

基于双重对抗自编码网络的红外目标建模方法

苗壮^{1,2}, 张湧^{1,3*}, 李伟华^{1,2}

¹中国科学院上海技术物理研究所, 上海 200083;

²中国科学院大学, 北京 100049;

³中国科学院红外探测与成像技术重点实验室, 上海 200083

摘要 提出一种基于深度学习的红外目标建模方法。将对抗与自编码相结合, 设计了双重对抗自编码网络。利用训练后的网络, 仅需输入类别标签和满足一定分布的随机变量即可生成相应类别的红外目标图像。在自建红外数据集上对模型的有效性进行验证, 实验表明, 生成的目标图像在真实性和多样性等各方面均取得了较高的评价结果。将随机生成的目标图像作为小数据集的补充, 可有效改善训练数据匮乏的问题, 提高红外成像系统识别算法的准确率。

关键词 成像系统; 红外成像; 目标建模; 深度学习; 自编码网络; 生成对抗网络

中图分类号 TP391

文献标志码 A

doi: 10.3788/AOS202040.1111002

Infrared Target Modeling Method Based on Double Adversarial Autoencoding Network

Miao Zhuang^{1,2}, Zhang Yong^{1,3*}, Li Weihua^{1,2}

¹Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai 200083, China;

²University of Chinese Academy of Sciences, Beijing 100049, China;

³Key Laboratory of Infrared System Detection and Imaging Technology, Chinese Academy of Sciences, Shanghai 200083, China

Abstract In this study, we propose an infrared target modeling method based on deep learning. Further, we design a conditional double adversarial autoencoding network by combining the adversarial concepts with autoencoding. By using the trained network, the expected infrared target images can be easily generated by inputting category labels and the random variables that satisfy a certain distribution. The effectiveness of the proposed model is verified using a self-built infrared dataset. The conducted experiments prove that the generated target images exhibit considerable authenticity and diversity. Finally, the randomly generated target images as supplement the small data set can effectively improve the problem of lack of training data and improve the accuracy of the recognition algorithm in the infrared imaging system.

Key words imaging systems; infrared imaging; target modeling; deep learning; autoencoding network; generative adversarial network

OCIS codes 110.3080; 100.4996; 110.3000

1 引 言

红外成像系统因具有探测距离远、抗干扰性强、可全天候工作等优势, 被广泛应用于医疗、安防、军事等诸多领域。近年来, 随着深度学习在机器视觉领域的飞速发展, 各类性能优异的框架和模型相继被提出, 与之相结合的高精度红外目标识别已成为

红外成像领域研究的热点内容^[1-3]。而当前基于深度学习的红外目标识别最大的不足就是获取已标注的真实目标数据集的过程十分繁琐, 高昂的人力、物力成本及有限的外场实验次数往往导致模型训练不充分, 进而影响系统的整体性能。传统的红外目标建模方法通常是基于目标温度场的辐射特性来计算探测器接收的红外辐射能量, 需要考虑的约束条件

收稿日期: 2020-01-09; 修回日期: 2020-02-21; 录用日期: 2020-03-10

基金项目: 国家十三五国防预研项目(Jzz2016-0404/Y72-2)、上海市现场物证重点实验室基金(2017xcwzk08)

* E-mail: zybyx@sina.com

非常多,整个建模与计算过程复杂、耗时,同时生成的红外目标在真实性和多样性上存在一定的缺陷^[4-5]。文献[6]首次提出了基于生成对抗网络(GAN)的红外目标仿真方法,但是由于GAN训练不稳定等问题,目标图像的生成效果并不理想。针对上述问题,本文提出了基于双重对抗自编码的建模方法,通过将GAN和变分自编码(VAE)相结合,在保证图像真实性和多样性的前提下生成了不同类

别的红外目标,为目标识别算法的训练提供了更加丰富的样本,提高了目标识别的准确率。

2 模型框架

2.1 生成对抗网络

GAN是一种隐式深度生成模型,于2014年首次被Goodfellow提出^[7]。GAN模型框架如图1所示,包含生成器G和判别器D两个网络。

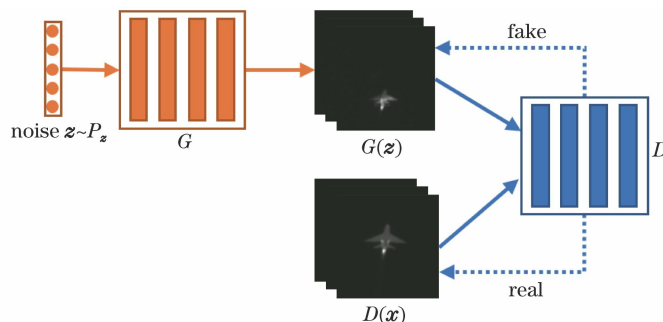


图1 GAN模型框架

Fig. 1 Structure of GAN model

生成器G将满足特定分布 P_z (如高斯分布或均匀分布)的噪声向量 z 作为输入,输出与真实数据 $D(x)$ 维度相同的生成数据 $G(z)$,同时期望 $G(z)$ 的数据分布与 $D(x)$ 的数据分布尽可能接近,达到“以假乱真”的效果。判别器D为一个二分类网络,假设真实数据给定标签为1,生成数据给定标签为0,将真实数据和生成数据分别送入判别器,期望真实数据经过判别器后输出值尽可能为1,而生成数据经过判别器后输出值尽可能为0,即最大可能地辨别数据来源的真伪。生成器和判别器之间的博弈可描述为

$$\min_D \max_G V(D, G) = -E_{x \sim P(x)} [\log D(x)] - E_{z \sim P(z)} \{\log \{1 - D[G(z)]\}\}, \quad (1)$$

式中: E 为期望值; x 为取自真实分布 P_x 中的

数据。

虽然目前基于GAN的各类衍生模型发展十分迅猛,但是相较于VAE等模型,在很多方面依然存在不足。GAN训练的最终目标为达到生成器与判别器的纳什均衡,即要求生成器生成的数据与真实数据无差异,且判别器无法对二者进行区分。由于目前尚无很好的方法达到该均衡,因此GAN训练的稳定性较难保证。同时由于模式崩溃现象的存在,GAN在生成数据的多样性上亦受到了限制。

2.2 变分自编码

VAE最早被Kingma等于2014年提出^[8]。作为现如今最热门的两大深度生成模型之一,VAE模型框架如图2所示,其由编码器E和解码器 D_g 两部分构成。

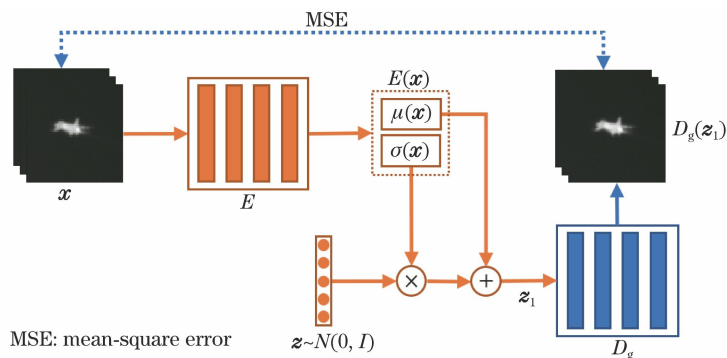


图2 VAE模型框架

Fig. 2 Structure of VAE model

编码器 E 将真实样本 x 作为输入,输出其在潜变量空间上的均值 $\mu(x)$ 和方差 $\sigma(x)$ 。假定潜变量空间的数据满足均值为 0, 方差为 I 的特定分布 $N(0, I)$, 利用重新参数化的技巧, 将 $z_1 = \mu(x) + \sigma(x) \times z$ 作为解码器 D_g 的输入, 最终输出重构样本 $D_g(z_1)$ 。通过反复迭代, 期望真实样本与重构样本的方均误差(MSE)足够小的同时, 真实样本在潜变量空间上的映射能够尽可能接近假定的数据分布, 其损失函数可表示为

$$L_{vae} = K_L\{N[\mu(x), \sigma(x)] \| N(0, I)\} + \|D_g(z_1) - x\|_2^2, \quad (2)$$

式中: $N[\mu(x), \sigma(x)]$ 为样本 x 经过编码器后在潜变量空间上的数据分布; K_L 为 Kullback-Leibler (KL) 散度, KL 散度越小, 数据分布越接近;

$\|D_g(z_1) - x\|_2^2$ 为 MSE。

相较于 GAN, VAE 的训练通常比较稳定且能够覆盖数据集中所有的样本。但是 VAE 在计算 KL 散度时实际优化的是对数似然函数的下界而非似然函数本身, 因此实际生成的图像通常比较模糊。结合 GAN 和 VAE 各自的优缺点^[9-10], 本文提出了一种双重对抗自编码(CDAAE)的红外目标建模方法, 该方法在 VAE 的基础上通过加入判别器和目标类别信息来生成多种类别的红外目标图像。

2.3 CDAAE 模型框架

CDAAE 模型框架共包含四个网络, 如图 3 所示, 分别为编码器 E 、解码器 D_g 、潜空间判别器 D_1 及样本判别器 D_s 。

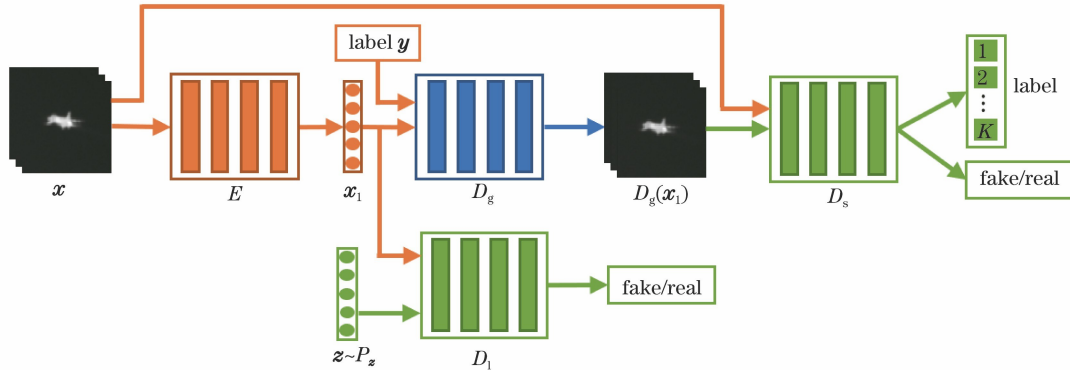


图 3 CDAAE 模型框架

Fig. 3 Structure of CDAAE model

在 CDAAE 模型中, 为了使映射到潜空间的数据 x_1 满足特定的数据分布 P_z , 不再使用 VAE 中近似计算 KL 散度的方式, 而是设计了一个潜空间判别器 D_1 , 通过神经网络来学习该数据分布, 形成了模型中的第一重对抗。编码器输出的潜空间数据 x_1 在和类别标签 y 结合后, 经过解码器 D_g 生成重构后的样本。除常规通过计算重构样本与真实样本的 MSE 来优化解码器网络的方式外, 提出了平均误差(MAE)的计算方法并设计了样本判别器 D_s , 用来对重构样本与真实样本进行逻辑判别和分类, 从而进一步约束解码器网络, 提高样本生成效果, 由此形成了模型的第二重对抗。通过对整个网络的训练, 最后仅需将类别标签和符合分布 P_z 的随机采样数据输入解码器中, 即可获得形貌逼真、样式丰富的红外目标图像。

3 算法原理

为了高效地生成符合需求的红外目标图像, 提

出了基于 CDAAE 的红外建模方法。整个算法分为模型训练和目标生成两个阶段。在模型训练阶段, 通过神经网络学习样本的潜在数据分布, 在目标生成阶段利用学习得到的网络参数生成期望类别的红外目标。

3.1 模型训练

第一重对抗为编码器 E 与潜空间判别器 D_1 之间的对抗, 图 4 给出了二者的网络结构。编码器 E 包含了 2 个卷积层和 2 个全连接层, 共 4 个中间层, 其中所有卷积层的卷积核尺寸均为 $4 \text{ pixel} \times 4 \text{ pixel}$ 、步长均为 2 pixel 。潜空间判别器 D_1 包含了 3 个全连接层并以其作为中间层。为了提高训练效率和稳定性, 所有的中间层在完成数据非线性变换后均使用批量归一化(BN)处理, 并将 Leaky ReLU 作为激活函数。

训练开始阶段, 首先从数据集中随机选取 m 张图片组成批处理样本集合 x , 并将其作为编码器 E 的输入。 x 在通过两层卷积(Conv1, Conv2)处理

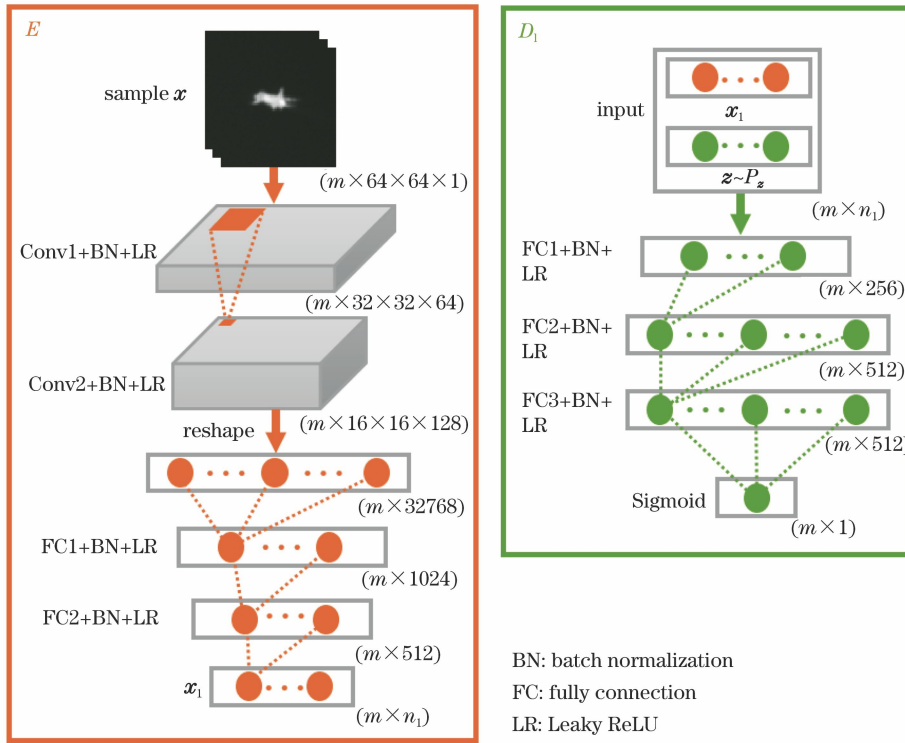


图 4 编码器与潜空间判别器的网络结构

Fig. 4 Network structures of encoder and latent-space discriminator

后,经维度变换,维度从四维降到二维,再通过 2 个全连接层(FC1,FC2)对数据进行压缩,输出维度为 $(m \times n_1)$ 的潜空间数据 \mathbf{x}_1 ,从而完成从原始数据到潜空间数据的映射。为了使映射后的数据满足特定的数据分布 P_z ,将潜空间数据 \mathbf{x}_1 和从 P_z 中随机采样的数据 \mathbf{z} 分别输入到潜空间判别器 D_1 中,经过 3 个全连接层(FC1,FC2,FC3)处理后,通过 Sigmoid 函数输出判别概率。最后利用判别概率分别计算编码器和潜空间判别器的熵损失函数 L_E 和 L_{D_1} ,并反向传播更新两个网络的参数 θ_E 和 θ_{D_1} ,以达到最小化损失函数的目标。 L_E 和 L_{D_1} 的表达式为

$$L_E = -E_{\mathbf{x} \sim P_x} \{ \log D_1[E(\mathbf{x})] \}, \quad (3)$$

$$L_{D_1} = -E_{\mathbf{x} \sim P_x} \{ 1 - \log D_1[E(\mathbf{x})] \} - E_{\mathbf{z} \sim P_z} \{ \log [D_1(\mathbf{z})] \}. \quad (4)$$

在完成第一重对抗的训练后,紧接着会进入第二重对抗的训练。第二重对抗为解码器 D_g 与样本判别器 D_s 之间的对抗,图 5 为二者的网络结构。由于解码实际上是编码的逆过程,因而解码器 D_g 采用了与编码器对称的网络结构,包含 3 个全连接层和 2 个反卷积层,实现了潜空间数据到重构样本的映射。样本判别器 D_s 由于输入数据与编码器相同,因而采用了与其类似的中间层网络结构,包含 2 个卷

积层和 2 个全连接层。其中,解码器 D_g 和样本判别器 D_s 的所有卷积层和反卷积层的卷积核大小均为 $4 \text{ pixel} \times 4 \text{ pixel}$ 、步长均为 2 pixel 。在所有中间层的最后均进行 BN 处理并使用 Leaky ReLU 作为激活函数。从图 5 不难看出,样本判别器 D_s 最终不仅会判别样本的真伪,同时还会输出样本的类别标签,因而该网络还具备标注未分类样本的功能。

第二重对抗开始,首先对第一重对抗中编码器 E 的输出 \mathbf{x}_1 与对应样本集合的类别标签 \mathbf{y} 进行连接,并以此作为解码器 D_g 的输入 \mathbf{x}_{1c} 。其中,类别标签 \mathbf{y} 是对原始类别标签 one hot 编码后的结果, \mathbf{y} 的维度为 $(m \times n)$ 。 \mathbf{x}_{1c} 在经过 3 个全连接层(FC1,FC2,FC3)处理后,再通过 2 层反卷积(Deconv1,Deconv2)完成数据的上采样,最终生成重构后的样本 $D_g(\mathbf{x}_{1c})$ 。利用重构样本,将 MSE 和 MAE 相结合的重构损失函数 $L_{D_g\text{-recon}}$ 作为解码器总损失函数 L_{D_g} 的一部分,表达式为

$$L_{D_g\text{-recon}} = \beta \times \|D_g(\mathbf{x}_{1c}) - \mathbf{x}\|_2^2 + (1 - \beta) \times \|D_g(\mathbf{x}_{1c}) - \mathbf{x}\|_1, \quad (5)$$

式中: $\|D_g(\mathbf{x}_{1c}) - \mathbf{x}\|_2^2$ 为重构样本和原始样本的 MSE; $\|D_g(\mathbf{x}_{1c}) - \mathbf{x}\|_1$ 为重构样本和原始样本的 MAE; β 为 MSE 和 MAE 的权重系数,取值范围为 $(0, 1)$,后续实验 β 取 0.9。

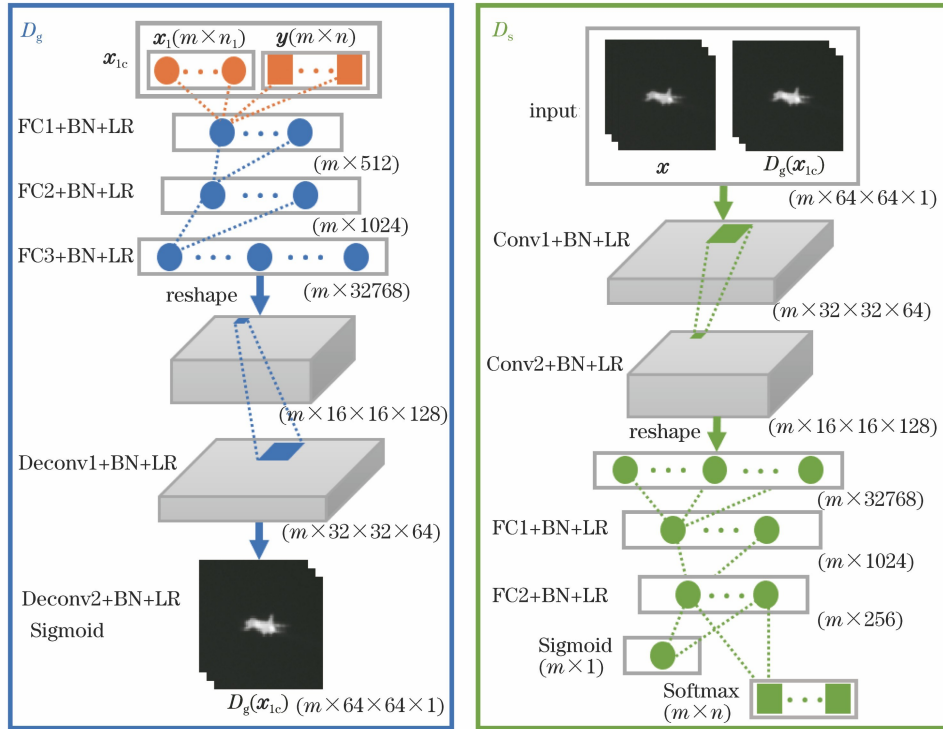


图 5 解码器与样本判别器的网络结构

Fig. 5 Network structures of encoder and sample discriminator

将生成的重构样本 $D_g(\mathbf{x}_{1c})$ 与真实样本 \mathbf{x} 分别送入样本判别器 D_s 。经过卷积层 (Conv1, Conv2) 和全连接层 (FC1, FC2) 一系列处理, 最终输出经 Sigmoid 函数变换的判别概率和经 Softmax 函数变换的类别概率。利用该输出结果, 可分别计算解码器和样本判别器的熵损失函数 $L_{D_g\text{-entropy}}$ 和 L_{D_s} , 表达式为

$$L_{D_g\text{-entropy}} = -E_{\mathbf{x}_{1c} \sim P_{\mathbf{x}_{1c}}} \{ \log P_{D_s} [1 | D_g(\mathbf{x}_{1c})] \} - E_{\mathbf{x}_{1c} \sim P_{\mathbf{x}_{1c}}} \{ \log P_{D_s} [\mathbf{y} | D_g(\mathbf{x}_{1c})] \}, \quad (6)$$

式中: 当输入为重构样本 $D_g(\mathbf{x}_{1c})$ 时, $P_{D_s} [1 | D_g(\mathbf{x}_{1c})]$ 为样本判别器判断重构样本为真的概率, $P_{D_s} [\mathbf{y} | D_g(\mathbf{x}_{1c})]$ 为样本判别器估计重构样本类别为 \mathbf{y} 的概率。

$$L_{D_s} = -E_{\mathbf{x} \sim P_{\mathbf{x}}} \{ 1 - \log \{ P_{D_s} [0 | D_g(\mathbf{x}_{1c})] \} \} - E_{\mathbf{x} \sim P_{\mathbf{x}}} \{ \log [P_{D_s} (1 | \mathbf{x})] \} - E_{\mathbf{x} \sim P_{\mathbf{x}}} \{ \log [P_{D_s} (\mathbf{y} | \mathbf{x})] \}, \quad (7)$$

式中: 当输入为重构样本 $D_g(\mathbf{x}_{1c})$ 时, $P_{D_s} [0 | D_g(\mathbf{x}_{1c})]$ 为样本判别器判断重构样本为假的概率; 当输入为真实样本 \mathbf{x} 时, $P_{D_s} [1 | D_g(\mathbf{x})]$ 为样本判别器判断真实样本为真的概率, $P_{D_s} [\mathbf{y} | D_g(\mathbf{x})]$ 为样本判别器估计真实样本类别为 \mathbf{y} 的概率。

结合(5)和(6)式, 解码器 D_g 的总损失函数为 $L_{D_g} = L_{D_g\text{-entropy}} + L_{D_g\text{-recon}}$ 。在第二重对抗的最后, 利用计算得到的 L_{D_g} 和 L_{D_s} , 分别反向传播更新解

码器和样本判别器的网络参数 θ_{D_g} 和 θ_{D_s} , 一次完整的训练至此结束。通过不断的迭代, 最终模型将收敛或局部收敛。

3.2 目标生成

在模型训练结束后, 根据需要生成的红外目标的类别标签和样本数量, 在潜空间对特定的分布 P_z 进行随机采样。将随机采样得到的数据 \mathbf{z} 和对应的类别标签输入解码器 D_g , 即可生成随机多样的红外目标。

4 实验与分析

4.1 平台与数据

所有实验均使用 tensorflow 建模, 并在一块 GTX1080Ti 的 GPU 上完成所有训练。为了充分验证所提方法的有效性, 使用自建的红外数据集进行实验。自建红外数据集采自外场实验数据, 以对空红外告警系统为依托, 选取战斗机、直升机、训练机、客机 4 大类共计 5500 张红外目标图像, 所有红外目标图像均为 8 位 $64 \text{ pixel} \times 64 \text{ pixel}$ 。同时从模型可重用性角度出发, 通过调整模型参数, 增加了基于 MNIST 手写数据集的相关实验。MNIST 数据集包含 10 类数字, 共计 55000 个训练样本和 10000 个测试样本, 所有样本均为 $28 \text{ pixel} \times 28 \text{ pixel}$ 的灰度图像。

4.2 参数与训练

在 CDAAE 模型训练过程中,超参数的取值会直接影响模型的训练效果,因此需要根据数据集的样本特性和模型的网络结构进行调节。实验表明,当数据集样本种类较多、样本细节较丰富时,适当增加潜空间映射数据的维度能够有效提高样本的生成质量。在使用 Adam 算法^[11]优化网络的损失函数时,单独对四个网络的学习速率进行调节,可提高训练效率。同时,在单次训练中,调整各个网络的优化次数亦可提高模型的训练速度。表 1 为实验所采用的超参数。

4.3 质量评价

生成图像的质量是评价建模方法的重要参考依据。根据评价方法的不同,可分为主观评价和客观评价两种。主观评价是指通过实验人员的视觉感受直接评价图像质量。图 6 为模型在红外数据集上训练后的实验结果,可以看到,虽然测试集中的目标图像并未参与训练,但是重构后的图像依然保持了与原始图像极高的相似度。对于随机生成的红外目

表 1 CDAAE 模型训练的超参数

Table 1 Hyper parameters used for CDAAE model training

Hyper parameter	Infrared dataset	MINIST dataset
Training loop	100000	100000
Class number	4	10
Batch size	64	256
Slope of Leaky ReLU	0.01	0.01
Learning rate of E/D_g	0.00002	0.00002
Learning rate of D_1/D_s	0.0001	0.0001
Data dimension in latent-space	8	4
Optimization of $E/D_g/D_1/D_s$ per training loop	2:4:1:1	2:4:1:1

标,首先其保持了与真实目标相似的红外辐射特性,例如喷射尾焰辐射、发动机高温机舱辐射等,具有真实性;其次生成的目标纹理清晰、样式丰富,例如飞行姿态、目标形貌等,其中许多特征是真实目标所不具备的,符合多样性的要求。

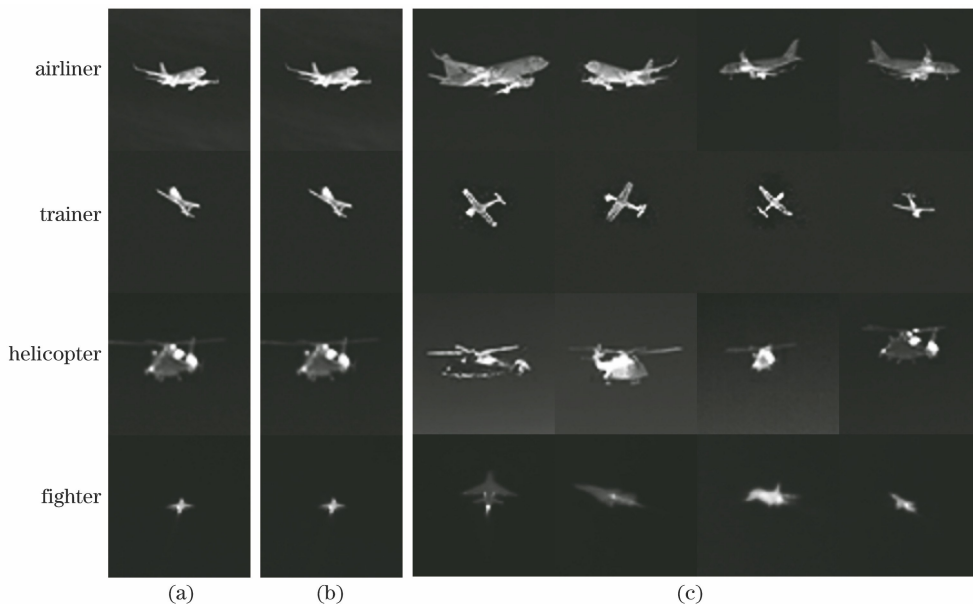


图 6 CDAAE 模型在红外数据集上的实验结果。(a)真实目标;(b)重构目标;(c)随机生成目标

Fig. 6 Test results of CDAAE model on infrared dataset. (a) Real samples; (b) reconstructed samples; (c) random samples

图 7 为模型在 MINIST 数据集上训练后的实验结果。可以看到,实验生成的样本,无论是在清晰度还是结构、轮廓,均达到了与真实样本近乎相同的视觉效果。

为了进一步验证所提方法的有效性,选取 FID (Fréchet inception distance)^[12]作为客观评价的指标。FID 利用特定网络提取真实目标数据集与生成目标数据集的特征向量,再统计并映射该特征向量

使其分别满足高斯分布,通过计算两个高斯分布的 Fréchet 距离来判别两个数据集的相似程度。FID 值越小,生成效果越好。

表 2 为 CDAAE 模型在两个数据集上的 FID 评估结果。其中,第一行为将真实样本集拆分成两个独立的样本集后相互比较的结果;第二行比较了真实样本与经 CDAAE 模型重构后的样本;第三行比较了真实样本与 CDAAE 模型随机生成的样本;

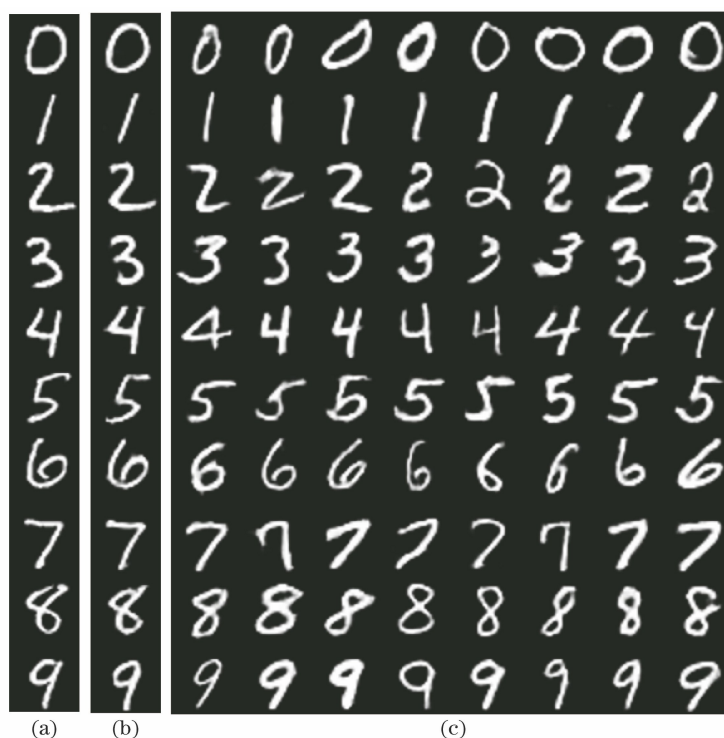


图 7 CDAAE 模型在 MNIST 数据集上的实验结果。(a)真实图像;(b)重构图像;(c)随机生成图像

Fig. 7 Test results of CDAAE model on MNIST dataset. (a) Real samples; (b) reconstructed samples; (c) random samples

表 2 CDAAE 模型的 FID 评估结果

Table 2 FID evaluation results of CDAAE model

Condition	Infrared dataset	MNIST dataset
Group 1 vs group 2	10.13	8.85
Real samples versus reconstructed samples	14.26	7.51
Real samples versus random samples	35.36	10.97
Real samples versus random samples from α -GAN	32.47	12.18
Real samples versus random samples from C-DCGAN	48.83	27.62
Real samples versus random samples from CVAE	52.74	25.55

最后三行对真实样本与经典模型 α -GAN^[10]、C-DCGAN^[6]、CVAE 随机生成的样本进行了比较。需要说明, α -GAN 由于没有考虑类别信息,因此不具备输出多种类别目标的能力,表 2 中对应结果均为使用训练集单一类别样本进行训练后的结果;C-DCGAN 使用了文献[6]中仿真方法的网络结构,将红外数据集中的目标图像分辨率从 64 pixel \times

64 pixel 降低为 32 pixel \times 32 pixel,并对该分辨率下随机生成的样本进行比较。从表 2 可知:CDAAE 模型重构样本的 FID 非常低,说明重构样本与真实样本非常相似,模型具有强大的样本重构能力;CDAAE 模型随机生成样本和重构样本与真实样本相互比较的 FID 比较接近,说明 CDAAE 模型能够学习真实样本的数据分布,能生成多样的高质量样本,这与主观评价的结论是一致的,从侧面说明了模型应用于小数据集数据增强的可行性; α -GAN 虽然得到了与 CDAAE 模型相似的结果,但由于是单一类别生成,因此从工程角度上来说,将 α -GAN 作为建模方法实用性不高;CDAAE 模型在保证输出图像分辨率更高的情况下,FID 依然远小于 C-DCGAN 的实验结果,优势十分显著;CVAE 的 FID 与 C-DCGAN 相当,与 CDAAE 模型相比差距较大。

4.4 数据增强

卷积神经网络(CNN)因具有强大的特征提取能力而在目标识别领域得到了广泛的应用。在网络训练过程中,训练样本的数量将直接影响网络识别的准确率。而在实际的工程中,大量真实样本的获取往往是十分繁琐的,这一问题在红外系统的应用中尤为突出。作为扩充小训练集的一种方法,数据增强根据训练集处理方式的不同,分为变换式和生

成式两大类。其中生成式的数据增强要求生成样本保持真实性。通过 4.3 节的质量评价可知,CDAAE 模型生成的红外目标可以满足数据增强的基本要求。

首先,分别从红外数据集和 MINIST 数据集中选取 1200 张和 3000 张真实样本用于分类器和 CDAAE 模型的训练,选取 600 张和 1500 张真实样本用于分类器的测试;然后依次将每类 300 张和 600 张随机生成的样本加入到训练集中进行对比训练。分类器网络采用与 CDAAE 模型中样本判别器类似的结构,实验结果如表 3 所示。实验表明,对于小数据集,随着训练集中生成样本数量的增加,分类器的准确率在不断提高,证明了所提方法的实用性。

表 3 数据增强过程中的准确率

Table 3 Accuracy in data augmentation process %

Number of pictures in Infrared dataset	Number of pictures in MINIST dataset	Infrared dataset	MINIST dataset
1200	3000	92.83	97.80
1200+1200	3000+3000	94.33	98.33
1200+2400	3000+6000	96.00	99.26

5 结 论

提出了一种基于 CDAAE 模型的红外目标建模方法。设计了编码器与潜空间判别器、解码器与样本判别器的双重对抗网络结构及相应的损失函数。在自建红外数据集上对模型进行训练,并分别从主观和客观两个角度对生成目标的图像质量进行评价,随机生成的目标形貌逼真且样式丰富,验证了所提方法的有效性。实验表明,使用随机生成的目标图像扩充数据集可以有效解决训练数据不足的问题,提高了系统识别算法的准确率。下一步将对如何优化模型生成更高质量的图像及利用模型对图像进行标注、复原等展开深入研究。

参 考 文 献

- [1] D'Acremont A, Fablet R, Baussard A, et al. CNN-based target recognition and identification for infrared imaging in defense systems[J]. *Sensors*, 2019, 19(9): 2040.
- [2] Kim S, Song W J, Kim S H. Infrared variation optimized deep convolutional neural network for robust automatic ground target recognition [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 195-220.
- [3] Liu Q, Lu X H, He Z Y, et al. Deep convolutional neural networks for thermal infrared object tracking [J]. *Knowledge-Based Systems*, 2017, 134: 189-198.
- [4] Zou T, Tong Z X, Wang C Z, et al. Research on infrared image simulation of air targets based on tactic fusion [J]. *Laser & Infrared*, 2016, 46(4): 444-451.
邹涛, 童中翔, 王超哲, 等. 基于战技融合的空中目标红外图像仿真研究 [J]. *激光与红外*, 2016, 46(4): 444-451.
- [5] Wang X, Gao S L, Li F M. Infrared imaging modeling and simulation of aerial target based on BRDF [J]. *Journal of Infrared and Millimeter Waves*, 2019, 38(2): 182-187.
王霄, 高思莉, 李范鸣. 基于 BRDF 的空中目标红外成像建模与仿真 [J]. *红外与毫米波学报*, 2019, 38(2): 182-187.
- [6] Xie J R, Li F M, Wei H, et al. Infrared target simulation method based on generative adversarial neural networks [J]. *Acta Optica Sinica*, 2019, 39(3): 0311002.
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): 0311002.
- [7] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets [C] // *Advances in Neural Information Processing Systems 27 (NIPS 2014)*, December 8-13, 2014, Montreal, Quebec, Canada. New York: Curran Associates, 2014: 2672-2680.
- [8] Kingma D P, Welling M. Auto-encoding variational bayes [EB/OL]. (2014-05-01) [2019-12-25]. <https://arxiv.org/abs/1312.6114>.
- [9] Bao J M, Chen D, Wen F, et al. CVAE-GAN: fine-grained image generation through asymmetric training [C] // 2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice. New York: IEEE, 2017: 2745-2754.
- [10] Rosca M, Lakshminarayanan B, Warde-Farley D, et al. Variational approaches for auto-encoding generative adversarial networks [EB/OL]. (2017-10-21) [2019-12-25]. <https://arxiv.org/abs/1706.04987>.
- [11] Kingma D P, Jimmy L B. Adam: a method for stochastic optimization [EB/OL]. (2017-01-30) [2019-12-25]. <https://arxiv.org/abs/1412.6980>.
- [12] Heusel M, Ramsauer H, Unterthiner T, et al. GANs trained by a two time-scale update rule converge to a local nash equilibrium [C] // *Advances in Neural Information Processing Systems 30 (NIPS 2017)*, December 4-9, 2017, Long Beach, CA, USA. New York: Curran Associates, 2017: 6626-6637.