像超分辨率（SISR）是将低分辨率图像映射到高分辨率图像的过程。作为一种提高卫星图像空间分辨率的方法，它在遥感图像中具有广泛的应用。本文采用深度学习算法：首先通过超分辨复原生成对抗网络（SRGAN）实现 SISR，在包括测试船舶本体的多个训练集中进行训练；然后采用 AlexNet 网络进行分类检测；最后采用 RetinaNet 网络进行目标识别。实验表明，和不包含测试本体的训练集相比，包含测试本体的训练集，超分辨率下的目标识别精确率最高。

关键词：图像超分辨率；遥感图像；SRGAN；目标识别；深度学习

中图法分类号：TP39

文献标志码 A

Research on the influence of training set on automatic target recognition under super-resolution of remote sensing images

WANG Yan¹, LI Ang^{1,2}, WANG Shengquan¹

(1.School of Electronic Engineering and Optoelectronic Technology, Nanjing University of Science and Technology ZiJin College, Nanjing City, 210023 China

2.School of Telecommunications & Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing City, 210023 China)

Abstract: Image super-resolution (SISR) is the process of mapping low-resolution image to high-resolution image. As a method to improve the spatial resolution of satellite image, it is widely used in remote sensing image. In this paper, deep learning algorithm is used: firstly, SISR is realized by srgan, which is trained in several training sets including testing ship ontology; secondly, alexnet network is used for classification detection; finally, retinanet network is used for target recognition. Experimental results show that, compared with the training set without test ontology, the training set with test ontology has the highest accuracy of target recognition under super-resolution.

Keywords: image super-resolution; satellite image; SRGAN; target detection; deep learning

收稿日期：2020-5-12

基金项目：江苏省高校自然科学基金面上项目 (18KJD510004), 江苏省普通高校学术学位研究生科研创新计划项目 (KYLX160661), 江苏省高等学校大学生创新创业训练计划项目 (201913654002Y)。

作者简介：王艳, 女, 硕士, 讲师, 主要从事数字图像处理研究, E-mail: kintywy@163.com。

1 概述

遥感成像技术的迅猛发展催生了大量遥感图像的产生。遥感图像在国民经济建设的众多领域中有着特殊的优势,例如战场态势监控、城市交通,以及矿物、陆地植被、惰性气体和人工材料的探测与识别等^[1]。自动目标识别(ATR)是计算机自动完成目标捕捉并分类的过程^[2-4]。该过程涉及图像滤波、图像增强、图像分割、图像变换、特征提取、特征选择、图像匹配等多个方向的专业知识^[5-6]。本文以船舶为目标,在遥感图像中实现船舶的自动目标识别。遥感图像中的船舶属于小目标,对遥感图像进行小目标识别并非容易的事情,原因在于遥感图像的空间分辨率较低,针对普通图像的自动识别算法无法直接应用于遥感图像。因此在对船舶进行识别之前,需要完成图像超分辨率(SISR),即将低分辨率图像映射到高分辨率图像^[7-9],该技术可以在不改变物理成像设备的限制下,获取更高分辨率的图像,在遥感图像处理得到了应用^[10-12],可以提高遥感图像的空间分辨率、自动识别的准确率。下面分别介绍 ATR 和 SISR 的发展现状。

随着深度学习的巨大进展,基于卷积神经网络(Convolution Neural Network,CNN)^[13]的自动目标识别算法取得巨大进步。相比于传统方法,目标识别性能得到巨大提升,特别是在复杂场景的自动目标识别方面。Wagner 等^[14]利用 CNN 提取数据特征,然后将得到的特征数据送入到可训练的 SVM 分类器中进行分类。文献[15]使用迁移学习解决了深度学习应用于 SAR 目标识别时数据不足的问题。Chen 等人^[16]用卷积层替换 CNN 中的全连接层,降低了过拟合风险。文献^[17]通过在损失函数中添加欧式距离限制(Euclidean Distance Restriction),使得类与类之间的特征更具有可分性。

SISR 是低分辨率图像映射到高分辨率图像的过程,比任何缩放算法都具有更高保真度,如图 1 所示。SISR 可以有效提高图像的空间分辨率,因此可以减少对更大、更昂贵卫星摄像机的需求。SISR 的核心是尝试将低分辨率图像映射到更高分辨率分辨率图像基于图像中的像素,如图 2 所示。SISR 分为三类:基于插值、基于模型和基于判别学习的方法,其中基于判别学习的方法最为主流^[18]。基于判别学习的算法利用 CNN 对于大量的图像训练数据集进行训练,学习 LR 图像和 HR 图像之间的映射关系,从而能够对于输入 LR 图像进行有效超分辨率重建。Kawulok 等人^[19]测试了用于训练 SR 网络的下采样方法,使用 DIV2K 图像测试了 SRResNet 和 FSRCNN,通过不同的下采样方法显示了不同的图像质量得分。Kim 等^[20]提出一个深度的超分辨率模型 VDSR (Very deep super-resolution),采用了 20 个卷积层。Kim 等通过实验验证了采用更深网络

模型能够带来更好的重建效果,但也会对收敛速度产生很大影响。



图 1 SR 提高空间分辨率的效果

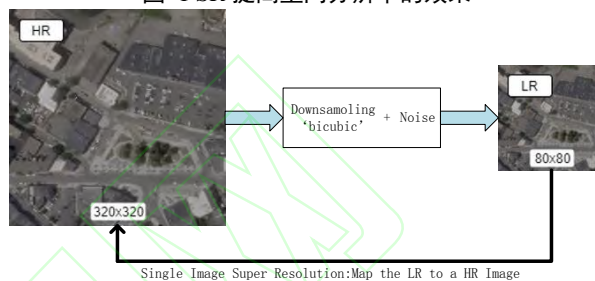


图 2 SISR 概述

本文采用超分辨率复原生对抗网络(SRGAN)在包括测试船舶本体的多个训练集中进行训练,多个训练集包括农业、城市、杂乱场景、石油、港口和船舶本体,然后采用 AlexNet 网络进行分类检测,最后采用 RetinaNet 网络进行目标识别。本文主要研究了训练集对 SISR、超分辨率下目标识别的影响,目前相关研究较少。实验表明,和不包含测试本体的训练集相比,包含测试本体的训练集, SISR 后的图像更加真实,超分辨率下的目标识别准确率更高。

2 超分辨率下目标自动识别算法

2.1 SRGAN 网络

SRGAN 的目的在于将一个低分辨率的图片转化为一个高分辨率的图片^[21]。SRGAN 是基于 GAN 方法进行训练,包括一个生成器和一个判别器,判别器的主体使用 VGG19,生成器是一连串的 Residual block 连接,同时在模型后部也加入了 subpixel 模块,借鉴了文献[22]改成中的 Subpixel Network 思想,让图片在最后的网络层增加分辨率,提升分辨率的同时,减少计算资源的消耗。

将均方误差(MSE)的感知损失 l^{SR} 表示为内容损失 l_X^{SR} 和对抗性损失成分 l_{Gen}^{SR} 的加权和,如式 1 所示。

$$l^{SR} = l_X^{SR} + 10^{-3} l_{Gen}^{SR} \quad (1)$$

像素方式的 MSE 损失 l_{MSE}^{SR} 计算方法如式 2 所示:

$$l_{MSE}^{SR} = \frac{1}{r^2 WH} \sum_{x=1}^{rW} \sum_{y=1}^{rH} (I_{x,y}^{HR} - G_{\theta_G}(I_{x,y}^{LR}))^2 \quad (2)$$

将 VGG 损函数失 $l_{VGG/i,j}^{SR}$ 定义为重建图像 $G_{\theta_G}(I_{x,y}^{LR})$ 特征表示与参考图像 I^{HR} 之间的欧氏距离，如式 3 所示：

$$l_{VGG/i,j}^{SR} = \frac{1}{W_{i,j} H_{i,j}} \sum_{x=1}^{W_{i,j} H_{i,j}} (\phi_{i,j}(I_{x,y}^{HR}) - \phi_{i,j}(G_{\theta_G}(I_{x,y}^{LR})))^2 \quad (3)$$

其中， $\phi_{i,j}$ 表示从 VGG-19 网络的第 j 层卷积(激活后)的第 i 个最大池化层之前获取的特征图； $W_{i,j}$ 和 $H_{i,j}$ 是 VGG 网络中对应特征图的维数。

生成损失 l_{GEN}^{SR} 是基于所有训练样本上鉴别器

Generator Network

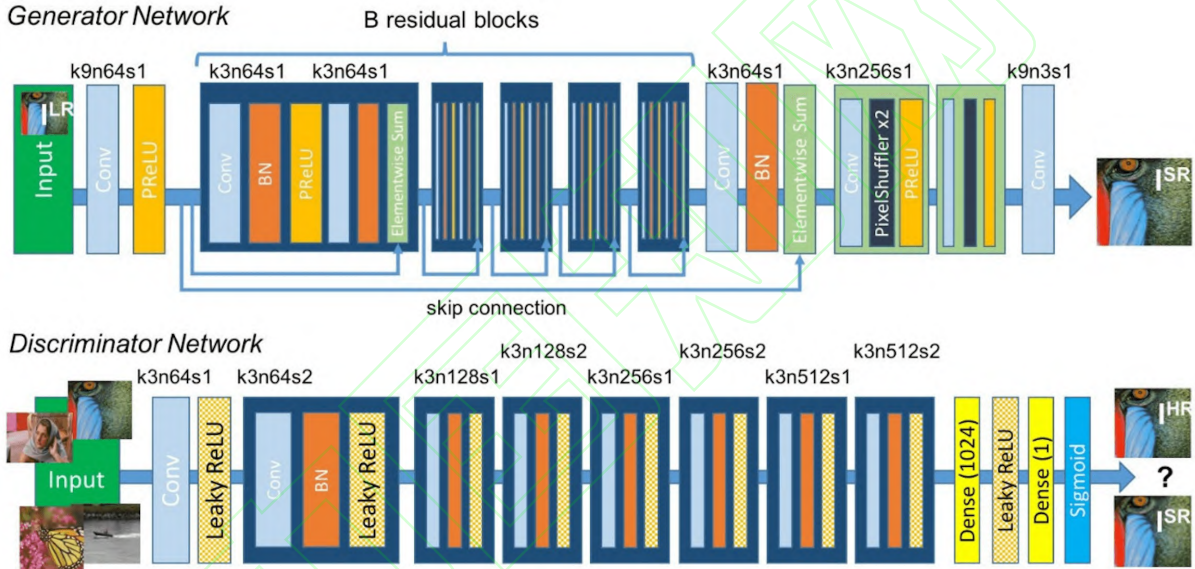


图 3 生成器和鉴别器网络的体系结构

2.2 分类网络

在后续的分类检测中，本文采用 AlexNet^[23] 网络作为分类网络的基本框架，前五层为卷积层，然后是全连接层，最后一层是分类层 softmax。传统网

$G_{\theta_G}(I^{SR})$ 的概率定义，如式 4 所示：

$$l_{GEN}^{SR} = \sum_{n=1}^n -\log(D_{\theta_D}(G_{\theta_G}(I^{SR}))) \quad (4)$$

$D_{\theta_D}(G_{\theta_G}(I^{LR}))$ 定义为重建图像 $G_{\theta_G}(I^{LR})$ 是自然 HR 图像的概率。为了更好的梯度表现，我们使用了 $-\log D_{\theta_D}(G_{\theta_G}(I^{LR}))$ 最小化生成损耗函数，而不是 $\log[1 - D_{\theta_D}(G_{\theta_G}(I^{LR}))]$ 。

SRGAN 结构图如图 3 所示，其中每个卷积层指示了相应的内核大小 (k)、特征图数量 (n) 和步幅 (s)。

络常用的激励函数 Sigmoid 和 tanh 函数被取代为修正线性单元 ReLU，同时为了减轻过拟合问题，采用了 dropout 方法，可以有效提高正确率，整体框架如图 4 所示。

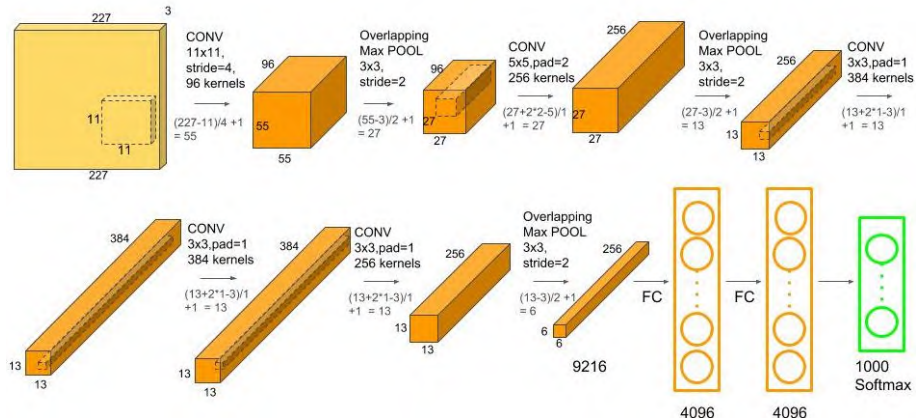


图 4 AlexNet

对于 80px 和 320px 的输入，使用深度均为 64，内核大小为 3×3 的 2D 卷积层，并激活 ReLU，然后再进行大小为 2×2 的 2D 最大池化层，直到获得 80×80 的特征空间。然后，我们改为使用深度为 32 的 Conv2D，然后使用大小为 2×2 的 Maxpool2D，以实现 38×38 的潜在空间。然后，应用速率为 0.30 的滤除层，展平要素，添加一个大小为 30 mm 的密集层 128 个，带 ReLU 激活，一个速率为 0.30 的退出层，一个 SoftMax 输出为 2 类，损失是分类交叉熵，优化器是默认的 Keras Adadelta。

2.3 目标识别网络

本文中采用了文献[24]基于 Focal Loss 的 RetinaNet 作为识别网络。RetinaNet 在识别领域有着重要的地位。Focal loss 是一种改进了的交叉熵 CE(cross-entropy) loss，但是它在原有的 CE loss 上乘了使得易检测目标对模型训练贡献削弱的指数式，从而成功地解决了在目标检测时，正负样本区域极不平衡、而目标检测 loss 易被大批量负样本影

响的问题，改进后的 bias: $b = -\alpha \log((1 - \pi) / \pi)$ ，其中 α 可以控制 bias 的值，使之能够根据实际的情况，取得相应的值，使其预测图像目标的准确性上升。

3 实验

3.1 训练集

为了避免数据集中的相机伪影变化太多，训练图像需使用单个长焦距镜的照片。为此，采用了 Planet 公司 0.8m Visual Skysat 卫星采集的图像。SkySat 卫星是美国 Planet 公司发展的高频成像对地观测小卫星星座，主要用于获取时序图像，并服务于高分辨率遥感大数据的应用。如图 5 所示，单个图像的可视化效果和 RGB 直方图，以及 5 种训练集（农业、城市、杂乱场景、石油和港口）中每种训练集的直方图平均值。

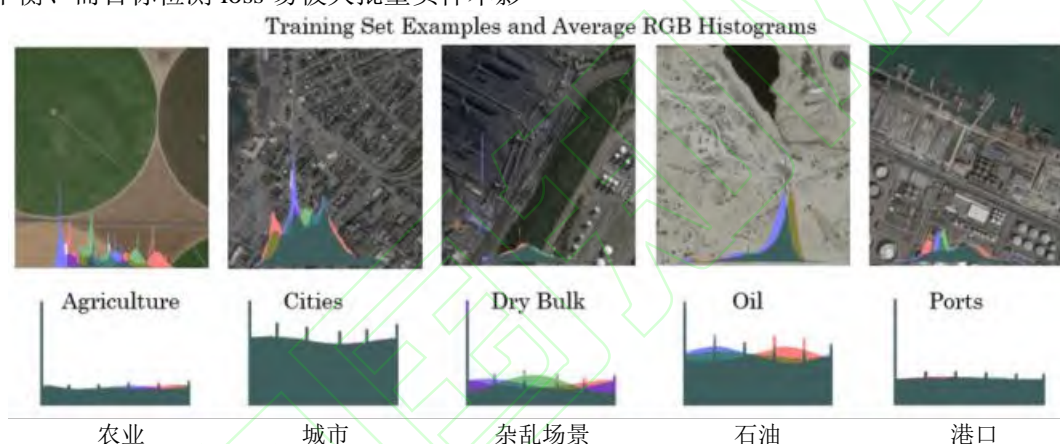
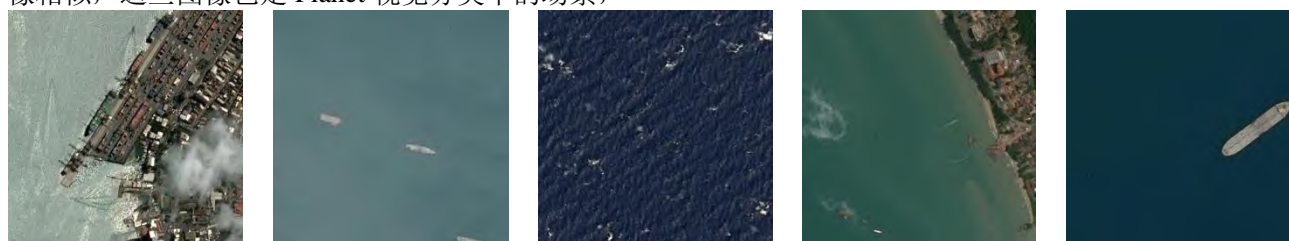


图 5 土地利用类别的示例图像和 RGB 直方图

3.2 测试集

测试集采用来自 Kaggle Shipsnet 竞赛的图像，示例图像如图 6 所示。这些图像来自 Planet 的 Planetscope 卫星（3.0m），其中包含加利福尼亚州的旧金山湾和圣佩德罗湾地区图像。与训练集的图像相似，这些图像也是 Planet 视觉分类中的场景，

这意味着它们也经过了锐化、正射和颜色校正的 RGB。但是，与训练集不同，测试集中的场景已被平铺为 80 px×80 px 的正方形。就数据集中的类别而言，只有全幅船被分类为船，所有其他图像均归类为“无船”。测试集包含 4000 张图像，其中 1000 张图像被分类为船。



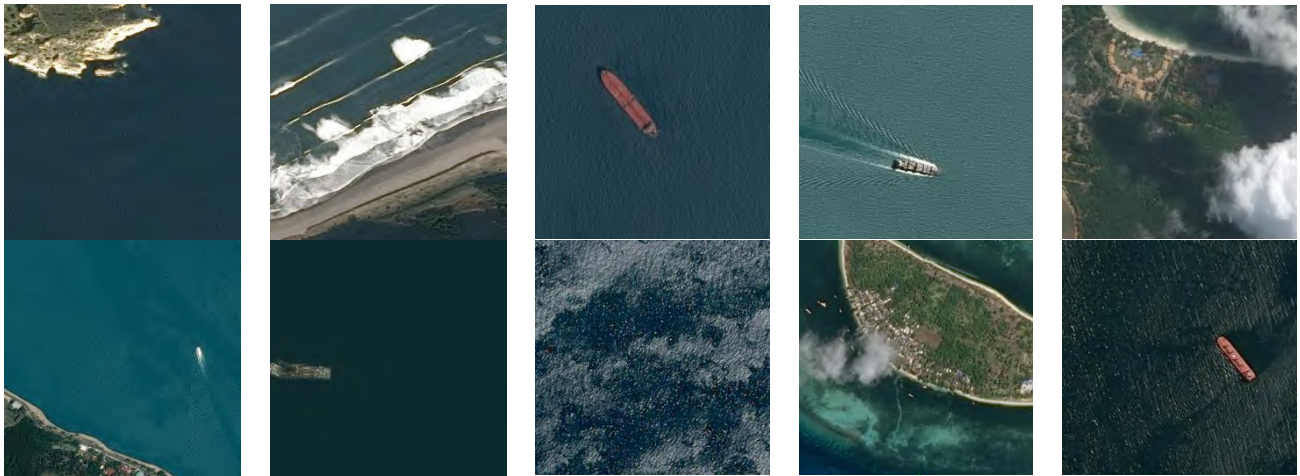


图 6 来自 Kaggle Shipsnet 竞赛的图像

3.3 模型训练

在农业、城市、杂乱场景、石油、港口和船舶本体这6个测试集训练了6个SRGAN,每个SRGAN均训练了10000次,每批次20个。目标检测和分类检测在 Intel i7-7700HQ 2.8 GHz 上运行,每组图片大约需要30 min。

3.4 SRGAN

使用所需分辨率的所有图像后,可以设置SRGAN网络,加载数据后,需要先采样以获得LR图像。再次使用双三次滤波器对图像进行下采样,以减少采样方法对结果的影响,通过GAN运行数据之前将-1和1之间的所有像素值标准化。

SRGAN在包括农业、城市、杂乱场景、石油、港口和船舶本体这6个训练集中进行训练。每个SRGAN运行的船舶图片如图7所示。对于船舶图

像,由农业、港口训练的SRGAN产生的图像色彩偏暗,由城市、杂乱场景训练的SRGAN产生的图像色彩偏亮,由石油训练的SRGAN在船舶的船体产生了黑色的人工痕迹,由船舶本体训练的SRGAN产生的图像色彩最亮、效果最好。

在训练了SRGAN之后,判别出一些流行的卫星图像样本的缩减样本,如: Xview (.3m), Pleiades (.5m), Quickbird (.65m), Triplesat (.8m)和 Ikonos (1m)。不同种类的训练集SRGAN对测试集产生的影响如表1所示,船舶本体和其他训练集相比,PSNR和SSIM这两个指标均表现最好。这表明针对本体图像需要进行定制训练。原因在于,超分辨复原的训练,是无监督学习,结果具有随机性,但是也与本身的图像显著性特征有关。

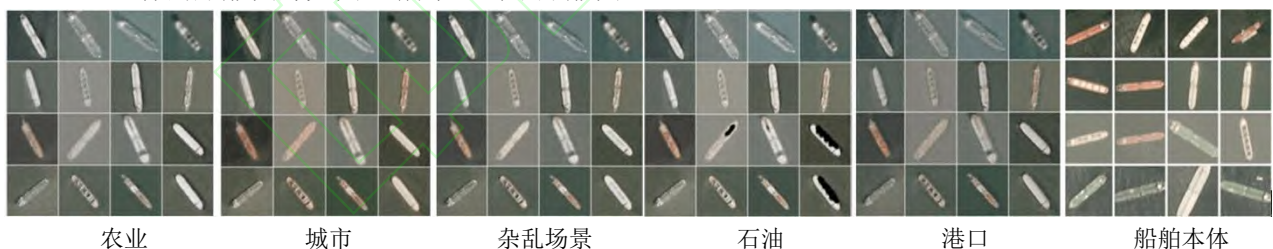


图 7 SRGAN 运行的船舶图片

表 1 SRGAN 分数应用于不同的图像

Dataset	Scoring	Agriculture	Cities	Dry bulk	Oil	Ports	Ships
Xview (.3m)	PSNR	17.011 623	16.981 220	17.307 997	16.801 039	17.162 341	17.170 23
	SSIM	0.491 469	0.477 119	0.451 492	0.483 934	0.477 334	0.493 614
Pleiades (.5m)	PSNR	22.770 306	19.560 364	21.841 543	21.544 979	18.182 035	22.987 76
	SSIM	0.625 712	0.413 790	0.541 247	0.562 504	0.293 370	0.642 325
Quickbird (.65m)	PSNR	22.087 340	18.617 457	18.647 200	17.486 228	17.244 123	22.638 25

Triplesat (.8m)	SSIM	0.583 851	0.371 042	0.451 841	0.463 415	0.270 718	0.576 446
	PSNR	20.925 616	18.843 711	20.599 660	20.514 242	17.674 832	21.275 64
	SSIM	0.471 851	0.423 649	0.504 717	0.509 565	0.342 879	0.510 532
Ikonos (1m)	PSNR	20.413 195	19.709 312	21.795 880	21.917 841	19.604 518	21.980 98
	SSIM	0.483 809	0.449 529	0.518 172	0.531 595	0.394 726	0.528 874

3.5 目标识别

在目标识别的过程中，采用了 3.2 节提到的 AlexNet 网络和 3.3 节提到的 RetinaNet 算法，图像的固定分辨率似乎并不能从增强中获益很多，并且显示出对基于尺寸的分辨率变化的敏感性。xView 对比例敏感度的一个示例是在较高的分辨率下，同一张图像可以错误地检测出汽车的建筑物，而在较低分辨率下，则可以将汽车误认为是游艇。总体而言，对于未更改的 xView 图像的基础模型，最佳

验证 mAP 为 300 个历时（3 天）后为 0.16，对应的训练 mAP 为 0.30。不同训练集下的识别评价指标如表 2 所示，其中 TP 是实际为正例、预测为正例的比例 FP 是实际为负例、预测为正例的比例，识别精确率 $Precision=TP/(TP+FP)$ 。从表 2 可见，在不同的训练集情况下 SRGAN 中有明显的差异，包含测试船舶本体的训练集，目标识别的精确率最高，达到了 98.79%，可以满足实际的识别应用需求。

表 2 不同训练集下的识别评价指标

Data Set	Agriculture	Cities	Dry Bulk	Oil	Ports	Ships
TP	93.27%	93.45%	95.37%	91.93%	92.34%	97.14%
FP	2.38%	1.67%	1.79%	2.12%	1.79%	1.19%
Precision	97.51%	98.24%	98.16%	97.75%	98.10%	98.79%

3.6 附加实验

本文算法对飞机目标进行识别和准确性的测试效果图如图 8 所示，图中绿色框表示标签位置、蓝色表示测试结果、红色表示误检的目标，准确率比较理想。但是在不同特征尺度下的飞机还是不能特别正确的标识，当飞机分辨率较低时，有一定的误

检率和漏检率。因此对遥感图像中的飞机进行识别之前，引入 SISR 算法是非常必要的，该算法可以将低分辨率图像映射到高分辨率图像。SRGAN 处理过的 google 遥感图像训练损失率比较理想，如图 9 所示。



图 8 对遥感机场的检测实验

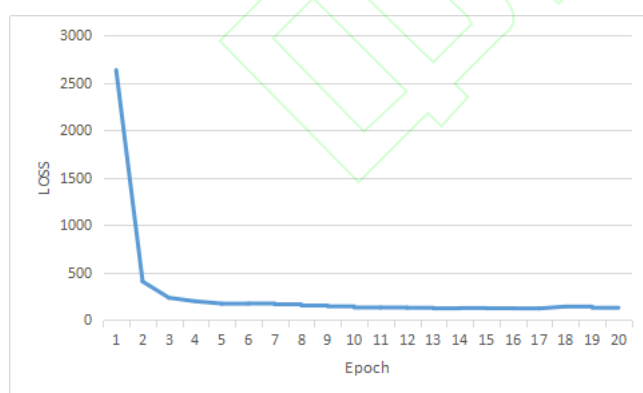


图 9 SRGAN 处理过的 google 遥感图像

4 结束语

当训练不同的图像集时，在 SRGAN 网络之间可以看到明显的差异。这突显了定制训练集需求的

重要性。在包括农业、城市、杂乱场景、石油、港口和船舶本体这 6 类训练集中进行 SRGAN 训练，然后采用 AlexNet 网络进行分类检测；最后采用 RetinaNet 网络进行目标识别。实验表明：在同一个船舶测试集中，受过测试本体图像训练的识别效果最好，精确率最高。

参考文献:

- [1] 罗峰. 基于超分辨率迁移学习的遥感图像车辆检测[D]. 厦门: 厦门大学, 2017.
- [2] BHANU B, DUDGEON D E, ZELNIOE G, et al. Guest editorial introduction to the special issue on automatic target detection and recognition [J]. IEEE Transactions on Image Processing, 1997, 6(1):1-6.
- [3] 贺丰收, 何友, 刘淮钊, 等. 卷积神经网络在雷达自动目标识别中的研究进展 [J]. 电子与信息学报, 2020, 42(1):119-131.
- [4] 戴玉超, 张静, 何明一, 等. 深度残差网络的多光谱遥感图

- 像显著目标检测[J].测绘学报,2018,47(6):873-881.
- [5] 刘士建,金璐.自动目标识别算法发展综述[J].电光与控制,2016,23(10):1-7.
- [6] 刘扬,付征叶,郑逢斌.高分辨率遥感影像目标分类与识别研究进展[J].地球信息科学学报,2015,17(9):1080-1091.
- [7] 李云红,王珍,张凯兵,等.基于学习的图像超分辨重建方法综述[J].计算机工程与应用.2018,54(15):13-21.
- [8] ANWAR S, KHAN S, BARNES N. A deep journey into super-resolution: A survey[J/OL]. (2019-07-17)[2019-11-30]. <https://arxiv.org/abs/1904.07523?context=cs>.
- [9] WANG Z, CHEN J, HOI S C H. Deep learning for image super-resolution: A survey[J/OL]. (2019-02-16)[2019-11-30]. <https://arxiv.org/abs/1902.06068>.
- [10] JI H, GAO Z, MEI T, et al. Vehicle detection in remote sensing images leveraging on simultaneous super-resolution[J]. IEEE Geoscience and Remote Sensing Letters, 2019. DOI: 10.1109/LGRS.2019.2930308.
- [11] SHERMEYER J, VAN ET TEN A. The effects of super-resolution on object detection performance in satellite imagery[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. [S.l.]: IEEE, 2019.
- [12] 杨超,杨斌,黄国玉. 基于多光谱图像超分辨率处理的遥感图像融合[J]. 激光与光电子学进展, 2016, 53(2): 88-95.
- [13] 王艳,李昂,王晟全,王建宇,胡家俊. 基于 CNN 与传统 S 盒结合的图像加密算法[J/OL]. 电光与控制 :1-9[2020-05-20].<http://kns.cnki.net/kcms/detail/41.1227.tn.20200526.1025.002.html>.
- [14] WAGNER S A. SAR ATR by a combination of convolutional neural network and support vector machines[J]. IEEE Transactions on Aerospace & Electronic Systems. 2017, 52(6): 2861-2872.
- [15] ZEILER M D, FERGUS R. Visualizing and understanding convolutional networks [J]. 2013, 8689:818-833.
- [16] CHEN S, WANG H, XU F, et al. Target Classification Using the Deep Convolutional Networks for SAR Images[J]. IEEE Transactions on Geoscience & Remote Sensing 2016, 54(8):4806-4817.
- [17] DENG S, DU L, LI C, et al. SAR Automatic Target Recognition Based on Euclidean Distance Restricted Autoencoder[J]. IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing, 2017,99:1-11.
- [18] 史振威,雷森.图像超分辨重建算法综述[J].数据采集与处理,2020,35(1):1-20.
- [19] MICHAL K, SZYMON P, KRZYSZTOF H, et al. On training deep networks for satellite image super-resolution. arXiv preprint arXiv:1906.06697, 2019.
- [20] KIM J, LEE K J, LEE M K. Accurate image super-resolution using very deep convolutional networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2016: 1646-1654.
- [21] LEDIG C, THEIS L, HUSZÁR F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 4681-4690.
- [22] TANG Z H, WANG J N. MOANA: An Online Learned Adaptive Appearance Model for Robust Multiple Object Tracking in 3D[J]. IEEE Access, 2019.
- [23] KRIZHEVSKY A, SUTSKEVER I, HINTON G. ImageNet Classification with Deep Convolutional Neural Networks[C]// NIPS. Curran Associates Inc. 2012.
- [24] 李昂,王晟全,郑宝玉,等.改进的用于军用车目标检测的 RetinaNet[J/OL]. 电光与控制 :1-8 [2019-12-05].<http://kns.cnki.net/kcms/detail/41.1227.TN.20191108.0934.002.html>.