文章编号:1674-2974(2022)04-0011-07 DOI:10.16339/j.cnki.hdxbzkb.2022277

基于对抗样本的深度学习图像压缩感知方法

王继良[1,2],周四望[1†],金灿灿[1]
(1. 湖南大学 信息科学与工程学院,湖南 长沙 410082;
2. 长沙环境保护职业技术学院,湖南 长沙 410004)

摘 要:压缩感知是研究数据采样压缩与重构的信号处理新理论,近年来研究人员将深度学习运用到图像压缩感知算法中,显著提高了图像重构质量.然而,图像信息常与隐私关联,高质量的重构图像在方便人们观赏的同时,带来了隐私保护的问题.本文基于深度学习理论,提出一种对抗的图像压缩感知方法.该方法将压缩理论和对抗样本技术统一于同一个压缩感知算法,通过设计损失函数,联合重构误差和分类误差来训练压缩感知深度神经网络,使得压缩感知重构样本同时也是一个对抗样本.因此,重构图像在保证重构质量的同时,也能对抗图像分类算法,降低其识别率,达到保护图像隐私的效果.在Cifar-10和MNIST图像集上进行的实验结果表明,和已有的压缩感知方法相比,我们提出的对抗压缩感知方法以损失仅10%的图像重构质量为代价,使得图像分类精度下降了74%,获得了很好的对抗性能.

关键词:对抗样本;深度学习;图像;压缩感知
中图分类号:TP391 **文献标志码**:A

Method of Deep Learning Image Compressed Sensing Based on Adversarial Samples

WANG Jiliang[1,2],ZHOU Siwang[1†],JIN Cancan[1]
(1. College of Computer Science and Electronic Engineering, Hunan University, Changsha 410082, China;
2. Changsha Environmental Protection College, Changsha 410004, China)

Abstract: Compressed sensing is a new signal processing theory focusing on data sampling compression and reconstruction. In recent years, researchers have applied deep learning to image compressed sensing algorithms, which significantly improves the quality of the recovered images. However, images are often associated with personal privacy, and high-quality recovered images often bring privacy protection problems while facilitating people's viewing. Based on deep neural network, this paper proposes an image compressed sensing algorithm with adversarial learning. This method integrates data compression and adversary sample technique into the compressed sensing algorithm. By training the neural network with a loss function combining reconstruction loss and classification loss, the output samples, i.e., the recovered images, become adversarial samples. The recovered images with our proposed algorithm can then be adversarial to image classifications algorithms, decreasing their recognition rate and achieving the perfor-

* **收稿日期**:2021-11-10
基金项目:国家自然科学基金资助项目(6217071837), National Natural Science Foundation of China(6217071837);湖南省自然科学基金资助项目(2020JJ7010), Natural Science Foundation of Hunan Province(2020JJ7010)
作者简介:王继良(1973—),女,湖南娄底人,副教授
† 通信联系人,E-mail: swzhou@hnu.edu.cn

mance of protecting image privacy while guaranteeing a reasonable image quality. Experimental results on Cifar-10 and MNIST show that, compared with the existing compressed sensing methods, the proposed adversarial algorithm achieves excellent adversarial performance, as the classification accuracy is decreased by 74% at the cost of 10% loss of image reconstruction quality.

Key words: adversarial sample; deep learning; image; compressed sensing

压缩感知是研究数据采样压缩与重构的信号处理新理论[1-3].压缩感知理论突破了奈奎斯特采样定理的限制,能降低图像获取成本、节省图像的存储空间和传输开销,在图像处理领域已经取得了成功应用.迄今为止,已有多种图像压缩感知算法被提出,目标是获得更高的图像重构质量.经典的图像压缩感知重构算法包括基于消息传递 AMP 框架的算法[4, 5]、应用于二进制图像的压缩感知算法[6]、自适应压缩感知算法[7].我们则提出了基于分块的图像压缩感知算法[8, 9].压缩感知理论有着严谨、完备的数学基础,但图像重构算法复杂度高,运行时间长.

受深度学习研究进展的鼓舞,近年来研究人员开始探索基于深度神经网络的图像压缩感知算法[10-12].深度学习压缩感知利用深度神经网络的学习能力,在有标签的训练集中学习从原始输入样本到重构样本的映射,实现压缩感知重构.ReconNet 是较早提出的压缩感知深度网络模型[13],文献[14]对此网络模型做了改进,通过联合学习测量进程和重构进程来优化压缩感知测量矩阵,在低采样率下有更好的重构性能.受分块压缩感知算法的启发,文献[15-16]提出 CSNet 网络结构,图像压缩采用分块方法,但用一个深度网络实现整体图像重构,从而提高了图像重构质量.我们对 CSNet 做了深入研究,根据图像各块的重要性自适应分配采样率,进一步提高了 CSNet 的重构效果[17].和传统压缩感知方法相比,深度学习算法有显著更快的重构速度,在低采样率时有更好的图像重构效果.

然而,高质量的重构图像更容易被图像分类算法自动识别,带来了隐私保护问题.图像识别是指通过特征提取算法提取图像样本的特征,再通过分类器将图像样本划分到一定的类别中,从而实现自动分类.Hinton 和 A. Krizhevsky 设计的深度神经网络 AlexNet 是机器识别发展的一个里程碑[18].在此基础上,VGG[19]、ResNet[20]和 EfficientNet[21]等深度网络陆续被提出,获取了更高的分类识别精度.以人脸图像识别为例,目前分类算法的识别精度已经超过了人类本身,“刷脸”进站、“刷脸”支付等极大地方便了人们的生活.然而,图像自动识别是一把“双刃剑”.更高的图像识别率往往意味着更多的图像隐私被暴露.图像被隐藏于网络中的机器模型自动识别,带来安全隐患.

本文研究图像压缩感知中的安全问题.我们注意到深度学习模型存在某种程度的脆弱性.在文献[22]中,一个“鲸鱼”图像样本被对抗算法修改,虽然视觉上依然是“鲸鱼”,但识别算法失效了,“鲸鱼”样本被误识别成了“乌龟”.对抗算法的核心思想是扰动样本,生成对抗样本,迷惑图像分类模型,使之失效[23-24].从中受到启发,本文提出对抗的图像压缩感知方法,利用机器模型的脆弱性来保护图像重构样本.我们提出的压缩感知方法同时兼具压缩和对抗的功能,其生成的重构图像也是一个对抗样本,在保证图像质量的同时,能对抗图像分类算法,保护图像隐私.

本文的主要内容组织如下:第 1 节提出一种基于对抗样本的图像压缩感知方法;第 2 节设计实验以验证所提方法的性能;在第 3 节给出结论.

1 基于对抗样本的图像压缩感知方法

我们的目标是将安全性融入压缩感知网络,使得压缩感知算法同时兼具压缩和对抗两项功能.压缩感知算法的安全性通过对抗样本技术加以实现,重构图像不会影响视觉效果,但能对抗图像分类算法,降低图像分类算法的识别率,客观上起到保护图像隐私的效果.

1.1 压缩感知深度网络模型

现有的深度学习压缩感知方法通常用一个压缩

子网来实现图像的采样压缩,再用一个重构子网实现图像重构.一般地,压缩感知算法的深度网络模型如图1所示.

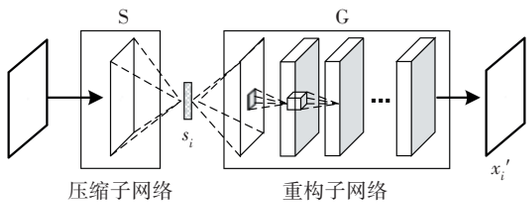


图1　压缩感知算法的深度网络模型

Fig.1　Deep learning model for compressed sensing

网络由压缩子网S和重构子网G组成.压缩子网S以原始景象x_i为输入,输出采样测量值s_i:

$$S(x_i)=s_i \tag{1}$$

重构子网则努力由采样测量值s_i恢复原始景象x_i,即

$$G(s_i)=x_i' \tag{2}$$

用$|s_i|$和$|x_i|$分别表示$|s_i|$和$|x_i|$的大小,则$|s_i|/|x_i|$的比值即为压缩子网S的采样率,或者称为压缩比.压缩感知深度网络模型期待x_i'能够以最小重构误差逼近原始图像x_i,即重构图像x_i'有好的重构质量和视觉效果.

1.2　基于对抗样本的压缩感知深度网络

在现有压缩感知深度网络模型的基础上,本小节提出一种对抗策略,使得重构子网G生成的压缩感知重构图像x_i'成为一个对抗本.

1.2.1　针对图像分类模型C的对抗模型

设C代表某一个图像分类网络.针对C,我们设计相应的压缩感知网络对抗模型,命名为Adv-G-C,如图2所示.Adv-G-C的目标是G网络生成的重构图像x_i'能对抗模型C.换句话说,模型C能正常识别一般的图像,但不能识别x_i'.

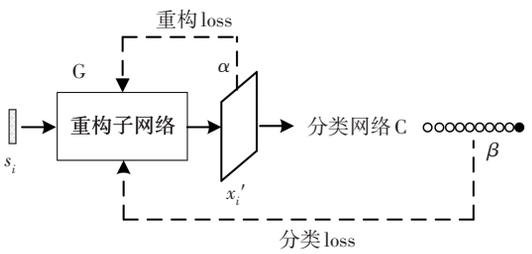


图2　针对模型C的压缩感知对抗模型Adv-G-C

Fig.2　Adversarial Adv-G-C against model C

正式地,针对图像分类模型C的压缩感知重构网络的对抗模型可定义为一个神经网络:

$$\text{Adv-G-C}: s_i \rightarrow x_i' \tag{3}$$

式中:"→"指神经网络的输出操作.该模型由压缩感知重构子网络G和一个已知的分类模型C组成.模型的输入是压缩感知测量值s_i,输出则是重构图像x_i'.在这里,C是预训练好的分类网络,C本身不参与对抗模型的训练.从另一角度说,我们提出的对抗模型Adv-G-C对图像分类模型没有额外的要求,即不需要改变现有的图像分类模型来适应本节提出的对抗模型.

为实现对抗,需要为Adv-G-C模型设计合理的损失函数.设$\{s_i, x_i\}_{i=1}^{B}$为训练集,B为一个批次训练集的数量.Adv-G-C的损失函数L_{adv}定义为

$$L_{adv}=\alpha L_G-\beta L_C \tag{4}$$

式中:α和β表示损失函数中L_G和L_C的相对重要性.L_G表示重构子网G的重构损失,定义为

$$L_G=\frac{1}{2B}\sum_{i=1}^{B}d(G(s_i),x_i) \tag{5}$$

式中:$d(\cdot,\cdot)$为距离函数.最小化L_G将保证图像的重构质量.L_C是分类损失,定义为

$$L_C=\frac{1}{B}\sum_{i=1}^{B}l_{ce}(C(G(s_i)),Y_i) \tag{6}$$

式中:$l_{ce}(\cdot,\cdot)$表示计算交叉熵的函数,Y_i是图像x_i对应的真实分类值.

注意式(4)中的"减"号,这使得分类损失L_C越大,L_{adv}越小.这就保证了在对抗模型Adv-G-C下,G网络会尽力输出一种重构图像,试图让图像分类模型C识别出错.也就是说,重构图像同时也是针对图像分类模型C的一个对抗样本.

2.2.2　生成对抗模型

在Adv-G-C的基础上,本小节提出一种更一般化的生成对抗模型Adv-G.Adv-G不只是针对模型C,而是能对抗任意的图像分类模型.

Adv-G模型如图3所示.在此模型中,压缩感知G网络称为重构样本的生成网络,生成重构图像.同时,我们引入一个新的被称为区分网络的D网络.D网络由子网络D_{rec}和子网络D_{rf}组成,D_{rec}和D_{rf}有相同的前面一部分卷积层,最后一层则由各自专属.D网络有两个设计目标:一是区分真图和假图,真图是训练集中的图像x_i,假图是生成网络G网络输出的重构图像x_i';二是区分真图和假图的标签,将真图分类到正确的类别中,但将假图归类为错误的类别.也就是说,D网络一方面让G网络生成高质量的重构图像,

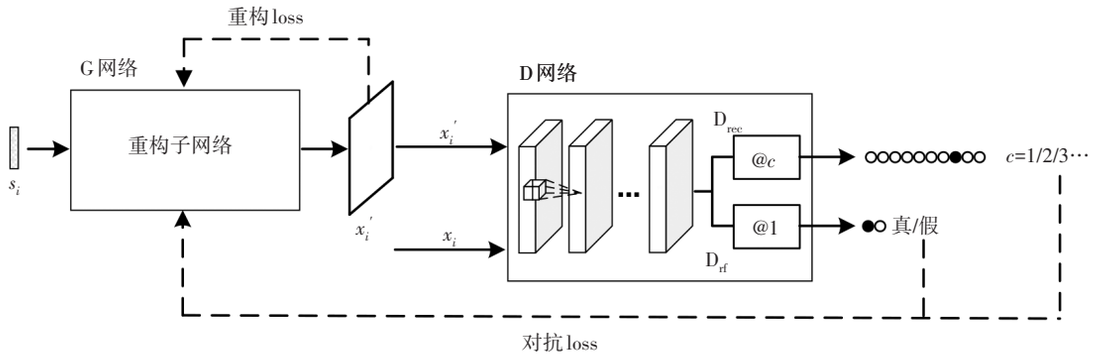


图3 生成对抗压缩感知网络模型Adv-G

Fig.3 Generative adversarial Adv-G model

另一方面则让该图像被分类错误,从而实现对抗.在本文的实验部分,D_{rec}和D_{rf}的网络层结构各自设计成4层,卷积核大小为5×5.

为此,Adv-G模型中D网络的损失函数L_{tot}定义为

$$L_{D_{tot}} = \lambda_D \cdot L_{D-D_{rf}} + \gamma_D \cdot L_{D-D_{rec}}, \tag{7}$$

式中:λ_D和γ_D分别表示网络D_{rec}和D_{rf}在整个损失函数中所占的比重.在式(7)中,子网D_{rf}的损失函数定义为

$$L_{D-D_{rf}} = \frac{1}{B}\sum_{i=1}^{B}(l_{ce}(D_{rf}(x_i),1) + l_{ce}(D_{rf}(s_i),0)) \tag{8}$$

式中:1代表真图,0代表假图,$l_{ce}(\cdot,\cdot)$的定义和式(6)相同.子网D_{rec}的损失函数则定义为

$$L_{D-D_{rec}} = \frac{1}{B}\sum_{i=1}^{B} l_{ce}(D_{rec}(x_i), Y_i) \tag{9}$$

即D_{rec}将真图归类为真实标签Y_i的概率.

Adv-G模型中G网络的损失函数定义为

$$L_{G_{tot}} = \alpha \cdot L_G + \lambda_G \cdot L_{G-D_{rf}} - \gamma_G \cdot L_{G-D_{rec}}, \tag{10}$$

式中:α、λ_G和γ_G分别表示G网络在整个损失函数中所占的比重,以及D_{rec}和D_{rf}对G网络的影响程度.式(10)由三个子项组成,其中L_G与式(5)的定义一致,用于保证图像的重构质量.$L_{G-D_{rf}}$定义为

$$L_{G-D_{rf}} = \frac{1}{B}\sum_{i=1}^{B} l_{ce}(D_{rf}(G(s_i)),1), \tag{11}$$

即G网络试图让D网络将生成图像$G(s_i)$识别为真图.$L_{G-D_{rec}}$定义为

$$L_{G-D_{rec}} = \frac{1}{B}\sum_{i=1}^{B} l_{ce}(D_{rec}(G(s_i)), Y_i), \tag{12}$$

即G网络同时试图让D网络将生成图像$G(s_i)$分类到正确的标签.注意到式(10)中第三项$L_{G-D_{rec}}$取值为负.也就是说,最小化$L_{G_{tot}}$会使得D网络分类错误,这正是我们设计Adv-G模型的目的.不像Adv-G-C模型,Adv-G模型不针对任何特定的图像分类网络,希望D_{rec}和D_{rf}的引入能实现泛化的效果.为此,在训练过程中,G网络、D_{rec}、D_{rf}交替训练,努力保证图像重构质量,同时让D_{rec}网络分类错误,达到对抗的目的.

2 实验与分析

以最新的深度学习压缩感知算法CSNet[16]为例,本小节阐述对抗模型Adv-CSNet-C和生成对抗模型Adv-CSNet的对抗性能.在实验中,两个经典的分类网络VGG-16和ResNet-110被用来测试原始图像和相应对抗样本的识别率.更低的识别率表明更高的对抗性能,从而达到更好的图像隐私保护效果.为叙述简单起见,后续描述中分别略去了这两个分类网络的网络层数16和110.实验平台采用Tensorflow1.14,硬件配置为NVIDIA GeForce GXT 1070单GPU, Intel Core i7-4790K 4.00GHz单处理器,配备32GB内存.

2.1 训练过程

训练数据集为MNIST和Cifar-10数据集.MNIST数据集是由手写数字图像和它们相对应的标签共同组成,共10个类别,分别对应阿拉伯数字的0~9.MNIST一共包含55 000张训练图像和10 000张测试图像,每张图像为28×28大小的灰度图.Cifar-10数据集则包含10个类别的图像,分别是“飞机”“汽车”“鸟”“猫”“鹿”“狗”“青蛙”“马”“船”和“卡车”.Cifar-10数据集用50 000张图像用于网络模型的训练集,剩下的10 000张图像用于模型的测试集.基于本实验平台,在Cifar-10数据集上训练一个Adv-G-C模型约需12 h,其中训练一个采样率完成2个epoch需

要 1 min,模型大约训练 500 个 epoch,耗时约 4 h,共训练了 3 个采样率,总计 3×4=12 h. 训练 Adv-G 模型的时间类似.

为了便于图像压缩性能和图像分类性能的比较,所有训练图像均处理成灰度图,并且将其原始标签采用热编码形式表示.我们也对网络的输入进行了统一的处理,将输入进网络的图像进行归一化处理,使灰度值范围从[0,255]区间线性映射到[0,1]区间.为公平比较性能,CSNet 网络分别用 MNIST 和 Cifar-10 重新训练,且训练过程中超参数的设置与文献[16]保持一致.同样地,我们用这两个数据集重新训练了两个分类网络 VGG 和 ResNet,获取这两个网络对原始图像的初始分类准确率.

对于 2.2.1 小节式(4)中的 α 和 β 参数,均设置为 1,式(7)中的参数 λ_D 设为 1 而 γ_D 设为 0,式(10)中的 α、λ_G 和 γ_G 则分别设置为 1、1 和 0.1.本实验设置 $\alpha=\beta=1$,原因是本文将图像重构效率和识别效率看成同等重要,因此设置了相同大小.我们设置 $\alpha=\lambda_G=1$ 但 $\gamma_G=0.1$,将 D 网络对 G 网络的影响降低了一个数量级,主要是考虑 D 网络和 G 网络在对抗训练时能逐渐增强,以达到训练效果.在对抗模型 Adv-CSNet-C_k 的训练过程中,我们预先训练好 CSNet 模型参数,以避免对抗模型陷入局部最优解.实验使用 Adam 优化器,其中两个参数 beta1 和 beta2 分别设置为 beta1=0.9,beta2=0.999. 使用 Cifar-10 数据集进行训练时,训练周期设置为 300,每个周期迭代 200 次,批量大小为 250. 前 125 个周期的学习率设置为 0.001,126 到 225 个周期的学习率设置为 0.000 1,剩下的训练周期学习率设置为 0.000 01. 使用 MNIST 数据集进行训练时,训练周期设置为 500,每个周期迭代 220 次,批量大小为 250. 前 200 个周期的学习率设置为 0.001,201 到 350 个周期的学习率设置为 0.000 1,剩下的训练周期学习率设置为 0.000 01. 对于 MNIST 的训练,我们先将式(4)中的 α 和 β 参数均设置为 1,待训练过程中训练集的分类准确率降低到 0.1 之下,再修改 α 为 1,β 为 0.1,目的是将 MSE 重构 loss 和交叉熵分类 loss 平衡在一个数量级上,这有利于稳定重构图像质量.

2.2　实验结果

表 1 给出了重构图像的平均识别精度对比结果,数据集为 Cifar-10,识别率数值是 VGG 和 ResNet 的识别率平均值,值越小表示对抗性能越好.从表 1 可以看出,我们提出的对抗方案有明显更低的识别率,而 Adv-CSNet 则有最低的识别精度.这表明我们提出的两种对抗学习方案均获得了预期的对抗性能.其中,生成对抗模型 Adv-CSNet 因为不针对指定的识别算法,平均识别率下降了 74.7%,获得了更好的对抗性能.

表 1　平均识别率对比

Tab.1　Comparison of recognition accuracy　　%

模型	采样率			
	0.1	0.2	0.3	平均
CSNet	28.92	50.76	69.43	49.70
Adv-CSNet-VGG	10.82	13.51	15.84	13.39
Adv-CSNet-ResNet	19.56	27.85	34.61	27.34
Adv-CSNet	**12.06**	**11.06**	**14.52**	**12.55**

表 2 测试对抗模型 Adv-CSNet-C 和 Adv-CSNet 的泛化能力,评价指标是识别精度,测试数据集为 Cifar-10. 从表 2 可以看出,对抗识别算法 VGG 的模型 Adv-CSNet-VGG 在 ResNet 上也有很低的重构图像识别率,表明该模型能迁移到不同的识别算法中,具有泛化性.但是,对抗 ResNet 的模型 Adv-CSNet-ResNet 在 VGG 上的重构识别率却明显高于 ResNet,这说明模型 Adv-CSNet-C 的泛化能力受到限制.从表 2 也容易看出,生成对抗模型 Adv-CSNet 对 VGG 和 ResNet 都取得了很低的识别率度,即 Adv-CSNet 比 Adv-CSNet-C 有更好的泛化性能.

表 2　泛化能力对比

Tab.2　Comparison of generalization capacity　　%

模　型		采样率		
		0.1	0.2	0.3
Adv-CSNet-VGG	VGG	10.69	12.95	14.75
	ResNet	10.95	14.07	16.92
Adv-CSNet-ResNet	VGG	25.91	42.94	53.36
	ResNet	13.20	12.76	15.86
Adv-CSNet	VGG	11.60	9.01	13.55
	ResNet	12.51	13.11	15.49

表 3 对比 CSNet 和相应对抗方案的重构图像质量,数据集为 Cifar-10. 从表 3 中的数据可以看出,相比于原始的 CSNet,我们提出的对抗方案 Adv-CSNet-C 和 Adv-CSNet 在重构质量方面均有所降低.相比之下,生成对抗模型降低幅度稍大,为 10.6%. 这

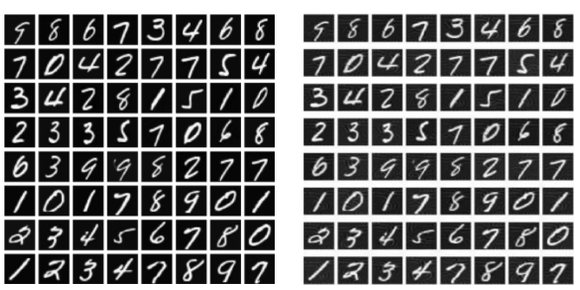
表明,我们提出的方案在取得对抗性能的同时,损失了一定的图像重构质量.

表3 图像重构质量对比(PSNR,dB)

Tab.3 Comparison of image reconstruction quality %

模型	采样率			
	0.1	0.2	0.3	平均
CSNet	25.01	28.51	31.19	28.24
Adv-CSNet-VGG	24.95	28.40	30.50	27.95
Adv-CSNet-ResNet	24.63	28.07	30.51	27.74
Adv-CSNet	23.75	25.52	26.45	25.24

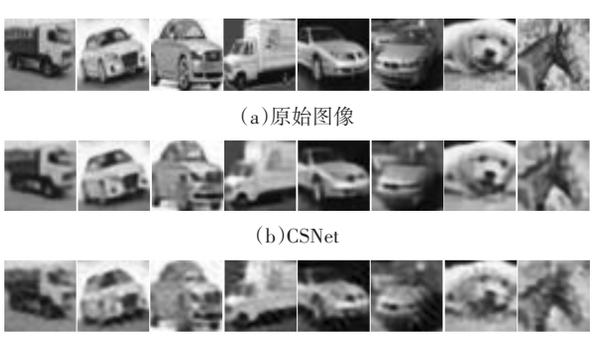
图4和图5展示对抗模型重构图像的视觉效果.图4以对抗模型Adv-CSNet-ResNet为例,数据集为手写数字MNIST,图5则是Adv-CSNet生成的重构图像,数据集为Cifar-10.采样率均为0.1.从图4可以看见,我们提出的对抗方案也能清晰地重构出手写数字,视觉效果完全可以接受.从图4可以看出,本文提出的Adv-CSNet重构图像的视觉效果则略差于原始的CSNet模型,这与表3的结果基本一致.



(a)CSNet (b)Adv-CSNet-ResNet

图4 MNIST重构图像视觉效果

Fig.4 Visual effect of the recovered MNIST images



(a)原始图像

(b)CSNet

(c)Adv-CSNet

图5 Cifar-10重构图像视觉效果

Fig.5 Visual effect of the recovered Cifar-10 images

从上述实验结果可以看出,相比于原始的CSNet,本文提出的对抗模型达到了预期的对抗性能.主要原因是我们在设计对抗模型时将对抗损失通过损失函数反向传播给生成网络,使得生成网络在训练的过程中自动学习对抗性能,从而达到对抗的效果.这样,模型生成的图像能够欺骗识别算法,使之识别错误,从而达到图像隐私保护的效果.

3 结 论

我们将安全性融入压缩感知重构算法,算法生成的重构图像同时也是一个对抗样本,对抗机器分类模型.本文基于对抗样本的思想,提出了一种新的深度学习图像压缩感知方法.降低分类模型的识别率,从而起到保护图像隐私的效果.也就是说,分类模型会识别错误,从而起到保护图像隐私的效果.实验结果表明,我们提出的对抗模型在Cifar-10数据集中有10.6%重构图像PSNR下降,但获得了74.7%的对抗性能提升.同时,我们提出的对抗模型体现出了较好的泛化能力,能对抗不同的分类模型.

参考文献

[1] DONOHO D L. Compressed sensing[J]. IEEE Transactions on Information Theory,2006,52(4):1289-1306.

[2] 戴琼海,付长军,季向阳. 压缩感知研究[J]. 计算机学报,2011,34(3):3425-3434.
DAI Q H,FU C J,JI X Y. Research on compressed sensing[J]. Chinese Journal of Computers,2011,34(3):3425-3434.(In Chinese)

[3] JIANG Q R,LI S,ZHU Z H,*et al*. Design of compressed sensing system with probability-based prior information[J]. IEEE Transactions on Multimedia,2020,22(3):594-609.

[4] METZLER C A,MALEKI A,BARANIUK R G. From denoising to compressed sensing [J]. IEEE Transactions on Information Theory,2016,62(9):5117-5144.

[5] MA Y T,ZHU J N,BARON D. Approximate message passing algorithm with universal denoising and Gaussian mixture learning [J]. IEEE Transactions on Signal Processing, 2016, 64(21): 5611-5622.

[6] AHN J H. Compressive sensing and recovery for binary images [J]. IEEE Transactions on Image Processing, 2016, 25(10): 4796-4802.

[7] WARNELL G, BHATTACHARYA S, CHELLAPPA R, *et al*. Adaptive-rate compressive sensing using side information [J]. IEEE Transactions on Image Processing, 2015, 24(11): 3846-3857.

[8] ZHOU S W,XIANG S Z,LIU X T,*et al*. Asymmetric block based

compressive sensing for image signals[C]//2018 IEEE International Conference on Multimedia and Expo (ICME). San Diego, CA, USA: IEEE, 2018: 1-6.

[9] ZHOU S W, CHEN Z N, ZHONG Q, *et al*. Block compressed sampling of image signals by saliency based adaptive partitioning[J]. Multimedia Tools and Applications, 2019, 78(1): 537-553.

[10] METZLER C A, MOUSAVI A, BARANIUK R G. Learned D-AMP: principled neural network based compressive image recovery [C]//Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS). 2017: 1770-1781.

[11] ZHANG J, GHANEM B. ISTA-net: interpretable optimization-inspired deep network for image compressive sensing [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. IEEE: 2018: 1828-1837.

[12] YANG Y, SUN J, LI H B, *et al*. ADMM-CSNet: a deep learning approach for image compressive sensing[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42 (3): 521-538.

[13] KULKARNI K, LOHIT S, TURAGA P, *et al*. ReconNet: non-iterative reconstruction of images from compressively sensed measurements [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016: 449-458.

[14] LOHIT S, KULKARNI K, KERVICHE R, *et al*. Convolutional neural networks for noniterative reconstruction of compressively sensed images[J]. IEEE Transactions on Computational Imaging, 2018, 4(3): 326-340.

[15] SHI W Z, JIANG F, ZHANG S P, *et al*. Deep networks for compressed image sensing [C]//2017 IEEE International Conference on Multimedia and Expo (ICME). Hong Kong, China: IEEE, 2017: 877-882.

[16] SHI W Z, JIANG F, LIU S H, *et al*. Image compressed sensing using convolutional neural network [J]. IEEE Transactions on Image Processing, 2020, 29: 375-388.

[17] ZHOU S W, HE Y, LIU Y H, *et al*. Multi-channel deep networks for block-based image compressive sensing [J]. IEEE Transactions on Multimedia, 2021, 23: 2627-2640.

[18] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6): 84-90.

[19] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [C]//Proceedings of the International Conference on Learning Representations (ICLR). 2015.

[20] HE K M, ZHANG X Y, REN S Q, *et al*. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016: 770-778.

[21] TAN M X, LE Q V. EfficientNet: rethinking model scaling for convolutional neural networks[EB/OL]. 2019: arXiv: 1905.11946 [cs. LG]. https://arxiv. org/abs/1905.11946

[22] MOOSAVI-DEZFOOLI S M, FAWZI A, FROSSARD P. DeepFool: a simple and accurate method to fool deep neural networks [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016: 2574-2582.

[23] 张思思,左信,刘建伟. 深度学习中的对抗样本问题[J]. 计算机学报,2019,42(8):1886-1904.
ZHANG S S, ZUO X, LIU J W. The problem of the adversarial examples in deep learning [J]. Chinese Journal of Computers, 2019, 42(8): 1886-1904. (In Chinese)

[24] WU Y F, YANG F, XU Y, *et al*. Privacy-protective-GAN for privacy preserving face de-identification [J]. Journal of Computer Science and Technology, 2019, 34(1): 47-60.