

基于渐进式生成对抗网络的舰船红外图像仿真

谢晓方¹, 刘厚君¹, 张龙杰¹, 孙涛¹, 张龙云²

(1. 海军航空大学, 山东 烟台 264001; 2. 山东大学, 山东 济南 250000)

摘要: 红外图像仿真在红外导引头设计、仿真训练中起到十分关键的作用。针对如何生成高分辨率、视觉特征可控的红外图像, 提出了一种基于渐进式生成对抗网络的红外图像仿真方法。本文利用舰船模型的红外图像数据集训练了图像合成网络, 输入随机特征向量, 输出高分辨率的红外仿真图像; 设计了图像编码网络, 实现红外图像到特征向量的转换; 利用 Logistic 回归方法, 在特征向量域找到了控制红外图像角度特征的方向向量, 并据此生成了不同角度的舰船模型仿真图像; 最后通过均值哈希算法和平均结构相似性算法来定量评价仿真图像和真实图像的差异, 实验结果表明仿真的红外图像和真实图像的相似度很高, 可以为真实舰船的可控化红外图像仿真提供参考。

关键词: 渐进式生成对抗网络; 图像编码网络; 图像特征向量; 样式控制; 舰船红外仿真

中图分类号: TP391 **文献标识码:** A **DOI:** 10.3969/j.issn.1001-5078.2021.04.012

Infrared image simulation for warship based on progressive generative adversarial network

XIE Xiao-fang¹, LIU Hou-jun¹, ZHANG Long-jie¹, Sun Tao¹, Zhang Long-yun²

(1. Naval Aeronautical University, Yantai 264001, China; 2. Shandong University, Ji'nan 250000, China)

Abstract: Infrared image simulation plays an important role in the design and training of infrared seeker. Aiming at how to simulate infrared images with high resolution and controllable visual features, a method of infrared image simulation based on progressive generative adversarial network is proposed. In this paper, the infrared images dataset of warship model is used to train the image synthesis network, only random feature vector is necessary for the generation of high resolution infrared simulated image. An image encoder network is designed to convert infrared images to feature vectors. By the logistic regression method, the direction vector which controls the angle feature of infrared images is found in the feature vector domain, and simulated images of warship model from different angles are generated accordingly. AHash algorithm and MSSIM algorithm are used to quantitatively evaluate the difference between simulated and real images. The experimental results show that the similarity between the simulated infrared images and the real ones is very high, and it can be used as a reference for the controllable simulation of real warship infrared images.

Keywords: progressive generative adversarial network; image encoder network; image feature vector; style control; infrared simulation of warship

基金项目: 国家自然科学基金项目 (No. 51809156); 中国博士后科学基金项目 (No. 2016M600537); 教育部产学研合作协同育人项目 (No. 201902295036) 资助。

作者简介: 刘厚君 (1996-), 男, 硕士, 主要从事红外目标仿真方面的研究工作。E-mail: liuhoujun_lhj@163.com

通讯作者: 谢晓方 (1962-), 男, 教授, 博士生导师, 硕士, 主要从事武器系统建模与仿真方面的研究工作。

Email: Xiexf208@sina.com

收稿日期: 2020-07-07

1 引言

红外成像具备全天候、高隐蔽、抗云雾等优势,在反舰导弹制导过程中,红外制导作为末段制导的主要方式,发挥着关键性的作用^[1-2]。利用计算机模拟生成舰船的红外图像是红外导引头视景仿真领域的重要研究课题,在红外导引头设计、仿真训练等领域有广泛的应用前景。

近年来,在深度学习理论的推动下,涌现出许多神经网络模型,其中生成对抗网络(GAN)由于具有较好的特征自动提取与图像重构能力,已经广泛应用于图像生成、风格迁移、降噪修复等领域,文献[3]提出了深度卷积对抗网络(DCGAN),在 LSUN、Faces 和 ImageNet-1k 数据集实现了 32×32 分辨率的图像生成。文献[4]提出了一种基于条件深度卷积对抗生成网络(C-DCGAN)的红外目标仿真框架,在 DCGAN 基础上引入了条件特征码,特征码的每一维度对应一类目标的红外图像,通过改变特征码来生成不同类别的红外目标图像;文献[5]针对目前多数 GAN 模型生成图像分辨率过小的问题,提出了一种渐进式生成对抗网络 proGAN,生成网络采用了动态增长的网络结构,由低分辨率到高分辨率逐渐学习图像特征,最终可以生成高分辨率的仿真图像。

较于传统的计算机建模仿真^[6-10]的方式,利用 GAN 技术生成的图像在纹理细节方面表现的更加真实,但存在着一个关键性的问题:无法建立生成网络中输入与输出之间准确的映射关系,不能通过调整网络的输入对生成图像的视觉特征进行准确的控制,虽然生成的图像真实度很高,但无法直接应用于红外视景仿真系统中去。因此,利用 GAN 技术进行可控化的红外图像生成是一项值得研究的课题。

本文基于改进的渐进式生成对抗网络模型,提出了一种可控化的红外图像仿真方法:通过调整图像特征向量,对仿真图像的视觉特征进行定量的控制。仿真实验实现了舰船角度可控的红外图像生成,通过 aHash 算法和 MSSIM 算法检测,仿真图像与真实图像相似度较高,结果表明提出的仿真方法可行。本文方法与传统的计算机建模仿真的方法相比,仿真图像更真实、纹理细节表现更丰富;与传统的 GAN 图像生成相比,图像的合成过程更加可控。可以为真实舰船的红外图像仿真

提供参考。

2 改进的渐进式生成对抗网络

图 1(a)是文献[5]提出的渐进式生成网络,采用动态增长的网络结构,由低分辨率到高分辨率逐级合成图像,最终可以实现高分辨率的图像生成。但是该网络直接将随机向量 z 作为生成网络的输入,属性不加限定,无法有效地控制最终生成图像的视觉特征。

针对上述问题,图 1(b)对原有生成网络的架构进行了改进,改进后的生成网络由映射网络和合成网络两部分组成。

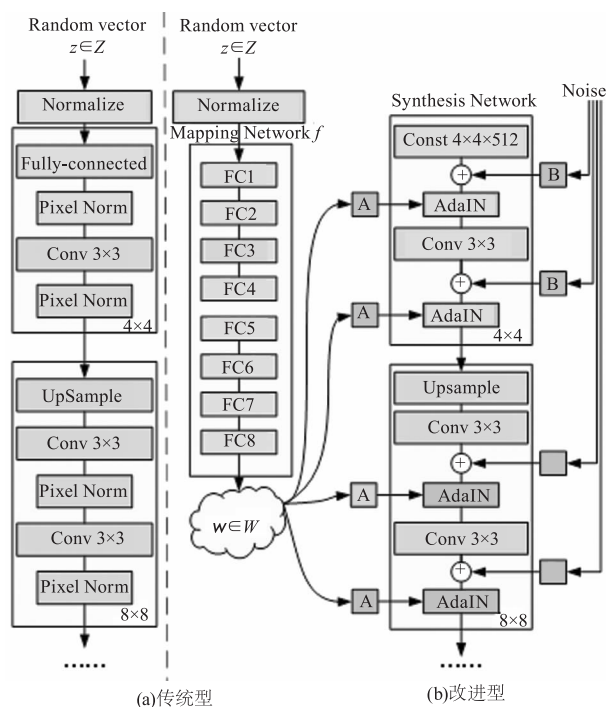


图 1 渐进式生成网络

Fig. 1 Progressive generator network

2.1 映射网络

映射网络由 8 个全连接层(FC1-FC8)组成,其目标是对随机向量 z 进行特征解耦。改进型的渐进式生成网络与传统型相比,不是直接将随机向量 $z \in Z$ 作为生成网络的输入,而是先通过映射网络将其映射为一个中间向量 $w, w \in W$ 。如果单纯依靠随机向量 z 来控制生成图像的视觉特征,能力十分有限,输入的随机向量 z 与输出图像的特征之间的映射往往是非线性的关系:稍微调整 z 某一维度的值,会影响生成图像的多个属性,出现特征纠缠,如图 2(a)所示;但经过映射网络的非线性变换: $f: Z \rightarrow W$ 后,实现了对 z 的特征解耦,输出的中间向量

w 控制图像特征生成能力更加线性,如图 2(b)所示,由于 w 与图像特征存在着线性的映射关系,将其命名为特征向量。

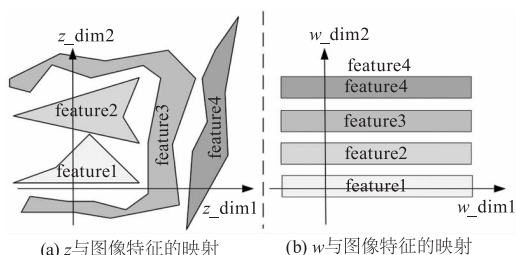


图2 特征解耦

Fig. 2 Decoupling of features

2.2 合成网络

如图 1(b)所示,合成网络不是直接将随机向量 z 直接作为生成网络的输入,而是先通过映射网络将图 z 特征解耦为图像特征向量 w ,通过仿射变换生成一种风格,然后通过自适应实例归一化,来控制该层级所表示的视觉特征。相比于图 1(a),合成网络新增加了仿射变换 A(Affine transfer)和噪声缩放通道 B(Scale),另外用自适应归一化模块 AdaIN 替代了原来的像素归一化 Pixel Norm 模块。通过加入这三个模块,合成网络可以实现对图像的样式控制,实现方式如图 3 所示,具体可分为五步。

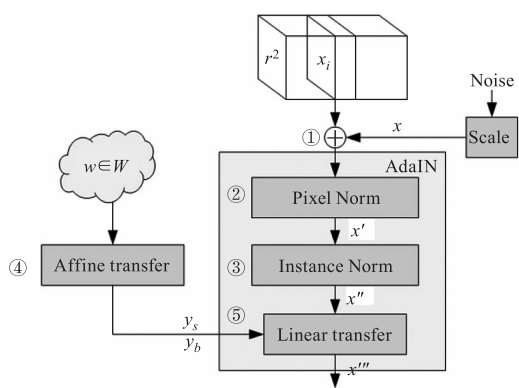


图3 图像样式可控化实现

Fig. 3 Realization of image style controlled

2.2.1 混合噪声

$$x = x + \alpha \times \text{noise} \quad (1)$$

其中, x 表示卷积输出; α 为比例缩放因子; noise 表示噪声。

引入经过比例缩放(Scale)的噪声 noise 与卷积层的输出(Conv output x)进行混合,可以微调对应的分辨率层级上图像的视觉特征,生成的红外舰船图像往往更加逼真与多样。

2.2.2 像素归一化(Pixel Norm)

$$x' = \frac{x}{\sqrt{\frac{1}{N} \sum_{i=0}^{N-1} x_i^2 + \varepsilon}} \quad (2)$$

其中, x 表示叠加噪声的卷积输出; x_i 是 x 第 i 个像素点的值; N 是 x 像素点的个数, ε 是归一化的参数,防止分母为 0。

像素归一化是局部响应归一化(LRN)的变种,可以有效避免生成网络训练过程中的梯度爆炸现象。

2.2.3 实例归一化(Instance Norm)

$$x'' = \frac{x' - \mu(x')}{\sqrt{\sigma(x')^2 + \varepsilon}} \quad (3)$$

其中, x' 是步骤②的输出; $\mu(x')$ 和 $\sigma(x')$ 是 x' 中所有像素点的均值和标准差; ε 是归一化的参数,防止分母为 0。

实例归一化是对特征图像的宽和高两个维度进行归一化,加速模型的收敛,并且可以保持样例之间的相对独立性。

2.2.4 仿射变换(Affine transfer)

$$y = \text{Dense}(w) \quad (4)$$

其中, w 是经过映射网络特征解耦的中间向量; $\text{Dense}(\cdot)$ 表示全连接操作; 输出 y 为样式因子,由缩放因子 y_s 和偏置因子 y_b 共同组成。

2.2.5 线性变换(Linear transfer)

$$x''' = y_s \cdot x'' + y_b \quad (5)$$

其中, y_s 和 y_b 是经过步骤④得到的缩放因子和偏置因子; x'' 是经过步骤③的输出; x''' 是自适应归一化模块(AdaIN)的最终输出。

合成网络之所以能实现对最终生成图像的样式控制,主要是生成图像的关键信息仅由上采样层(Upsample)和卷积层(Conv 3×3)来决定,卷积输出的归一化(Pixel Norm、Instance Norm)操作抹去了对图片局部信息的可见性,而 w 经过仿射变换成 y_s 与 y_b ,与标准化的卷积输出进行加权求和,仅仅影响图片的全局信息。

2.3 损失函数

对 GAN 网络进行训练,影响 G 和 D 训练效果的关键在于损失函数的选取,这里选择的是 logistic 损失函数。

2.3.1 生成网络的损失函数

生成网络的损失函数是非饱和 logistic 损失:

$$\text{Loss}_G = \log(1 + e^{-D(G(w))}) \quad (6)$$

其中, $D(G(w))$ 表示判别网络对生成图像的评分, 由于生成网络的目标只是生成足够真实的图像, 因此损失函数只需对生成图像 $G(w)$ 计算损失。

2.3.2 判别网络的损失函数

判别网络的损失函数是含梯度惩罚项的 logistic 损失:

$$\text{Loss}_D = \log(1 + e^{-D(X)}) + \log(1 + e^{D(G(w))}) + \frac{\gamma_1}{2} \sum \nabla_{T_{\text{real}}}^2 + \frac{\gamma_2}{2} \sum \nabla_{T_{\text{fake}}}^2 \quad (7)$$

$\log(1 + e^{-D(X)}) + \log(1 + e^{D(G(w))})$ 为判别网络的 logistic 损失, 其中, $D(X)$ 和 $D(G(w))$ 分别表示判别网络对真实图像和生成图像的评分, 由于判别网络需要鉴别图像的“真伪”, 因此损失函数需同时对真实图像 X 和生成图像 $G(w)$ 计算损失。

$\frac{\gamma_1}{2} \sum \nabla_{T_{\text{real}}}^2$ 和 $\frac{\gamma_2}{2} \sum \nabla_{T_{\text{fake}}}^2$ 分别为真实图像 X 和生成图像 $G(w)$ 的梯度惩罚项, 其中, $\nabla_{T_{\text{real}}}$ 和 $\nabla_{T_{\text{fake}}}$ 分别为 $D(X)$ 和 $D(G(w))$ 的梯度值, γ_1 和 γ_2 为梯度惩罚项的系数。引入梯度惩罚项主要是为了解决判别网络训练不收敛的问题。

3 特征可控的红外图像生成

首先通过图像编码网络找到红外图像在特征向量域 W 对应的特征向量 w , 然后利用 logistic 回归方法在 W 找到控制红外图像视觉特征的方向向量 e , 沿着方向向量 e 以不同的幅度 coeff 对 w 进行调整, 通过合成网络生成图像, 可以对图像的视觉特征进行控制, 实现红外图像的可控化生成。

3.1 图像编码网络架构

借鉴了双向对抗生成网络 (BiGAN)^[11] 的思想, 利用 Resnet50 模型^[12] 较好的图像特征提取能力, 选取其作为图像编码网络的基础架构, 然后加入几个转换层 (Conv2D: 卷积层; reshape: 形状重定; Locally_connected: 局部连接层; Permute: 维度置换), 构建了图像编码网络: 输入红外图像, 输出图像特征向量。网络结构如图 4 所示。

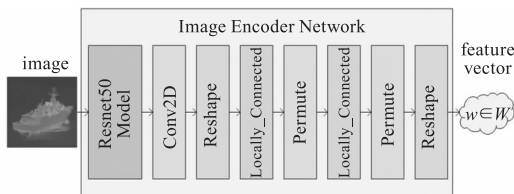


图 4 图像编码网络架构

Fig. 4 Structure of image encoder network

3.2 图像编码网络的训练

图像编码网络的训练分为模型的预训练和模型微调两个阶段, 如图 5 所示。

在预训练阶段, 将随机产生的图像特征向量 w 通过合成网络 (Synthesis network) 生成红外图像 $G(w)$, 以 $G(w)$ 和 w 分别作为编码网络的输入和输出, 训练得到一个预训练模型, 如图 5(a)。

预训练阶段仅仅是在生成图像域上的训练, 而具有实际应用价值的是能够找到真实红外图像的特征向量。因此在微调阶段, 选取真实的红外图像作为编码网络的输入, 如图 5(b) 所示, 针对特定的图像 w , 对预训练模型进行微调, 以便找到其在特征向量域 W 的准确位置 (target)。

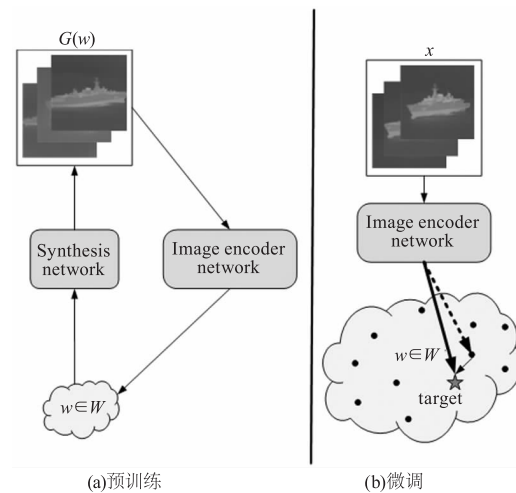


图 5 分阶段训练图像编码网络

Fig. 5 Partition training of image encoder network

图像编码网络训练的算法流程如下:

Algorithm: Train image encoder network

Input: real image to be encoded x ; initialized feature vector w ; synthesis network $G(\cdot)$.

Output: feature vector w' ; generated image $G(w')$

- 1 Initialize feature vector $w' = w$
- 2 **while** L is not converged **do**:
- 3 put w' into $G(\cdot)$ to generate $G(w')$
- 4 calculate perceptual loss between $G(w')$ and x :

$$L \leftarrow L_{\text{perceptual}}(G(w'), x) + \lambda_{l_2} || G(w') - x ||_2$$
- 5 apply SGD algorithm to update feature vector w' :

$$w' \leftarrow w' - \eta F(\nabla_{w'} L)$$
- 6 **end**

其中, $L_{\text{perceptual}}(G(w'), x)$ 为感知损失, 是真实图像 x

和生成的图像 $G(w')$ 经过 VGG16 模型^[13] 特征提取后两者的特征相似度 loss, 计算方式如下:

$$L_{\text{perceptual}}(G(w'), x) = \sum_{j=1}^4 \lambda_j \|F_j(G(w')) - F_j(x)\|_2 \quad (8)$$

其中, $F_i(\cdot)$ 表示 VGG16 中卷积层 conv1_1, conv1_2, conv3_2, conv4_2 输出的 feature map;

$\|G(w') - x\|_2$ 是 L2 正则项, 避免模型训练出现的过拟合, λ_{L2} 是正则项系数。

3.3 Logistic 回归求解方向向量

针对红外图像中的某一视觉特征, 探究其在特征向量域 W 的方向向量。

首先对采集的红外图像样本添加属性标签, 然后探究标签值的变化与对应特征向量分布变化的关系, 将这一变化关系用方向向量 e 表示, 其维度与特征向量 w 一致。对于 e 的求解, 有两种方法。

方法一是每次取两个样本 p 和 q , 用两者的特征向量的差值除以标签的差值, 得到一个方向向量:

$$e_i = \frac{w_p - w_q}{l_p - l_q}, l_p \neq l_q \quad (9)$$

然后多次重复这一过程, 将所有的方向向量累加起来并求平均值, 就得到了估计的方向向量 e :

$$e = \frac{1}{n} \sum_{i=1}^n e_i = \frac{1}{n} \sum_{i=1}^n \text{random}_{p,q} \left(\frac{w_p - w_q}{l_p - l_q} \right), l_p \neq l_q \quad (10)$$

其中, w_p 和 w_q 分别为样本 p 和 q 的特征向量; l_p 和 l_q 分别为样本 p 和 q 的标签。

方法一只适用于小样本的处理, 对于大样本处理效率很低。方法二采用了中位数二分法, 以标签的中位数 (Median) 为分界线, 低于中位数的置 0, 高于中位数的置 1, 将问题转化为二分类问题, 构造目标函数:

$$a \cdot x + b = y, y \in \{0, 1\} \quad (11)$$

然后用 Logistic 回归的方法求解, 求得的向量 a 近似作为方向向量 e :

$$e \approx a = \arg \max_a P(a \cdot x + b = y) \quad (12)$$

3.4 调整特征向量重构图像

给定一幅图像, 如果要对图像的某一视觉特征进行调整, 首先将图像通过图像编码网络转化为特征向量 w , 然后按照该视觉特征变化对应的方向向量 e 以一定的幅度 coeff 调整 w , 得到新的特征向量 w_{new} :

$$w_{\text{new}} = w + \text{coeff} \cdot e \quad (13)$$

将 w_{new} 通过合成网络重构图像, 合成的新图像与原图像相比仅在这一维视觉特征上发生了变化, 并且可以通过改变 coeff 控制视觉特征变化的幅度。

要控制图像实现多视觉特征的变化, 只需将上式可以改写为:

$$w_{\text{new}} = w + \text{coeff}_1 \cdot e_1 + \text{coeff}_2 \cdot e_2 + \dots + \text{coeff}_n \cdot e_n, i = 1, 2, \dots, n \quad (14)$$

其中, e_i 和 coeff_i 分别为视觉特征 i 对应的方向向量和幅度。

4 实验与分析

4.1 红外图像数据集

实验用 HF-160-2 型双光吊舱的红外热像仪系统对 1:400 的舰船模型进行取样, 从不同的水平角度 ($0 \sim 360^\circ$, 间隔 10°)、摄目距离 ($1.8 \sim 3.4$ m, 间隔 0.1 m) 对舰船模型进行拍摄, 获得 612 张原始的真实红外图像, 分辨率为 720×576 , 通过图像形态学处理 (镜像翻转、中心旋转、引入 Gauss 噪声) 进行样本增广, 最终获得 4284 张红外样本图像, 随机选取其中的 3672 张图像作为训练集, 612 张作为测试集。

4.2 训练渐进式生成对抗网络

4.2.1 训练参数设置

设计的合成网络共 8 个分辨率层级: $4 \times 4; 8 \times 8; \dots; 512 \times 512$, 输入生成网络的随机向量 z 和解耦后的特征向量 w 的维度为 (1, 512)。训练采用 Adam 优化算法, 设置学习率 $lr = 0.0001$, 一阶矩估计的指数衰减率 $\beta_1 = 0$, 二阶矩估计的指数衰减率 $\beta_2 = 0.99$; Loss_G 为非饱和的 Logistic 损失; Loss_D 是含梯度惩罚项的 Logistic 损失, 惩罚项系数 $\gamma_1 = 10$, $\gamma_2 = 0$; 训练的批尺寸 $\text{batch} = 32$, 在 Nvidia TITAN V 单 GPU 显卡 (显存 12G, 训练占用显存 9.8G) 上训练, 训练迭代次数 $\text{steps} = 164000$ 。

4.2.2 迭代次数与损失函数的关系

如图 6 所示, 随着迭代次数 steps 的增加, Loss_G 逐渐下降, 在第 20000 次左右达到了收敛; Loss_D 由三部分组成: $\text{Loss}_{D_{\text{real}}}$; $\text{Loss}_{D_{\text{fake}}}$ 和 gradient_penalty , 对应于图 6 中的 (b) (c) (d)。初始 Loss_D 较小, 是因为 G 生成图像的能力较弱, D 可以轻松鉴别出真实图像和生成的图像; 当 $0 < \text{step} < 20000$, G 生成的图像越来越“逼真”, D 很难鉴别真伪, 此时 Loss_D 上升; 当 $20000 < \text{step} < 100000$ 时, D 通过训练使得鉴别能力得到提升, 能够区分

出一些“高仿”的生成图像, $Loss_D$ 缓慢下降; 当 $step > 100000$ 时, G 和 D 趋向成熟, 两者处于一种动态较量的状态, 性能同时得到了提升: G 生成的图像越来越真实, D 鉴别真伪的能力越来越强。

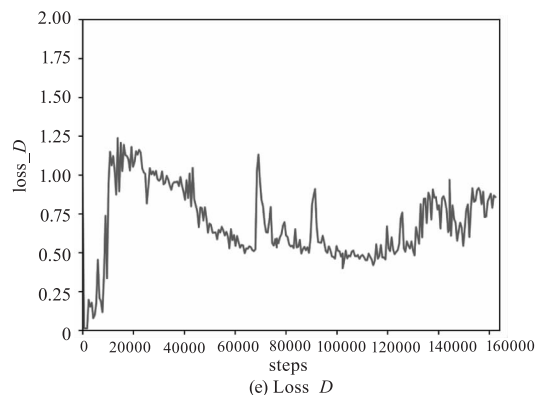
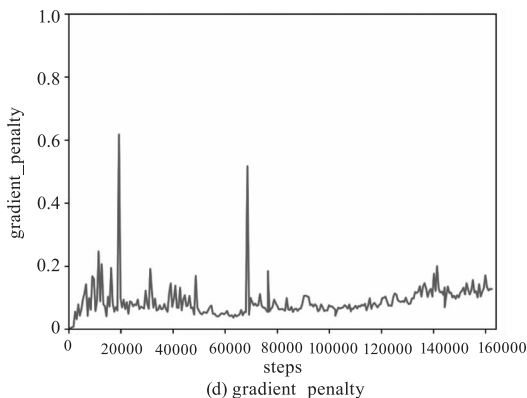
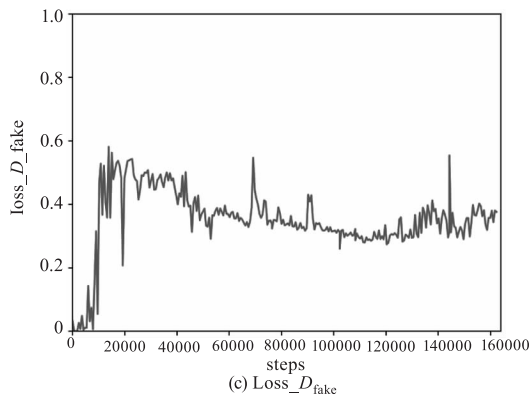
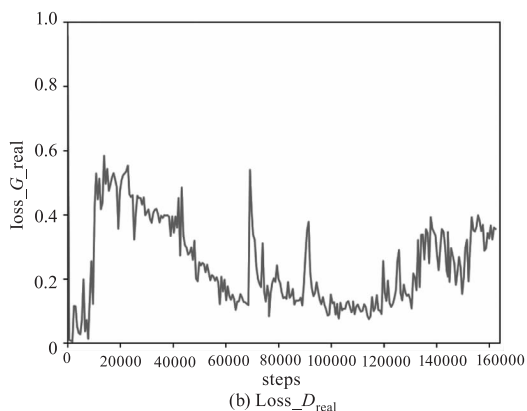
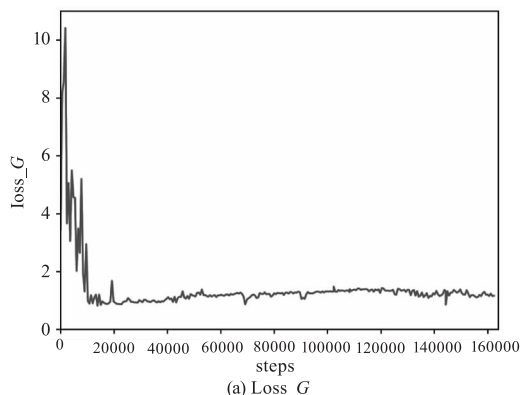


图6 红外数据集上损失函数的变化趋势

Fig. 6 Loss function's trend on infrared dataset

4.2.3 图像质量评价

评价网络训练的好坏是从视觉特征上对生成的图像样本进行评价, 为展现生成的图像在各分辨率层级上的训练效果, 在合成网络的 32×32 , 64×64 , 128×128 , 256×256 的四个分辨率层级的输出上添加了 `toImage()` 模块^[14], 负责将该层级的特征图转换为红外图像。指定一个固定的随机向量 $z(1 \times 512)$, 经过不同次数训练的生成网络, 得到在各分辨率层级上的图像输出, 如图7所示。

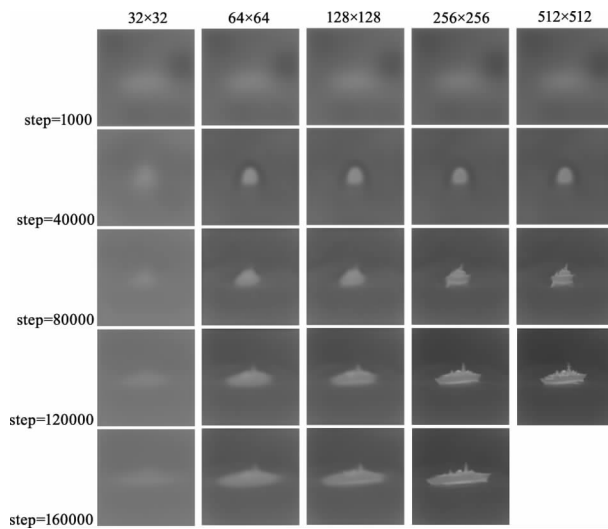


图7 不同分辨率层级图像的可视化结果

Fig. 7 Visual results of images at various resolutions

随着训练次数的增加, 合成网络生成的图像在各分辨率层级上的输出也愈加清晰。当训练进行到 160000 次, 生成网络最终生成的 512×512 图像轮廓清晰, 结构合理。

4.3 训练图像编码网络

4.3.1 训练参数设置

训练采用 Adam 优化算法, 学习率 $lr = 0.0003$,

一阶矩估计指数衰减率 $\beta_1 = 0$, 二阶矩估计指数衰减率 $\beta_2 = 0.99$; 感知模型用 vgg16_zhang_perceptual 的 conv1_1, conv1_2, conv3_2, conv4_2 四层, 损失函数正则项系数 $\lambda_{l2} = 3$, 模型经过 steps = 115000 的预训练和 steps = 18000 的微调, 损失最终收敛到 0.082。

4.3.2 相似度检测

实验选取了红外图像测试集上的 100 张图像, 经图像编码网络得到对应的特征向量, 然后将特征向量输入合成网络, 得到重构的图像。如图 8 所示, 与原始图像相比, 重构的图像保持了原有的轮廓结构, 但部分位置(如图 8(b)中的 1、2 和 3)的细节表现较为模糊。

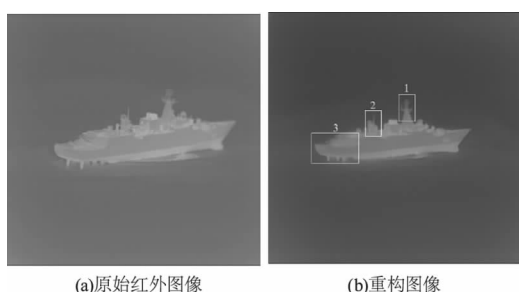


图 8 原始红外图像及其重构图像

Fig. 8 Original infrared image and reconstructed image

为了定量比较重构图像与原始红外图像的差异, 利用均值哈希算法(aHash)^[15]和平均结构相似性算法(MSSIM)^[16]对 100 张原始红外图像与对应的 100 张重构图像进行相似度检测, 结果见表 1。

表 1 重构图像的质量评价(S: Similarity)

Tab. 1 Quality evaluation of reconstructed images

	Low $S < 0.8$	Medium $0.8 \leq S < 0.9$	High $S \geq 0.9$
aHash	10	5	85
MSSIM	0	17	83

aHash 算法检测出有 85 张重构图像的相似度超过 90%, 为高质量图像, 有 10 张的相似度低于 80%, 为低质量图像; MSSIM 算法检测出所有的重构图像与原始红外图像的相似度均在 80% 以上。上述结果表明, 图像编码网络能够准确找到红外图像在特征向量域 W 对应的特征向量 w , w 通过合成网络重构的图像与原始图像有很高的相似度。

4.4 角度可控的红外图像生成

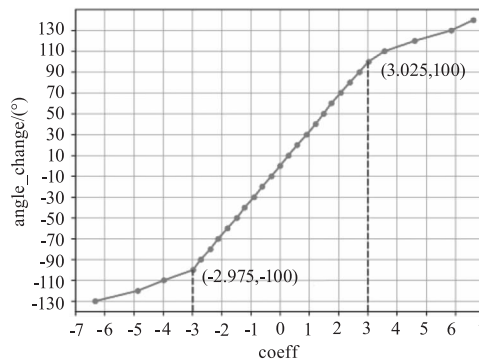
4.4.1 寻找角度属性的方向向量

在红外图像样本训练集上, 根据图像中舰船的角度添加标签: 以舰首正对热像仪为基准, 舰首左偏的图像记为 0; 舰首右偏的图像记为 1。利用 logistic 回归模型对图像对应的特征向量进行分类, 求得角度自左向右变化的方向向量 e , 并通过交叉验证的方式对模型的性能进行评估, 准确率最终达到了 94.87%。

4.4.2 调整特征向量重构图像

实验在红外图像测试集中随机选取了 100 张不同舰船角度的真实红外图像, 对每一幅红外图像: 首先经过图像编码网络转换为特征向量 w , 然后沿着方向向量 e 以不同的幅度 coeff_i 调整, 得到新的特征向量 w'_i , 将 w'_i 通过合成网络重构图像, 得到不同舰船角度的红外仿真图像 $G(w'_i)$, 然后将 $G(w'_i)$ 与真实的红外图像进行相似度检测, 对仿真图像的生成质量作出评价。

经过 100 次的重复实验, 图 9(a) 给出了 coeff 与角度变化的映射关系, 图 9(b) 给出了重构的红外图像与真实红外图像的相似度检测结果。在 $[-3.0, 3.0]$ 区间内, coeff 与舰船角度的变化近似为线性关系: coeff 每改变 0.3, 角度改变约 10° ; 重构的图像与真实图像的相似 $\text{similarity} > 0.8$ 。而在其他区间内, coeff 与舰船角度的变化为非线性关系, 且重构的图像均为低质量图像 ($\text{similarity} < 0.8$)。因此 coeff 一个合理的取值区间为 $[-3.0, 3.0]$, 在这一区间内可以对红外图像中舰船的角度进行定量控制, 且能保证重构图像的生成质量。对应的, 舰船角度改变的调整范围为 $[-100^\circ, 100^\circ]$, 如需生成舰船 360° 的全方位红外仿真图像, 则至少需要两张角度差大于等于 160° 的真实图像。



(a) coeff 与角度改变的关系

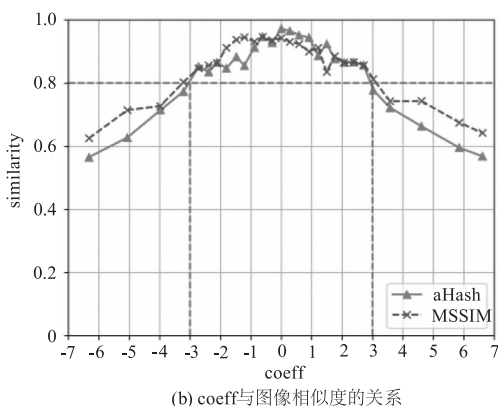


图 9 调整红外图像中舰船角度的可控化实验

Fig. 9 The experiment of adjusting warship's angle in infrared image in a controlled way

图 10 选取了其中一次实验的仿真结果,实验选取一幅舰首左偏 40° 的真实红外图像,以 0.3 的间隔调整 coeff 的值,调整特征向量并重构图像,实现了不同舰船角度的红外图像生成。在笔者实验的计算机平台上,利用训练好的合成网络进行图像生成,生成每幅图像的平均用时约为 27 ms,满足实时性仿真的需求。

5 结 语

本文提出了一种可控化的红外图像仿真技术,在渐进式对抗生成网络模型的基础上进行改进:引入映射网络,将输入向量转换为特征向量实现特征解耦;引入自适应归一化模块,实现对生成图像的样式控制,最终通过合成网络输出 512×512 的高分辨率红外仿真图像。基于 Resnet50 模型设计与合成网络逆向的图像编码网络,实现红外图像到特征向量的转换。在图像特征向量域,运用 logistic 回归的方法找到了控制红外图像中舰船角度的方向向量,沿着方向向量以不同的幅度对特征向量进行调整,通过合成网络重构图像,便可得到不同舰船角度的红外仿真图像,最后通过 aHash 算法和 MSSIM 算法对仿真图像和真实图像进行相似度检测,两者具有很高的相似度。实验结果表明:通过调整图像特征向量实现角度控制的仿真方法是可行的,这为小样本图像的数据增广技术提供了一种手段,并可以为真实舰船的红外图像仿真技术提供参考。

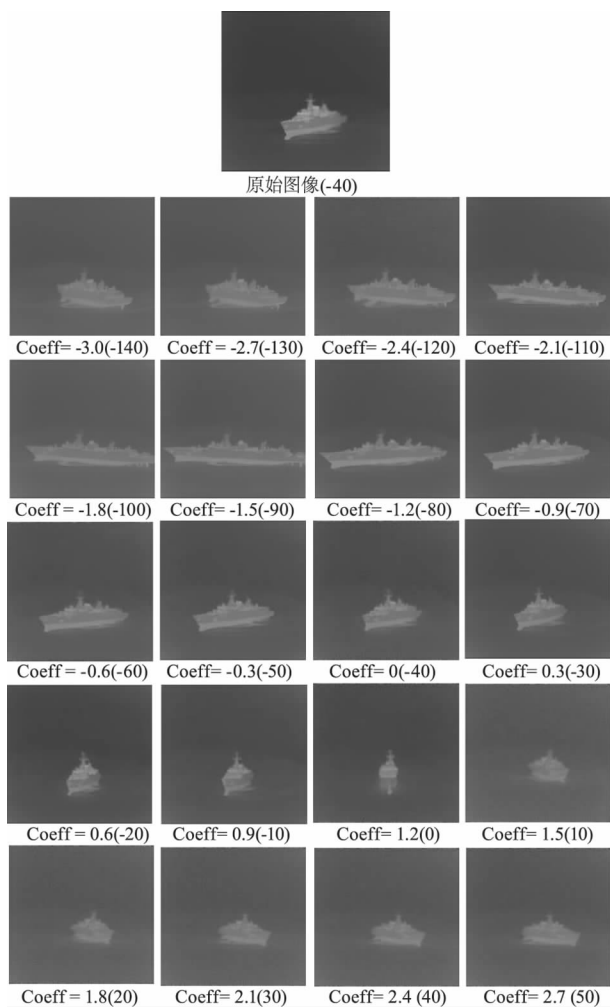


图 10 不同舰船角度的红外图像生成

(舰船角度以舰首正对热像仪为基准:左偏取负,右偏取正)

Fig. 10 Infrared image generation of different warship angles

(warship's angle is based on the bow over against thermal imager;
on left is negative; on right is positive)

参考文献:

- [1] Luo Haibo, Shi Zelin. Statlls and prospect of infrared imaging guidance technology [J]. Infrared and Laser Engineering, 2009, 38(4): 565 - 573. (in Chinese)
罗海波, 史泽林. 红外成像制导技术发展现状与展望 [J]. 红外与激光工程, 2009, 38(4): 565 - 573.
- [2] Zhao Xunjie, Li Chengjin. Key-tech in a hardware-in-the-loop IR simulation system [J]. Infrared and Laser Engineering, 2007(3): 41 - 43, 47. (in Chinese)
赵勋杰, 李成金. 红外半实物仿真系统的关键技术 [J]. 红外与激光工程, 2007(3): 41 - 43, 47.
- [3] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversari-

- al networks[J]. Computer Science, 2015.
- [4] Xie Jiangrong, Li Fanming, Wei Hong, et al. Infrared target simulation method based on generative adversarial neural networks[J]. Acta Optica Sinica, 2019, 39(3): 142–148. (in Chinese)
- 谢江荣, 李范鸣, 卫红, 等. 基于生成对抗式神经网络的红外目标仿真方法[J]. 光学学报, 2019, 39(3): 142–148.
- [5] Karras T, Aila T, Laine S, et al. Progressive growing of GANs for improved quality, stability and variation[J]. Arxiv Preprint Arxiv, 2017: 1710.10196.
- [6] Hu Haihe, Bai Tingzhu, Han Qiang, et al. Infrared illumination model and infrared scene simulation based on blinn-phong model[J]. Acta Optica Sinica, 2013, 33(6): 80–86. (in Chinese)
- 胡海鹤, 白廷柱, 韩强, 等. 基于 Blinn-Phong 模型的红外辐照模型及其红外场景仿真[J]. 光学学报, 2013, 33(6): 80–86.
- [7] Ren Liang, Chen Qian, Qian Weixian. Infrared target modeling and simulation based on ANSYS[J]. Infrared and Laser Engineering, 2013, 42(6): 1426–1431. (in Chinese)
- 任亮, 陈钱, 钱惟贤. 基于 ANSYS 的目标红外特性建模与仿真方法[J]. 红外与激光工程, 2013, 42(6): 1426–1431.
- [8] Wang Yan, Xie Xiaofang. Real-time IR imaging simulation of the surface warship[J]. Laser & Infrared, 2015, 45(8): 982–986. (in Chinese)
- 王彦, 谢晓方. 水面舰船目标的实时红外成像仿真[J]. 激光与红外, 2015, 45(8): 982–986.
- [9] Shi Chuanling. Real-time simulation of infrared land and sea scene[D]. Hangzhou: Zhejiang University, 2016. (in Chinese)
- 石川凌. 海陆场景的红外实时仿真研究[D]. 杭州: 浙江大学, 2016.
- [10] Yan M, Yu T. Modeling method of warship radiation model for infrared simulation[J]. Tactical Missile Technology, 2013.
- [11] Donahue J, Krähenbühl, Philipp, Darrell T. Adversarial feature learning[J]. Arxiv Preprint Arxiv, 2016; 1605.09782.
- [12] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//IEEE Conference on Computer Vision & Pattern Recognition. IEEE Computer Society, 2016.
- [13] Simonyan K, Zisserman A. Very deep convolutional networks for Large-Scale image recognition[J]. Computer Science, 2014.
- [14] Simonyan K, Vedaldi A, Zisserman A. Deep Inside convolutional networks: visualising image classification models and saliency maps[J]. Computer Science, 2013.
- [15] Li En, Wen Zhisheng, Zhang yulin, et al. Face similarity recognition system based on OpenCV and mean hash algorithm[J]. Industrial Design, 2016, (8): 132–133. (in Chinese)
- 李恩, 温志盛, 章育林, 等. 基于 OpenCV 与均值哈希算法的人脸相似识别系统[J]. 工业设计, 2016, (8): 132–133.
- [16] Alain Horé, Ziou D. Image quality metrics: PSNR vs. SSIM[C]//20th International Conference on Pattern Recognition, ICPR 2010, Istanbul, Turkey. IEEE Computer Society, 2010.