



计算机应用
Journal of Computer Applications
ISSN 1001-9081, CN 51-1307/TP

《计算机应用》网络首发论文

题目：基于改进 DCGAN 数据集增强方法
作者：甘岚，沈鸿飞，王瑶，张跃进
收稿日期：2020-07-20
网络首发日期：2020-10-16
引用格式：甘岚，沈鸿飞，王瑶，张跃进. 基于改进 DCGAN 数据集增强方法. 计算机应用. <https://kns.cnki.net/kcms/detail/51.1307.TP.20201015.1715.017.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。

基于改进 DCGAN 数据集增强方法

甘岚, 沈鸿飞*, 王瑶, 张跃进

(华东交通大学 信息工程学院, 南昌 330013)

(*通信作者电子邮箱 15207112943@163.com)

摘要: 针对小样本数据在深度学习中训练难的问题, 为提高 DCGAN 训练效率, 提出了一种改进的 DCGAN 算法对小样本数据进行增强。该方法首先使用 Wasserstein 距离替换原模型中的损失模型; 其次在生成网络和判别网络中加入谱归一化, 得到稳定的网络结构; 最后通过极大似然估计算法和实验估算出样本的最佳噪声输入维度, 提高生成样本的多样性。在三个数据集上的实验结果表明: 改进后的 DCGAN 所生成样本的清晰度以及识别率比改进前均得到了明显提高, 其中平均识别率 MNIST 提高了 8%, CelebA 提高了 16.4%, Cartoon 提高了 16.7%, 在几种清晰度评价指标上各数据集也均有不同程度的提高。表明该方法能够有效地实现小样本数据增强。

关键词: 小样本; 数据增强; DCGAN; Wasserstein 距离; 谱归一化; 内在维数

中图分类号: TP391.4

文献标志码: A

Dataset enhancement method based on improved deep convolutional generative adversarial networks

GAN Lan, SHEN Hongfei*, WANG Yao, ZHANG Yuejin

(School of Information Engineering, East China JiaoTong University, Nanchang Jiangxi 330013, China)

Abstract: A modified DCGAN algorithm was proposed in this article to enhance the small sample data so as to solve the training difficulty of small sample data in the in-depth learning and enhance the training efficiency of DCGAN. According to the method, Wasserstein distance was used to replace the loss model in the original model first and then spectral normalization was added in the spanning network and discrimination network so as to acquire a stable network structure. Finally, the optimal noise input dimension of sample was estimated through maximum likelihood estimation and experimental estimate to make the samples generated more diversified. The experimental result of the three data sets indicated that the improved DCGAN could generate samples with higher definition and recognition rate than the DCGAN prior to improvement. In particular, the average recognition rate MNIST, CelebA and Cartoon were improved by 8%, 16.4% and 16.7% respectively. All data sets regarding several definition assessment indexes were higher to different degrees, suggesting that the method could enhance the small sample data dramatically.

Keywords: small sample; data enhancement; DCGAN; Wasserstein distance; spectrum normalization; inherent dimension

0 引言

深度学习(Deep Learning, DL)^[1]作为一种数据驱动技术, 在样本充足的情况下能够取得很好的训练效果。而实际研究中更多的是小样本^[2]数据, 若直接将小样本数据带入深度学习训练, 将会出现严重的过拟合问题。为了改善小样本数据在深度学习中的训练效果, 需要对小样本数据进行处理, 使其能够在深度学习中取得好的训练结果。

处理小样本的主要方法可以归为以下几点: ①统计学习(Statistic Learning Theory, SLT)^[3], SLT 是一种专门研究小样本学习规律的机器学习方法, 但是 STL 在实际实现时有计算困难、泛化性差等问题; ②神经网络模型(Neural Networks, NN)^[4], NN 是近几年比较流行的一种研究小样本方法, 但是 NN 对于数据的依赖性较强, 需要大量的样本才能有较好的结果。③数据增强(Data Augmentation, DA)^[5], DA 的思想是通过对小样本图像进行不同的变换, 从而得到更多的样本, 以此来提高样本的多样性。DA 相比前两种方法, 拥有很好

收稿日期: 2020-07-20; 修回日期: 2020-09-15; 录用日期: 2020-09-27。

基金项目: 国家自然科学基金地区项目(11862006)。

作者简介: 甘岚(1964—), 女, 江西南昌人, 教授, 硕士, 主要研究方向: 模式识别、图像处理; 沈鸿飞(1992—), 男, 安徽滁州人, 硕士研究生, CCF 会员, 主要研究方向: 图像处理; 王瑶(1994—), 女, 安徽淮南人, 硕士研究生, 主要研究方向: 图像识别; 张跃进(1978—), 男, 湖北钟祥人, 教授, 博士, 主要研究方向: 计算生物力学。

的泛化性,也被广泛用于小样本问题的研究,因此,本文采用 DA 的方法对小样本数据进行处理。

DA 方法很多,如 PCA(Principal Component Analysis, PCA)抖动^[6]、随机裁剪^[7-9]、翻转等。这些方法虽然能实现数据增强,但是其效果有限,因为这些方法局限在原图的基础上进行变换,具有重复性,且数据分布单一。2014 年 Goodfellow 等^[10]受博弈模型的启发提出生成对抗网络模型(Generative Adversarial Networks, GAN),随后其广泛应用在图像合成、超分辨率、DA 等问题中。相比早期 DA, GAN 有了很大的改进,主要表现在 GAN 脱离了原图的限制,通过一组随机噪声学习真实样本分布,生产新的样本数据,故 GAN 是解决小样本数据增强问题的有效方法。自 GAN 提出以后,围绕其思想又衍生出很多改进型算法,其中 2015 年 Radford A^[11]等提出的 DCGAN(Deep Convolutional Generative Adversarial Networks, DCGAN)是众多改进算法中较好的一种,其通过引入步长卷积,使其代替原有的感知机层,提高了训练的稳定性以及生成样本的质量,因此在 DA 方面得到了广泛应用。

2019 年 Juan W 等^[12]通过 DCGAN 对帕金森声纹样本进行增强,有效提高了其识别率; Alyafi B 等^[13]通过 DCGAN 实现对乳腺 X 线摄影病变的质量分析;戚永军等^[14]提出通过生成对抗网络实现肺结节图像数据增强,并在论文中使用了 DCGAN 方法。上述文献表明了 DCGAN 在 DA 应用中的可行性和有效性,但是 DCGAN 的应用也存在一些问题,诸如梯度消失、模式崩塌、训练不稳定、样本多样性不足等,这些问题使得 DCGAN 模型在实际训练时难以稳定地进行,无法进行大规模的 DA。

本文从以下几个方面进行改进:①针对 DCGAN 中的梯度消失、模式崩溃问题,采用 W(Wasserstein)^[15]距离代替原模型中损失函数^[16];②针对模型中的卷积网络有参数振荡问题,提出在每层卷积网络中加入谱归一化(Spectral Normalization, SN)^[17];③DCGAN 的噪声输入维度会影响生成样本多样性,为此采用极大似然估计算法估算数据的最佳噪声输入。实验结果表明,使用上述改进方法,DCGAN 生成样本的多样性得到了明显提高、参数振荡、梯度消失、模式崩塌等问题有了较好的改善。

1 DCGAN 原理

与 GAN 一样, DCGAN 是一种无监督式深度学习模型。DCGAN 网络结构主要由生成器 Generator (G)和判别器 Discriminator (D)构成,其核心思想是通过 G 和 D 相互竞争,最终达到纳什均衡^[18]。其中 G 通过采集和学习真实样本数据的数学分布,最终将给定的噪声输入生成新的样本数据。D, 判别器,其作用是判别输入样本是真实样本还是生成的样本,通过这样不断的学习, G 提高自身样本数据的生成能力,最佳的状态是由 G 生成的样本判别器 D 不能判别其真

实来源。相反 D 提高其判别能力,能判别出输入数据为生成样本而不是真实样本。基本结构如图 1, z 为输入的随机噪声、G 为生成器、 $G(z)$ 表示生成的图片、 x 为真实数据、D 为判别器。

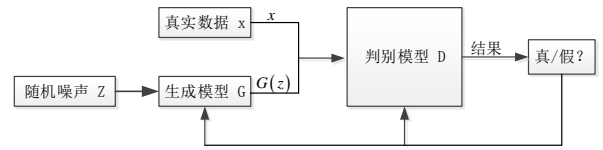


图1 DCGAN 基本结构

Fig. 1 DCGAN basic structure

1.1 训练过程

在模型 G 和 D 对抗训练过程中,使 D 判别输入样本来源于真实数据还是生成样本的正确率最大化,同时使 G 生成的样本数据接近真实数据的程度最大化。训练开始后,随机噪声 z 输入到 G 中开始学习模型参数,此时希望 G 生成的样本数据 $G(z)$ 经过 D 判别后的概率为 1, 数学模型为 $D(G(z))=1$, 也就是 $1-D(G(z))=0$, 即最小化模型 G。而在训练模型 D 时,若输入样本数据为真实数据 x 时,期望目标是经过 D 判别后的概率为 1, 数学模型为 $D(x)=1$; 若输入样本数据为 $G(z)$ 时,则希望经过 D 判别后的概率为 0, 即 $D(G(z))=0$, 所以 $1-D(G(z))=1$, 因此最大化模型 D。故 G 和 D 的训练问题可以用值函数 $V(G,D)$ 表示,从而变成 $V(G,D)$ 的最大化最小化问题。可表示为式(1):

$$\min_G \max_D V(G,D) = E_{x \sim P_{data}(x)} [\lg D(x)] + E_{z \sim P_g(z)} [\lg (1 - D(G(z)))] \quad (1)$$

其中 P_{data} 为真实数据分布, P_g 为噪声分布。实际训练中,模型 G 和 D 交替迭代更新,先固定 G 训练 D, 更新 D 模型中的参数,然后再更新迭代 G 模型的网络参数,不断重复这个过程,直到对抗网络稳定。根据文献[10]中可知,当且仅当 $P_{data} = P_g$ 时, $V(G,D)$ 的最大化最小化问题有全局最优解,即达到纳什均衡状态。此时可以停止训练,模型 G 学习了真实样本数据 P_{data} 的分布,而此时模型 D 的准确率稳定在 0.5。

1.2 损失函数

根据 1.1 节分析, DCGAN 最终转化为值函数 $V(G,D)$ 的极大化极小化问题,其公式(1)即为 DCGAN 的优化目标。但是由于 DCGAN 由生成器模型和判别器模型两部分组成,所以将式(1)拆成两个部分的损失函数模型。在判别器模型中,损失函数如下:

$$\max_D V(D,G) = E_{x \sim P_{data}(x)} [\lg (D(x))] + E_{z \sim P_g(z)} [\lg (1 - D(G(z)))] \quad (2)$$

在式(2)中,前项是真实样本的判别结果,后项是生成样本的判别结果,其中前项越趋近1说明越好,后项则是越趋近0越好。所以最终期望的结果是趋近1,故这里加上对数函数得到判别器损失函数原型。与判别器模型一样,通过处理得生成器损失函数如下:

$$\min_G V(G, D) = E_{z \sim P_g(z)} [\lg(1 - D(G(z)))] \quad (3)$$

同样的使用对数函数作为损失函数。与判别器相比,生成器只需要对生成数据的结果趋近1即可。

1.3 可行性证明

本节将从理论角度证明 DCGAN 方法是可行的,为后面改进提供理论依据。

首先从目标函数出发,由于值函数 $V(G, D)$ 是连续的,故将式(1)写成微积分形式来表示数学期望。如式(4):

$$V(G, D) = \int_{-x}^x P_{data}(x) [\lg D(X)] dx + \int_{-z}^z P_g(z) [\lg(1 - D(G(z)))] dz \quad (4)$$

假设生成器 $G(Z)$ 生成的数据是 x , 分别求出噪声点 z 和微分 dz 的表达式:

$$x = G(z) \Rightarrow z = G^{-1}(x) \quad (5)$$

$$dz = (G^{-1})'(x) dx \quad (6)$$

将式(5)、(6)带入式(4)得:

$$V(G, D) = \int_{-x}^x P_{data}(x) [\lg D(X)] dx + \int_{-z}^z P_g(G^{-1}(x)) [\lg(1 - D(x))] (G^{-1})'(x) dx \quad (7)$$

定义噪声输入 z 的生成分布为 $P_g(x)$, 则可得:

$$P_g(x) = P_z(G^{-1}(x)) (G^{-1})'(x) \quad (8)$$

其中 P_z 表示 z 的生成分布, 将式(8)带入式(7)得:

$$V(G, D) = \int_{-x}^x P_{data}(x) [\lg D(X)] dx + \int_{-x}^x P_g(x) [\lg(1 - D(x))] dx \quad (9)$$

求值函数 $V(G, D)$ 的最大值, 固定 G 求 D 的偏导数得:

$$D^*(x) = \frac{P_{data}(x)}{P_{data}(x) + P_g(x)} \quad (10)$$

从 $D^*(x)$ 的表达式可以看出, 当 G 生成的虚拟样本数据分布和真实数据分布一致时, 即 $P_{data}(x) = P_g(x)$ 时, 在这个条件下 $D^*(x) = 1/2$, 而此时 D 网络已经无法判断出 G 网络生成样本数据的真假性, 即 D 已经取得最优解。

经过推导得到了 D 的最优解, 此时只需证明在 D 取得最优解的同时 G 也能取得最优解, 便说明在目标函数(1)的作

用下, 网络 D 和 G 经过不断的对抗训练能够达到纳什均衡状态。

将 $D^*(x)$ 带入式(9)得:

$$C(G) = \int_{-x}^x P_{data}(x) \left[\lg \frac{P_{data}(x)}{P_{data}(x) + P_g(x)} \right] + P_g(x) \lg \frac{P_g(x)}{P_{data}(x) + P_g(x)} dx \quad (11)$$

其中 $C(G)$ 为 G 的积分表达式, 对式(11)进行等效变换得:

$$C(G) = \int_{-x}^x P_{data}(x) \left[\lg \frac{P_{data}(x)}{\frac{P_{data}(x) + P_g(x)}{2}} \right] + P_g(x) \lg \frac{P_g(x)}{\frac{P_{data}(x) + P_g(x)}{2}} dx - \lg 4 \quad (12)$$

对式(12)引入连续函数的 Kullback Leibler (KL)^[19] 离散度, 整理成离散度表达式得:

$$C(G) = KL \left[P_{data}(x) \parallel \frac{P_{data}(x) + P_g(x)}{2} \right] + KL \left[P_g(x) \parallel \frac{P_{data}(x) + P_g(x)}{2} \right] - \lg 4 \quad (13)$$

其中: $KL(P_1 \parallel P_2) = E_{x \sim P_1} \left(\lg \frac{P_1}{P_2} \right)$

根据离散度定义可知, 当虚拟样本 $P_g(x)$ 的数学分布等于真实数据 $P_{data}(x)$ 的分布时, 即 $P_{data}(x) = P_g(x)$ 时, $KL = 0$, G 网络的最小值是 $-\lg 4$ 。同理引入 Jensen-Shannon (JS)^[20] 离散度得:

$$2JS(P_{data} \parallel P_g) - \lg 4 \quad (14)$$

$$JS(P_1 \parallel P_2) = \frac{1}{2} KL \left(P_1 \parallel \frac{P_1 + P_2}{2} \right) + \frac{1}{2} KL \left(P_2 \parallel \frac{P_1 + P_2}{2} \right)$$

其中:

可以发现, 无论是 KL 离散度还是 JS 离散度 G 网络的最小值都是 $-\lg 4$ 。由此证明在 D 取得最优解^[21]的同时, G 网络也无限逼近其最小值, 即取得最优解。证明, 在目标函数的限制下, 能达到纳什均衡状态, DCGAN 可行。

2 DCGAN 的改进方法

相比其他生成对抗网络算法, DCGAN 在稳定性和生成样本质量方面有一定的优势, 但是梯度消失、多样性不高等问题依然存在。因此本文针对这些问题提出以下改进方法: 损失函数替换、加入谱归一化、最佳噪声输入估算。

2.1 损失函数的改进

2.1.1 问题引出

KL 散度通常被用来衡量分布之间的距离,但 KL 散度是不对称的。这意味着对于同一个距离,观察方式不同,获取的损失值也不同,那么整体 loss^[22]下降的方向就会趋于某个特定方向,这在 DCGAN 中非常容易造成模式崩塌。JS 离散度是另一种衡量分布之间距离的方法,在 KL 离散度的基础上修改而来,解决了不对称问题。但 JS 离散度存在一个严重的问题:两个分布没有重叠时,JS 散度为零,而在训练初期,JS 散度极大可能为 0。所以如果 D 被训练的过强,损失值会经常收敛到 $-2\lg 2$ 而没有梯度。

不难看出,无论是 KL 离散度还是 JS 离散度,DCGAN 都不能稳定的训练出好的结果。

2.1.2 方案选择

根据上述问题,需要找到合理的损失模型替换 KL/JS 散度。与 KL/JS 类似的有欧氏距离、马氏距离、W 距离等。

欧氏距离计算的是同一分布中的两组测量,由定义可知,欧氏距离假设测量的不同维度之间是独立的,而在大多数样本生成时各测量之间有很强的关联性,因此欧氏距离替换损失模型有很大的局限性。马氏距离是在欧氏距离的基础上修改而来,考虑到了各种特性之间的联系,并且尺度是无关的,表示的是数据的协方差距离。但是马氏距离受协方差矩阵不稳的影响,并不能有效快速的计算距离,故马用氏距离替换模型同样存在不稳定性问题。W 距离能稳定并且有效地计算两个分布之间的距离,规避了无相关性、不稳定的问题,这对于前面提出的问题是很好的解决方案。因此采用 W 距离作为 DCGAN 损失函数原型。

2.1.3 Wasserstein 距离相关理论

W 距离又称 Earth-Mover(EM)距离,其定义如下:

$$W(P_{data}, P_g) = \inf_{\gamma \sim \Pi(P_{data}, P_g)} E_{(x,y) \sim \gamma} [\|x - y\|] \quad (15)$$

其中 $\Pi(P_{data}, P_g)$ 是 P_{data} 和 P_g 组合后的所有可能的联合分布的集合。对于可能的联合分布 γ , 可以从 $(x, y) \sim \gamma$ 中得到一个真实样本 x 和生成样本 y , 计算出样本的距离 $\|x - y\|$, 因此 γ 下样本对距离的期望值为 $E_{(x,y) \sim \gamma} [\|x - y\|]$ 。在联合分布中,对这个距离期望值能够取到的下界(式 15)就定义为 W 距离。

W 距离与 KL 散度、JS 散度相比其优势性在于,即使两个分布没有重叠 W 距离仍然能够反映它们的远近。通过一个简单的方法来验证这一点,在二维空间中的两个分布 P_1 和 P_2 , P_1 在线段 MN 上均匀分布, P_2 在线段 PQ 上均匀分布,现在通过参数 λ 控制 P_1 和 P_2 的距离,如图 2。

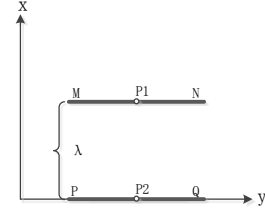


图2 P1, P2 距离分布

Fig. 2 P1, P2 distance distribution

此时很容易得到如下关系式:

$$KL(P_1 // P_1) = \begin{cases} +\infty & \lambda \neq 0 \\ 0 & \lambda = 0 \end{cases}$$

$$JS(P_1, P_2) = \begin{cases} \lg 2 & \lambda \neq 0 \\ 0 & \lambda = 0 \end{cases}$$

$$W(P_1, P_2) = \lambda$$

可以看到 KL 散度和 JS 散度是突变的,而 W 距离是平滑的。当使用梯度下降法优化参数时,KL 散度和 JS 散度无法提供合理的梯度信息,W 距离却可以。同样的,在高维空间中如果两个分布不重叠或者重叠部分可忽略,那么 KL 和 JS 无法直观反映两个分布的远近关系,也不能提供可靠的梯度信息,而 W 距离不仅可以反映出距离的远近,同时也能提供可靠的梯度信息,从理论层面解决了模式崩溃以及梯度消失问题。前期实验也表明,引入 W 距离后模式崩溃问题得到了有效的解决。

2.1.4 改进策略

W 距离的定义中要求解式(15),通过推导给出(15)的变换形式,如下:

$$W(P_{data}, P_g) = \frac{1}{K} \sup_{\|f\|_L \leq K} E_{x \sim P_{data}} [f(x)] - E_{x \sim P_g} [f(x)] \quad (16)$$

(16)式需要满足 Lipschitz^[23]连续。即在一个连续的函数 $f(x)$ 上面施加一个限制条件,要求存在一个常数 K ($K \geq 0$) 使得定义域内的任意两个元素 x_1 和 x_2 之间的距离 $\|f\|_L = |f(x_1) - f(x_2)|$ 都满足 $\|f\|_L \leq K \|x_1 - x_2\|$, 此时称函数 $f(x)$ 的 Lipschitz 常数为 K 。

在式(16)中要求函数 $f(x)$ 的 Lipschitz 常数 $\|f\|_L$ 在小于等于 K 的条件下,所有可能满足条件的 f 取 $E_{x \sim P_{data}} [f(x)] - E_{x \sim P_g} [f(x)]$ 的上界,然后再除以 K 。现在用参数 β 定义一组函数 f_β , 此时式(15)可近似变换成如下形式进行求解:

$$K \cdot W(P_{data}, P_g) \approx \max_{\beta: \|f_\beta\|_L \leq K} E_{x \sim P_{data}} [f_\beta(x)] - E_{x \sim P_g} [f_\beta(x)] \quad (17)$$

f_β 虽然无法包含所有可能, 但是也高度满足 $\sup_{\|f\|_L \leq K} \leq K$ 的要求。最后还要满足 $\|f_\beta\|_L \leq K$ 这个条件, 这里采用一个简单的方法, 限制神经网络 f_β 的所有参数 β_i 的值在范围 $[-n, n]$, 这里 n 可以根据实际训练情况取值, 此时输入样本 x 的导数 $\frac{\partial f_\beta}{\partial x}$ 同样不会超出限定的范围, 所以一定存在一个常数 K 使 f_β 的局部振幅不会超过它, 至此 Lipschitz 连续条件满足。下面即可根据 W 距离模型构建新的损失函数。

在算法实现时, 构造一个含参数 β 、网络的最后一层不是非线性激活层的判别器 f_β , 并且限制 β 不超过限定范围的条件使式(18) 取得最大值, L 就会近似真实分布与生成分布之间的 W 距离。

$$L = E_{x \sim P_{\text{data}}} [f_\beta(x)] - E_{x \sim P_g} [f_\beta(x)] \quad (18)$$

这里需要注意的是 DCGAN 的判别器是做的二分类任务, 所以最后一层加了 sigmoid 函数, 但是在改进 DCGAN 中判别器的作用是近似拟合 W 距离, 属于回归任务, 所以要去掉最后一层的 sigmoid 函数。对于生成器则需要近似最小化 W 距离, 故最小化 L 。这样就得到了改进后 DCGAN 的两个损失函数:

$$-E_{z \sim P_g} [f_\beta(G(z))] \quad (19)$$

$$E_{x \sim P_g} [f_\beta(x)] - E_{x \sim P_{\text{data}}} [f_\beta(x)] \quad (20)$$

式(19)为生成器的损失函数, (20)为判别器的损失函数。这里将式(18)取反得到式(20), 这样可以起到指示训练程度的作用, 损失值越小训练的越好。

2.2 谱归一化

2.2.1 问题引出

在前期的实验中发现, 尽管替换损失函数之后训练的稳定性有了很大的提高, 但是训练中依然会有较大的波动现象。在 2.1 节中为了满足 Lipschitz 连续条件, 直接对参数矩阵中的元素进行限制, 不使其大于限定的值。这种方法简单有效, 但是也存在问题, 因为在满足 Lipschitz 条件的同时, 也使得整个参数矩阵的结构不稳定。通过研究发现谱归一化在满足 Lipschitz 连续的同时, 又可以维持参数矩阵的稳定性。为了更好地发挥 W 距离优势, 提出通过谱归一化改善训练波动的现象。

2.2.2 理论推导

便于分析, 将 DCGAN 看作多层网络。对于多层网络的第 n 层, 输入输出关系可表示为:

$$x_n = a_n(W_n x_{n-1} + b_n) \quad (21)$$

其中, $a_n(\cdot)$ 是所在网络层的非线性激活函数, 可采用 ReLU(Rectified Linear Unit, ReLU)或者 Leaky ReLU; W_n 是网络的参数矩阵, b_n 是网络的偏执项, 这里为了便于推导, 省去偏置项 b_n 。则式(21)可写成:

$$x_n = D_n W_n x_{n-1} \quad (22)$$

D_n 表示对角矩阵, 用于表示激活函数的作用。当输入为负数时对角值为 0, 相反输入为正数时对角值为 1。这样多层网络的输入输出关系便可写成:

$$f(x) = D_n W_n \cdots D_1 W_1 x \quad (23)$$

根据 2.1 节可知, 在满足 Lipschitz 连续条件时对 $f(x)$ 的梯度提出新的要求:

$$\|\nabla_x(f(x))\|_2 = \|D_n W_n \cdots D_1 W_1\|_2 \leq \|D_n\|_2 \|W_n\|_2 \cdots \|D_1\|_2 \|W_1\|_2 \quad (24)$$

这里 $\|W\|$ 表示矩阵 W 的谱范数, 定义为:

$$\sigma(A) := \max_{\|h\| \neq 0} \frac{\|Ah\|_2}{\|h\|_2} = \max_{\|h\|=1} \|Ah\|_2 \quad (25)$$

$\sigma(w)$ 是矩阵 W 的最大奇异值, 对于对角矩阵 D , 有

$$\sigma(D) = \max(d_1, \dots, d_n)。由此式(24)可写成:$$

$$\|\nabla_x(f(x))\|_2 \leq \prod_{i=1}^N \sigma(W_i) \quad (26)$$

为了让 $f(x)$ 满足 Lipschitz 约束, 现对式(24)进行归一化处理:

$$\|\nabla_x(f(x))\|_2 = \|D_n \frac{W_n}{\sigma(W_n)} \cdots D_1 \frac{W_1}{\sigma(W_1)}\|_2 \leq \prod_{i=1}^N \frac{\sigma(W_i)}{\sigma(W_i)} = 1 \quad (27)$$

可见, 只要使每层网络的参数除以该层参数矩阵的谱范数即可满足 Lipschitz 连续的条件。因此谱归一化的引入很好地解决了因直接限制参数矩阵中的元素带来的问题。

2.2.3 实现步骤

在算法实现时, 需要求解 W_i 的奇异值, 这会耗费大量的计算资源, 而这在计算资源有限的情况下是不允许的。为此采用幂迭代法来近似求取, 减少计算资源的消耗。其迭代过程如下:

Step1. 获取随高斯向量 v_i^0

Step2. 以下归一化步骤循环 K 次:

$$\mu_i^k \leftarrow W_i v_i^{k-1}, \text{Normalization: } \mu_i^k \leftarrow \frac{\mu_i^k}{\|\mu_i^k\|},$$

$$v_i^k \leftarrow (W_i)^T \mu_i^k, \text{Normalization: } v_i^k \leftarrow \frac{v_i^k}{\|v_i^k\|},$$

结束循环

Step3. $\sigma_l(W) = (u_l^k)^T W v_l^k$

其中 μ^k 是该矩阵的最大奇异值对应的特征矢量。得到谱范数后, 每个参数矩阵上的参数都除以它, 从而达到归一化的目的。实验表明在加入谱归一化后, 改进的 DCGAN 训练更加稳定, 并且生成的图像质量得到明显的提高。

2.3 最佳噪声输入估算

2.3.1 问题引出

噪声的存在是为了让网络具有随机性, 能产生均匀分布, 从而可以对此进行采样。使用一个简单噪声 z , 经过生成器训练可以产生复杂的分布 P_{data} , 目标是让它接近真实数据分布。前期实验表明: 若噪声输入维度低于某个值时, 生成样本单一, 多样性低; 若噪声输入维度不断增加, 模型生成样本多样性会逐渐变好, 但是训练所需时间也越来越长。因此如何设置噪声输入维度是生成具有多样性样本的关键, 对此本文采用极大似然估计算法估算最佳噪声输入维度。

2.3.2 相关理论

通常可以认为图像、文字、音频等数据是分布在低维流^[24]形空间上的。对于 DCGAN 而言, 噪声的维数至少要达到数据流形的内在维数, 才能产生足够多样性的生成样本。下面给出内在维数的定义: 欧式空间 (\mathfrak{R}^n) 的子集称为一个内在维数 $m = m(M)$, 具有 P -光滑结构的流形, 如果存在常数 $c_p(M)$, 使对于任意给定的 $x \in M$, 存在一组 m 个向量 $v_1(x), v_2(x), \dots, v_m(x) \in \mathfrak{R}^n$,

$$\inf_{\gamma \in \mathfrak{R}^n} \left\| x - x - \sum_{j=1}^m \gamma_j v_j(x) \right\| \leq c_p(M) \|x - x\|^{1+p}$$
对任意的 $x \in M$ 都成立。

如果对任意点, 都能用 m 个附近的点进行逼近, 那么流形的内在维数就是 m 。需要注意的是, 与向量空间的维数定义不同, 向量空间要求基底是固定的, 而流形的基底是局部的, 因点而异。隐空间维数不能太低, 太低了容易丢失模式, 也会产生模式崩溃。也就是说, 隐空间的维数有个下界, 高于这个下界才能避免模式丢失的问题。这个下界就是流形的内在维数。

2.3.3 极大似然估计

一个合适的噪声输入对生成样本的多样性以及质量有着很大的影响。本文采用极大似然估计的方法来找到最佳的噪声输入。2004 年 Levina 等^[25]人提出通过极大似然估计内在维数的方法, 而 DCGAN 的噪声输入即为流形中的内在维数, 为本方法提供了理论依据。设来自流形 M 的样本点 $\{x_j\}_{j=1}^k$, 则内在维数 m_k 可表示为:

$$m_k(x) = \left[\frac{1}{k-1} \sum_{j=1}^{k-1} \lg \frac{T_k(x)}{T_j(x)} \right]^{-1} \quad (28)$$

其中, $T_k(x)$ 表示 x_i 的 k 近邻与 x_i 之间的距离。实验表明, 隐空间维数越高, 会有更多的信息对内在维度做出贡献, 即噪声输入的维数越高, 生成样本的多样性会越好。但是隐空间维数太高占用的显存会更大, 模型更强, 搜索空间更大, 存在更多的鞍点, 学的越慢, 因此不能无限的取大。

2.4 改进后模型可行性证明

根据上述三点改进方案对模型的可行性是否有影响进行论证说明。

① 损失函数替换: 下面给出改进后模型的目标函数:

$$WGAN: L = \max_D E_{x \sim P_{data}} f_w(x) - E_{z \sim P_z} f_w(g_\theta(z)) \quad (29)$$

其中 g_θ 为生成分布, 记 $G = g_\theta, D = f_w$, 则式(29)变为:

$$L = \max_D E_{x \sim P_{data}} D(x) - E_{z \sim P_z} D(G(z)) \quad (30)$$

可以看到, 改进后模型的目标函数和原模型的差别既是取对数与否, 两者只相差一个常数。

为进一步说明, 现假设 \Pr 是任一分布, P_θ 是 $g_\theta(z)$ 的分布, 其中生成分布 Z 的概率密度为 $P(z)$, 局部 Lipschitz 常数为 $L(\theta, z)$, 且 $E_{z \sim P(z)} L(\theta, z) < +\infty$, 则对于式(17)存在解 $f: \mathfrak{R} \rightarrow \mathfrak{R}$ (实数集), 因此 W 距离下的模型是有解的。证明了替换损失函数后模型是可行的。

② 谱归一化: 谱归一化的作用是对模型中结构参数进行限制, 属于过程参数的优化, 因此在模型的可行性证明中对目标函数无影响。

③ 噪声输入: 噪声输入估算仅是对输入参数进行优化, 即目标函数公式中参数 z , 而 z 的变化不影响模型的可行性。综上所述, 改进后的 DCGAN 模型是有效可行的。

2.5 改进的 DCGAN 算法

根据上述损失函数的替换、卷积层引入谱归一化、噪声输入的估算等改进, 给出改进后的 DCGAN 模型和算法。

2.5.1 模型

根据前面三节可知, 对 DCGAN 的改进主要体现在损失函数、噪声输入维度大小以及网络中的归一化改变。噪声输入维度会因数据集的不同而改变, W 距离为反向传播提供梯度信息, 如图 3 DCGAN 模型及训练结构所示。谱归一化引入细节见表 1 和表 2, SN 表示谱归一化。

表1 DCGAN 生成网络细节

Tab. 1 Details of DCGAN generation network				
网络层	步长	卷积核	激活函数	SN
输入层	0	4×4	ReLU	否

Conv1	2	5×5	ReLU	是
Conv2	2	5×5	ReLU	是
Conv3	2	5×5	ReLU	是
Conv4	2	5×5	ReLU	是
输出层	2	5×5	tanh	否

表2 DCGAN 判别网络细节

Tab. 2 Details of DCGAN discriminates network

网络层	步长	卷积核	激活函数	SN
输入层	2	5×5	Leaky ReLU	否
Conv1	2	5×5	Leaky ReLU	是
Conv2	2	5×5	Leaky ReLU	是
Conv3	2	5×5	Leaky ReLU	是
Conv4	2	5×5	Leaky ReLU	是
输出层	0	4×4	Leaky ReLU	否

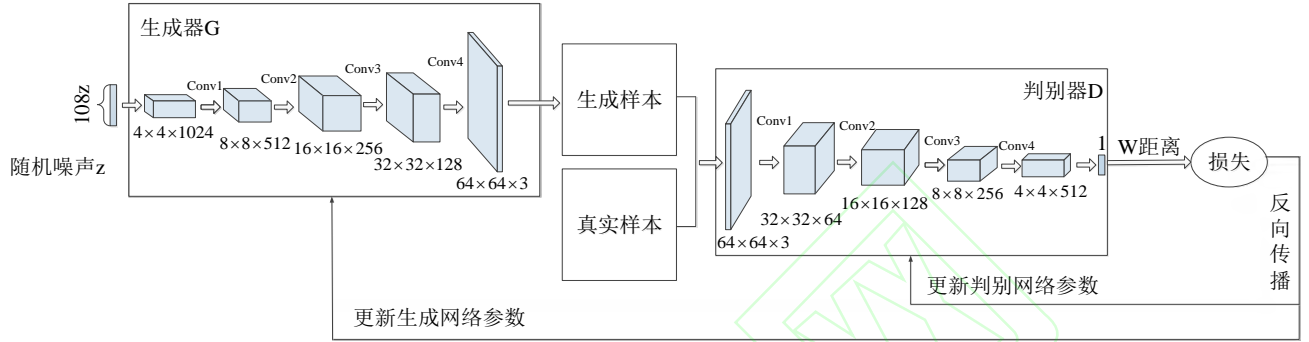


图3 改进后模型及训练结构

Fig. 3 Improved model and training structure

2.5.2 算法

根据前面改进模型，给出改进后 DCGAN 的算法，伪代码如下：

算法：改进后的 DCGAN 伪代码。

参数说明： ∇ 表示梯度算子；

```

1: While set_steps do:
2:   for k_steps do:
3:     在真实样本  $P_{data}(x)$  中采用  $m$  个真实样本  $\{x^i\}_{i=1}^m$ 
4:     在噪声分布  $P_z(z)$  中采样  $m$  个噪声样本  $\{z^i\}_{i=1}^m$ 
5:     根据梯度更新判别器:
6:     
$$\nabla_m \left[ \frac{1}{m} \sum_{i=1}^m f_{\beta}(x^i) - \frac{1}{m} \sum_{i=1}^m f_{\beta}(G(z^i)) \right]$$

7:   end for
8:   在噪声分布  $P_z(z)$  中采样  $m$  个噪声样本  $\{z^i\}_{i=1}^m$ 
9:   根据梯度更新生成器:
10:  
$$-\nabla_m \frac{1}{m} \sum_{i=1}^m f_w(G(z^i))$$

11: end while

```

其中 set_steps 和 k_steps 是提前设置好的超参数。set_steps 是最大的迭代次数，以保证足够的训练次数来达到训练要求。k_steps 是判别器的训练次数，一般是通过多次更新判别器的方法来保持两个网络训练平衡。

3 实验过程及分析

3.1 实验环境及参数设置

对于所提出的算法有效性验证，硬件平台在配置为 Intel(R)Core(TM)i7-9700k-3.60GHZ 处理器，运行内存 16GB，12GB 的 RTX2080TI (单个) 显卡，Windows 10 操作系统下进行。代码实现在 2018.3 x64 版本的 pycharm，3.6.3 版本的 python，1.14.0 版本的 TensorFlow 下进行。

实验发现 0.0001 的学习率收敛效果更好。故将学习率设置为 0.0001、Batch Size 大小为 128、迭代训练次数为 15000 轮。使用的是随机梯度下降 (Stochastic Gradient Descent, SGD)，Adma 优化器。数据集选用的是公共数据集 MNIST、CelebA 以及动漫人脸。

3.2 训练结果及分析

3.2.1 损失变化

对损失变化部分，本文取前 15000 轮实验数据作对比。根据实验绘制了生成器损失变化曲线，其中 Or_DCGAN(Original DCGAN) 为原 DCGAN 模型，IM_DCGAN(Improved DCGAN) 为改进后模型。Or_DCGAN2 曲线是实验中常出现的一种情况，在 10000 轮之前损失值基本无变化，没有任何梯度信息，随后虽然有梯度信息，但是随着训练的不断进行，其参数振荡幅度越来越大，呈发散趋势无法收敛。Or_DCGAN1 曲线是另一种训练中常出现的情况，3000 轮以前训练相对平稳，但随后振荡幅度开始加大，呈现发散没有收敛趋势，此种现象无法训练出理想的样本。而改进后模型 IM_DCGAN 在训练到 1000 轮的时候损失值快速下降，并且在后面的训练中整体呈稳定趋势，波动幅度较小，训练 15000 轮以后生成的样本多样性、清晰

度都比改进前有明显提高。变化曲线如图 4 训练损失对比所示。表明对 DCGAN 的改进方法有效。

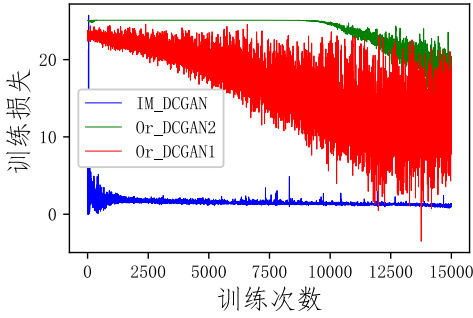


图4 训练损失对比

Fig. 4 Comparison of training losses

3.2.2 谱归一化影响

为说明谱归一化产生的影响,采用单一变量原则,除卷积层中是否加入谱归一化外,其余变量、参数均保持一致,

其中 SN 表示加入谱归一化, N_SN 表示没用加入谱归一化。可见在没有加入谱归一化时,前期训练基本稳定,但是从 8000 轮以后,波动逐渐加大,呈发散趋势,难以稳定训练出理想结果。而在加入谱归一化之后,可以明显看到整个训练过程比较稳定,无大幅度振荡现象,如图 5 谱归一化对生成器损失影响。表明该方法能有效地改善参数振荡问题。

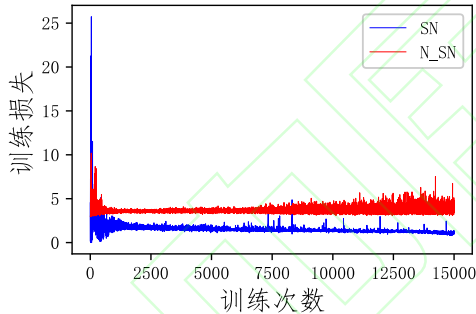


图5 谱归一化对损失的影响

Fig. 5 Effect of spectral normalization on loss

3.2.3 噪声输入对样本的影响

为估算样本的内在维数,实验随机抽取 3000 个样本通过极大似然算法进行计算。三种数据集的内在维数变化曲线如图 6 所示, m 为数据的维数变化, K 为样本点。其中 MNIST 的内在维数约为 6.8、Cartoon 的内在维数约为 8.3、CelebA 的内在维数最大约 10.5。需要说明的是,实际样本的内在维数远不止所估算的这么低,主要由于某些少见的特征样本很少,例如 MNIST 数据集每张样本的底色都是黑色,这些特征就被忽略,对内在维数没有贡献,从而导致实验的内在维数很低。

内在维数的确定是为了给噪声输入一个约定,即噪声输入必须大于内在维数才能生成高质量的样本。现将不同噪声输入生成样本带入卷积神经网络 (Convolutional Neural Networks, CNN) 训练,根据其识别率确定最佳噪声输入维度。随着噪声输入的逐渐增大,生成样本的识别率越来越高,但是增加的幅度越来越小,如图 7 所示,噪声输入为均匀随机分布,大小为 $[batch_size, noise_size]$, 其中 $batch_size$ 为固定值 128, $noise_size$ 即横坐标值。其中 MNIST 在较小的噪声输入时就能生成质量较高的样本,原因是 MNIST 样本复杂程度低,没有冗余信息。而 CelebA 和 Cartoon 直到噪声输入到 80 以后才能生成较高的样本,因为 CelebA 和 Cartoon 冗余信息多复杂程度高,如人物微表情、脸型、背景等,因此实际内在维数较大。

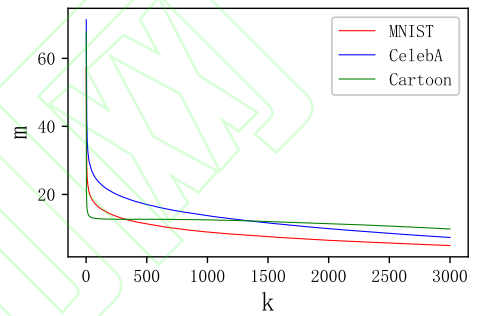


图6 数据内在维数变化曲线

Fig. 6 curve of data intrinsic dimension change

综上所述,为使 DCGAN 生成具有多样性的样本,应该根据内在维数合理设置噪声输入。现根据实验给出噪声参数设置建议,若样本特征简单信息复杂度低,则设置为 100 即可生成高质量样本;若样本信息复杂冗余信息多,则将噪声输入大小设置和样本本身大小相同或者更高。例如 MNIST 训练时噪声输入设置为 100、CelebA 设置为 150、Cartoon 设置为 140。需要说明的是考虑到内存和训练速度的原因,在设置噪声输入时需要根据硬件条件作出相应的调整,不能无限增大。

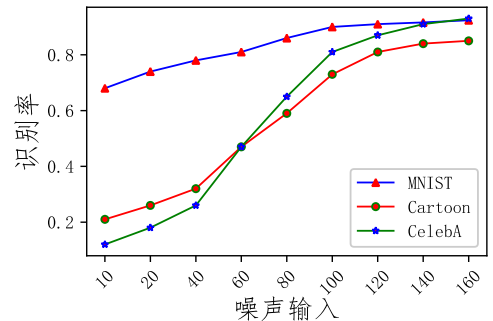


图7 噪声输入对成样本别率的影响

Fig. 7 Influence of noise input on recognition rate of generated samples

3.2.4 生成样本对比

为直观的展示模型改进前后生成样本质量,选取实验最终生成样本对比。改进后 DCGAN 生成的样本辨识度高、轮廓信息清晰、重复样本少。而改进前后生成的样本都比较模糊,学习到的特征少,其中动漫人脸和 CelebA 数据集出现了明显的模式崩塌现象,如图 8 训练过程生成图像所示。表明改进后的 DCGAN 有效地解决了模式崩塌、生成样本多样性不高等问题。

3.3 增强效果分析

对于增强效果分析,本文通过生成样本在 CNN 网络中的识别率和图像质量两个方面进行阐述。

3.3.1 样本量化评价

本文选取灰度方差(Sum of Modulus of gray Difference, SMD)函数、Laplacian^[26]梯度函数、Tenengrad^[27]梯度三种方法评价图像的清晰度,使用弗雷歇距离(Fréchet Inception Distance, FID)^[28]评价样本间的相似度。清晰度方法从生成的样本中随机选取 3000 张图像,然后计算其平均值作为量化评价指标,相似度计算 15000 轮实验 FID 的均值作为评价指标。

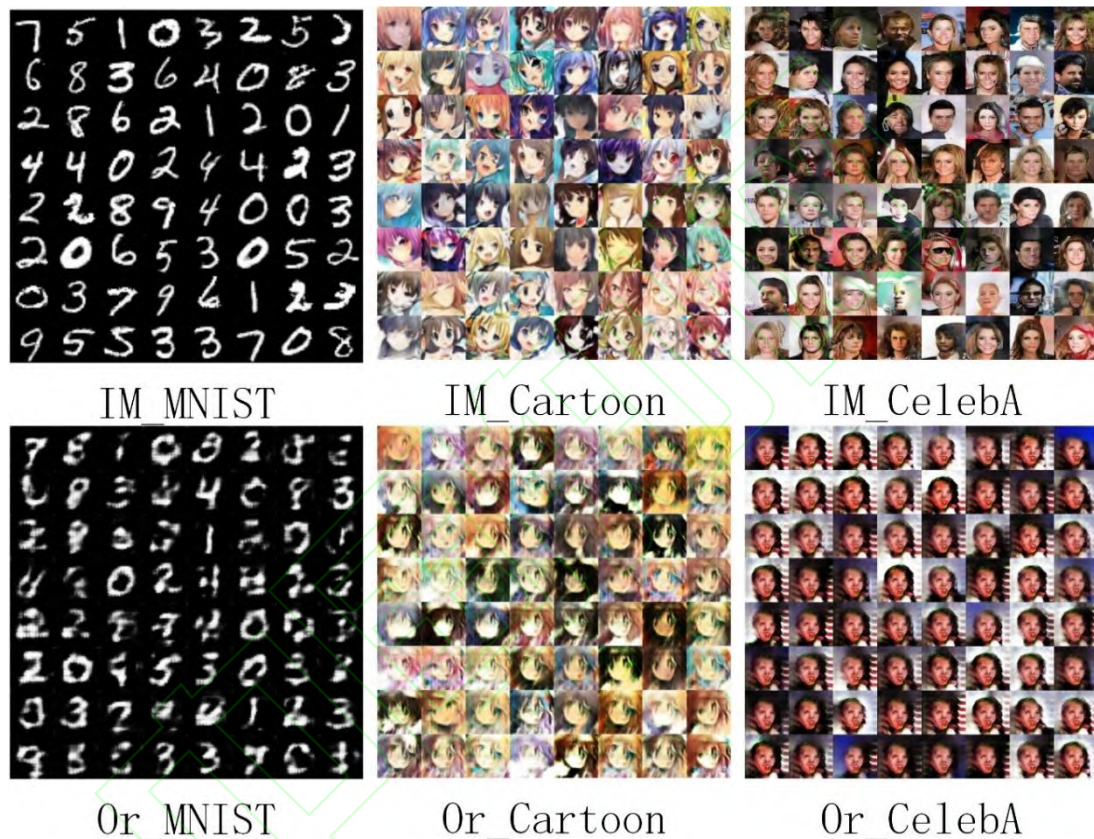


图8 训过程中生成图像对比

Fig. 8 Image comparison during training

表3 样本质量评价指标对比

Tab. 3 Comparison of sample quality evaluation indicators

数据集	训练模型	无参考图像质量评价方法		
		SMD	Laplacian	Tenengrad
MNIST	DCGAN	15.3	40.0	188.3
	IM_DCGAN	38.3	88.3	211.3
CelebA	DCGAN	6.3	16.9	90.1
	IM_DCGAN	28.3	37.6	113.4
Cartoon	DCGAN	10.5	36.4	145.3
	IM_DCGAN	23.1	53.6	167.4

三种方法均是量化后的数值越大,样本的清晰度越高。

容易发现,在各评价方法中,其改进后的 DCGAN 生成样本质量均得到了明显提高,如表 3 所示。表明改进后的 DCGAN 生成样本清晰度更高。

为进一步的展现改进后模型效果,采用 FID 作为客观评价指标。FID 更具有原则性和综合性,能准确的反映生成样本和真实样本间的相似性。

FID 值越小表明样本间相似度越高。可见,改进后 DCGAN 的 FID 值较改进前有明显的降低,其中 MNIST 数据集最为明显,降低了大约 44%,如表 4 所示。表明改进后 DCGAN 模型在提升样本质量上的有效性。

表4 在各数据集上 FID 值比较

Tab. 4 Comparison FID values across data sets

数据	FID 值/无单位	
	DCGAN	IM_DCGAN
MNIST	34.6	19.1 ↓
CelebA	87.8	82.5 ↓
Cartoon	76.3	68.6 ↓

3.3.2 样本在分类识别性能上的分析

前面通过实验验证了改进后的 DCGAN 生成样本的多样性、清晰度都有所提高。但是 DA 的最终目的是使小样本数据能够在深度学习算法上取得好的分类识别效果。所以对于改进后的 DCGAN 的有效性还需要进一步的验证。在生成样本的识别实验中, 各数据集的训练集和验证集通过训练好的模型生成样本, 按 9 比 1 进行随机选取, 最终训练集为 18000 张, 验证集为 2000 张, 测试集从原数据集中随机选取 4500 张; 无生成样本的实验中, 按同等比例在原数据集选取。

表5 生成样本识别率对比

单位: %

Tab. 5 Comparison of recognition rate of generated samples unit: %

数据集	识别率	DCGAN	IM_DCGAN	NO_DCGAN
MNIST	平均	84.1	90.9	91.7
	最高	89.6	94.2	95.8
Cartoon	平均	70.8	82.6	83.2
	最高	76.7	87.5	90.6
CelebA	平均	78.5	91.4	91.6
	最高	86.4	93.6	94.3

DCGAN 改进前后生成样本在 CNN 网络下的训练识别率如表 5 所示, NO_DCGAN 表示不使用生成样本的识别率, 选取的是 GoogLeNet。可以发现, 改进后 DCGAN 的识别率较改进前均有明显提高; 相比没有使用生成样本数据, 改进后的识别率有大幅度降低, 但是依然有较高的识别率。表明改进方法有效, 较好地提高了识别率, 实现了小样本数据在深度学习网络中的稳定训练, 并且有较高的识别率。

4 结语

本文对小样本问题进行分析与研究。首先讨论了处理小样本问题的方法, 提出采用 DA 方法对小样本数据进行增强。然后对比传统 DA 和生对抗网络的优缺点, 提出使用 DCGAN 作为本文 DA 方法。最后分析 DCGAN 存在的问题, 提出一种改进的 DCGAN 对小样本数据进行增强。在理论上, 证明了采用 W 距离作为损失函数、在卷积层中加入谱归一化、通过极大似然估计噪声输入维度三种方法是可行的。在实验上, 通过在不同数据集上对比实验, 得出改进方法确实有效可行,

提高了样本的识别率和 DCGAN 的稳定性。在未来的工作中, 还需更进一步的对小样本的分类识别进行研究, 提高其识别率, 实现小样本数据增强、分类识别一体化的功能。

参考文献

- [1] SCHMIDHUBER J. Deep learning in neural networks[J]. Neural Networks, 2015, 61: 85-117.
- [2] LI D C, WEN I H. A genetic algorithm-based virtual sample generation technique to improve small data set learning[J]. Neurocomputing, 2014, 143(nov.2):222-230.
- [3] HA M H, LI J, TIAN J, et al. The key theorem of learning theory on $g/\text{sub/spl lambda}/\text{measure spaces}[C]//\text{Proceedings of the 2004 International Conference on Machine Learning and Cybernetics (IEEE Cat. No. 04EX826)}. \text{IEEE}, 2004, 3: 1904-1907.$
- [4] 高智勇, 黄金镇, 杜程刚. 基于特征金字塔网络的肺结节检测[J/OL]. 计算机应用: 1-8[2020-08-31]. <http://kns.cnki.net/kcms/detail/51.1307.TP.20200713.1123.008.html>. (GAO Z Y, HUANG J Z, DU C G. Lung nodule detection based on feature pyramid network [J/OL]. Computer application: 1-8[2020-08-31].)
- [5] SALAMON J, BELLO J P. Deep Convolutional Neural Networks and Data Augmentation for Environmental Sound Classification[J]. IEEE Signal Processing Letters, 2017, 24(3): 279-283.
- [6] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]// Proceedings of the 2012 International Conference on Neural Information Processing Systems(NIPS). Lake Tahoe, Nevada: MIT Press, 2012: 1097-1105.
- [7] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]// Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Piscataway: IEEE, 2015, pp. 1-9.
- [8] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[EB/OL]. [2020-9-14] <https://arxiv.org/pdf/1409.1556.pdf>
- [9] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Piscataway: IEEE, 2016, pp. 770-778.
- [10] GOODFELLOW I J, POUGETABADIE J, MIRZA M, et al. Generative adversarial nets [C]// Proceedings of the 27th International Conference on Neural Information Processing Systems Volume 2. December 2014. New York: ACM, 2014: 2672-2680.
- [11] RADFORD A, METZ L, CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks[EB/OL]. (2015-11-19) [2020-9-13]. <https://arxiv.org/pdf/1511.06434.pdf>.
- [12] JUAN W, ZHI-JING X U. Study on Augmentation and Recognition of Parkinson's Voiceprint Samples by HR-DCGAN Method[J]. Journal of Chinese Computer Systems, 2019, 40(09): 2026-2032.
- [13] ALYAFI B, DIAZ O, ELANGO VAN P, et al. Quality analysis of DCGAN-generated mammography lesions[C]// Proceedings of the 15th International Workshop on Breast Imaging (IWBI2020). International Society for Optics and Photonics, 2020, 11513: 115130B.
- [14] 戚永军, 顾军华, 栗位勋, et al. 基于生成对抗网络的肺结节图像数据增强方法[J]. 中国计算机用户协会网络应用分会 2019 年第二十三届网络新技术与应用年会论文集, 2019: 5. (QI Y J, GU J H, LI W X, et al. Lung nodule image data enhancement method based on generative confrontation network[J]. Proceedings of the 23rd annual meeting of network application branch of China Computer Users Association in 2019, 2019: 5.)
- [15] ZHANG Y, FANG Q, QIAN S, et al. Knowledge-aware Attentive Wasserstein Adversarial Dialogue Response Generation[J]. ACM

- Transactions on Intelligent Systems and Technology, 2020, 11(4):1-20.
- [16] HERNÁNDEZ JUAN C, RODRÍGUEZ JOSÉ M, SIGARRETA JOSÉ M. Mathematical Properties of the Hyperbolicity of Circulant Networks[J]. Advances in Mathematical Physics, 2015, 2015:1-11.
- [17] MIYATO T, KATAOKA T, KOYAMA M, et al. Spectral Normalization for Generative Adversarial Networks[EB/OL]. [2020-05-08]. <https://arxiv.org/pdf/1802.05957.pdf>.
- [18] 孙钰, 李林燕, 叶子寒, et al. 多层次结构生成对抗网络的文本生成图像方法[J]. 计算机应用, 2019, 39(11):3204-3209. (SUN Y, LI L Y, YE Z H, et al. Multi-level structure generation against network text generation image method [J]. Computer Applications, 2019, 39(11): 3204-3209.)
- [19] 张变兰, 路永钢, 张海涛. 基于 KL 散度和近邻点间距离的球面嵌入算法[J]. 计算机应用, 2017, 37(03):680-683+690. (ZHANG B L, LU Y G, ZHANG H T. Spherical embedding algorithm based on KL divergence and distance between neighboring points [J]. Computer Applications, 2017, 37(03): 680-683+690.)
- [20] 丁月, 汪学明. 基于改进特征加权的朴素贝叶斯分类算法[J]. 计算机应用研究, 2019, 36(12):3597-3600+3627. (DING Y, WANG X M. Naive Bayes classification algorithm based on improved feature weighting [J]. Application Research of Computers, 2019, 36(12): 3597-3600+3627.)
- [21] GUTJAHR W J. ACO algorithms with guaranteed convergence to the optimal solution[J]. Information Processing Letters, 2002, 82(3): 145-153.
- [22] PALMER C N, IRVINE A D, TERRONKIATKOWSKI A, et al. Common loss-of-function variants of the epidermal barrier protein filaggrin are a major predisposing factor for atopic dermatitis[J]. Nature Genetics, 2006, 38(4): 441-446.
- [23] TOROUS J, LIPSCHITZ J, NG M, ET AL. Dropout rates in clinical trials of smartphone apps for depressive symptoms: A systematic review and meta-analysis[J]. Journal of Affective Disorders, 2020, 263: 413-419.
- [24] YU K, ZHANG T, GONG Y. Nonlinear learning using local coordinate coding[C]//Advances in neural information processing systems. New York: ACM, 2009: 2223-2231.
- [25] LEVINA E, BICKEL P J. Maximum likelihood estimation of intrinsic dimension[C]//Advances in neural information processing systems. New York: ACM, 2005: 777-784.
- [26] 代强, 程曦, 王永梅, et al. 基于轻量自动残差缩放网络的图像超分辨率重建[J]. 计算机应用, 2020, 040(005):1446-1452. (DAI Q, CHENG X, WANG Y M, et al. Image super-resolution reconstruction based on lightweight automatic residual scaling network [J]. Computer applications, 2020, 040 (005): 1446-1452)
- [27] CHERN N N K, NEOW P A, ANG M H. Practical issues in pixel-based autofocus for machine vision[C]//Proceedings of the 2001 International Conference on Robotics and Automation (Cat. No. 01CH37164). Piscataway: IEEE, 2001, 3: 2791-2796.
- [28] ZHANG H, GOODFELLOW I, METAXAS D, et al. Self-attention generative adversarial networks [EB/OL]. [2019-9-14]. <http://www.ivs-group.com/seminar/2019/PDF/LYF20190530.pdf>.

This work was supported in part by the National Natural Science Foundation of China under Grant 11862006.

GAN Lan, born in 1964, M.S., professor. Her research interests include pattern recognition, image processing.

SHEN Hongfei, born in 1992, M. S. candidate. His research interests include image processing.

WANG Yao, born in 1994, M. S. candidate. Her research interests include image recognition.

ZHANG Yuejin, born in 1978, Ph.D., His research interests include Computational Biomechanics.