

# 基于条件梯度 Wasserstein 生成对抗网络的图像识别

何子庆<sup>1</sup>, 聂红玉<sup>2</sup>, 刘月<sup>1</sup>, 尹洋<sup>1</sup>

(1. 西南交通大学 信息科学与技术学院, 成都 610097;

2. 重庆交通职业学院 大数据学院, 重庆 402247)

**摘要:** 生成式对抗网络 GAN 功能强大, 但是具有收敛速度慢、训练不稳定、生成样本多样性不足等缺点; 该文结合条件深度卷积对抗网络 CDCGAN 和带有梯度惩罚的 Wasserstein 生成对抗网络 WGAN-GP 的优点, 提出了一个混合模型—条件梯度 Wasserstein 生成对抗网络 CDCWGAN-GP, 用带有梯度惩罚的 Wasserstein 距离训练对抗网络保证了训练稳定性且收敛速度更快, 同时加入条件  $c$  来指导数据生成; 另外为了增强判别器提取特征的能力, 该文设计了全局判别器和局部判别器一起打分, 最后提取判别器进行图像识别; 实验结果证明, 该方法有效的提高了图像识别的准确率。

**关键词:** 生成式对抗网络; 条件模型; Wasserstein 距离; 梯度惩罚; 全局和局部一致性; 图像识别

## Image Recognition With Conditional Wasserstein Generative Adversarial Networks with Gradient Penalty

He Ziqing<sup>1</sup>, Nie Hongyu<sup>2</sup>, Liu Yue<sup>1</sup>, Yin Yang<sup>1</sup>

(College of Information Science and Technology, Southwest Jiaotong University, Chengdu 610097, China;

School of Big Data, Chongqing Vocational College of Transportation, Chongqing 402247, China)

**Abstract:** Generated adversarial net GAN is powerful, but it has some disadvantages such as slow convergence, unstable training, and insufficient sample diversity. This paper presents a conditional gradient Wasserstein generation confrontation network model CDCWGAN-GP by Combining the advantage of conditional deep convolution adversarial net CDCGAN and Wasserstein generated adversarial net with gradient penalty WGAN-GP. Using the Wasserstein distance training against the network with gradient penalty guarantees training stability and faster convergence, while adding condition  $c$  to guide data generation. In addition, in order to enhance the ability of the discriminator to extract features, the paper designs a global discriminator and a local discriminator to score together, and finally extracts the discriminator for image recognition. The result of simulation experiments show that this method effectively improves the accuracy of image recognition.

**Keywords:** GAN; conditional model; Wasserstein distance; gradient penalty; global and local consistency; image recognition

## 0 引言

图像识别一直是计算机视觉领域内的热门研究, 大多数图像识别的方法都是利用卷积神经网络模型 (CNN) 来实现, 但是卷积神经网络模型需要使用规模巨大的训练数据集来进行训练, 导致网络模型收敛速度很慢, 并且需要很多训练技巧来提高识别率。文献 [1] 提出的生成式对抗网络 (GAN), 已经被证明功能强大, 既可以利用生成器模拟复杂的数据分布, 又可以利用判别器进行特征提取。GAN 应用广泛, 包括语言处理<sup>[2]</sup>、图像修复<sup>[3]</sup>等。但是如何将 GAN 强大的特征提取能力应用到图像识别任务上是一

个具有挑战性的问题。文献 [4] 针对 GAN 不能生成具有特定属性图片的问题提出了条件生成对抗网络 CGAN, 利用数据标签指导数据生成, 构建了有监督的生成对抗网络模型。文献 [5] 针对 GAN 的网络模型训练不稳定的缺点, 提出一种深度卷积生成对抗网络 (DCGAN) 模型, 可以通过无监督学习训练深度卷积生成对抗网络模型, 然后在有监督学习任务中重复利用判别模型的一部分作为特征提取器。条件深度卷积对抗网络 CDCGAN 结合了 CGAN 和 DCGAN 的优点, 在 DCGAN 的基础上引入条件  $y$  指导训练。但是这几种方法, 都没有从根本上解决 GAN 产生的收敛速度慢、训练不稳定、生成样本多样性不足、collapse mode 等问题。

文献 [6] 提出了一种用 Wasserstein 距离代替 GAN 中 JS 散度作为惩罚函数的对抗网络模型 WGAN, 解决了 GAN 训练不稳定问题。文献 [7] 又对 WGAN 模型进行了改进, 提出了带有梯度惩罚的 WGAN: WGAN-GP, 比 WGAN 拥有更快的收敛速度, 也可以产生更高质量的生成

收稿日期: 2018-12-13; 修回日期: 2019-01-10。

基金项目: 国家自然科学基金资助项目 (61461048); 重庆市教委科学技术研究项目 (KJQN201805702); 四川省科技创新苗子工程资助项目 (2018102)。

作者简介: 何子庆 (1993-), 男, 新疆乌鲁木齐人, 硕士研究生, 主要从事数字图像处理、深度学习方向的研究。

样本。

本文结合 CDCGAN 和 WGAN-GP 的优点，提出一种混合网络模型：带有梯度惩罚的条件 Wasserstein 生成对抗网络模型 (CDCWGAN-GP)。首先训练此对抗网络模型，然后从训练好的网络模型中提取判别器 D 来进行图像分类任务。在这里，本文设计了一个具有识别全局和局部一致性的判别器：全局判别器需要完整图像的输入来提取整体图像的特征，局部判别器只需要从小一块区域上提取更加细微的特征，这样可以大大增强判别器提取特征的能力。通过实验结果可以证明，相比于其他类似的网络模型，本文提出的 CDCWGAN-GP 网络模型在图像识别任务中获得了更出色的结果。

## 1 相关工作

### 1.1 生成式对抗网络模型

生成式对抗网络模型 GAN 已经被证明了既可以生成高质量的图像，也拥有强大的特征提取能力。我们用生成器 (G) 代表生成模型，判别器 (D) 代表判别模型<sup>[8]</sup>。在图像识别的任务中，GAN 网络同时训练 G 和 D，生成器 G 把从先验分布  $P_z$  中采样得到的随机矢量  $z$  映射到图像空间；判别器 D 判断输入图像是不是“真实的”，优化公式如下：

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

D 的输入参数  $x$  代表一张图片，输出  $D(x)$  代表  $x$  为真实图片的概率，如果为 1，就代表 100% 是真实的图片，如果为 0，就代表不可能是真实的图片；类似  $D(x)$ ， $D(G(z))$  代表判别器判断由生成器产生的生成图片  $G(z)$  是否为真实图片的概率；这样，G 和 D 构成了一个动态的“博弈过程”。在最理想的状态下，G 可以生成足以“以假乱真”的图片  $G(z)$ ，对于 D 来说，它难以判定 G 生成的图片究竟是不是真实的，此时  $D(G(z)) = 0.5$ 。GAN 模型流程图如下：

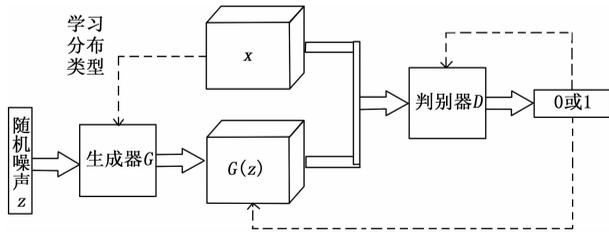


图 1 GAN 模型的流程图

### 1.2 条件生成对抗网络

条件生成对抗网络 CGAN 将属性信息  $y$  融入生成器 G 和判别器 D 中，属性  $y$  可以是任何标签信息，例如图像的类别、人脸图像的面部表情等。优化公式如下：

$$\min_G \max_D V(G, D) = E_{x \sim P_{data(x/y)}} [\log(D(x/y))] + E_{z \sim P_z} [\log(1 - D(G(z/y)))] \quad (2)$$

CGAN 的目标函数类似 GAN 的目标函数，在图像识别任务中， $y$  代表图片的标签信息。 $D(x/y)$  就代表真实输

入  $x$  与标签  $y$  融合后输入到判别器中得到是否为真实数据的概率； $D(G(z/y))$  就代表判别器判断由生成器产生的生成数据与  $y$  融合后的数据是否为真实的概率。

带条件的输入可以指导网络生成我们想要的某种图片信息，同时也相当于是从无监督模型到有监督模型。

### 1.3 条件深度卷积对抗网络

深度卷积对抗网络 DCGAN 是把 GAN 与 CNN 结合起来的网络模型，而条件深度卷积对抗网络 (CDCGAN) 就是把 CGAN 与 DCGAN 结合起来的网络模型。CDCGAN 是在 DCGAN 的框架下，将生成器每层的输入都加上标签  $y$ ，同样将判别器的每层输入也加上标签  $y$ ，利用标签  $y$  指导样本的生成。

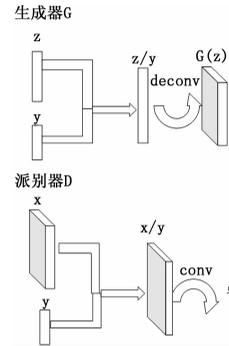


图 2 CDCGAN 网络结构示意图

### 1.4 带有梯度惩罚的 Wasserstein 生成对抗网络

#### 1.4.1 WGAN 网络

Wasserstein 距离又叫 Earth-Mover (EM) 距离，定义如下：

$$W(P_r, P_g) = \inf_{\gamma \in \Pi(P_r, P_g)} E_{(x, y) \sim \gamma} [\|x - y\|] \quad (3)$$

WGAN 网络中构造了一个包含参数  $\omega$  且最后一层不是非线性激活层的判别器网络 D，在限制  $\omega$  不超过某个范围的前提下，使得：

$$L = E_{x \sim P_r} [D(x)] - E_{x \sim P_g} [D(x)] \quad (4)$$

尽可能取到最大，此时  $L$  就近似等于真实分布与生成分布之间的 Wasserstein 距离 (忽略常数倍数  $K$ )。

WGAN 与 GAN 相比，其实只改了 4 点，如下<sup>[9]</sup>：

- 1) 判别器最后一层去掉 sigmoid。
- 2) 生成器和判别器的 loss 不取 log。
- 3) 每次更新判别器的参数之后把它们的绝对值截断到不超过一个固定常数  $c$ 。

4) 不要用基于动量的优化算法 (包括 Momentum 和 Adam)。推荐 RMSProp 和 SGD。

#### 1.4.2 WGAN-GP 网络

WGAN-GP 网络改进了 WGAN 网络，使用一个当且仅当其梯度小于或等于 1 的范数时的可微函数 1-Lipschitz 对约束函数  $L$  进行梯度惩罚。算法如下：

$$L = E_{\bar{x} \sim P_r} [D(\bar{x})] - E_{x \sim P_g} [D(x)] + \lambda E_{\bar{x} \sim P_r} [\|c \nabla_{\bar{x}} D(\bar{x})\|_2 - 1]^2 \quad (5)$$

类似于 WGAN 的目标函数, WGAN-GP 中生成器和判别器的 loss 不取 log,  $\hat{x} - P_g$  代表来自真实数据的样本,  $x - P_g$  代表来自噪声  $z$  通过生成器产生的样本,  $L$  近似等于真实分布与生成分布之间的 Wasserstein 距离。

梯度惩罚函数就是使用 Lipschitz 函数限制判别器的梯度不超过  $K$ , 即先求出判别器的梯度, 然后建立梯度与  $K$  之间的二范数。对于梯度惩罚项来说, 本文在每一个批次的真实图片数据和生成器生成的假图数据中做一个插值  $\hat{x}$ :

$$\hat{x} = \tilde{x} * \theta + x * (1 - \theta) \quad (6)$$

这里  $\theta$  代表  $U[0, 1]$  的随机数。相比于 WGAN, WGAN-GP 有以下几个特点:

- 1) 该模型提出了在原来 WGAN 模型的基础上利用梯度惩罚项来约束 Wasserstein 距离, 避免 WGAN 模型有可能产生的训练不稳定和生成数据病态的问题。
- 2) 该模型收敛速度更快, 生成样本质量也更高
- 3) 该模型适应性更强, 在不同架构下都可以得到稳定的训练结果。

## 2 本文网络模型

本文将 CDCGAN 和 WGAN-GP 结合起来, 从而得到本文的条件梯度 Wasserstein 生成对抗网络 (CDCWGAN-GP) 模型。为了提高判别器的提取图像特征能力, 本文将判别器 D 分为全局判别器和局部判别器, 把两个判别器连接起来一起打分。然后, 我们将训练好的网络中的判别器 D 提取出来, 构建一个新的卷积网络进行微调, 从而完成图像分类任务。

### 2.1 CDCGAN-GP 的惩罚函数

在 CDCWGAN-GP 中, 判别器 D 希望真实数据分布与生成数据分布之间的距离  $L$  越大越好, 相反, 生成器 G 希望  $L$  越小越好, 优化公式为:

$$L = E_{x \sim P_x} [D(x)] - E_{x \sim P_g} [D(x/y)] + \lambda E_{x \sim P_x} [(|| \nabla_x D(x) ||_2 - 1)^2] \quad (7)$$

$$\hat{x} = \tilde{x} * \theta + x * (1 - \theta) \quad (8)$$

$y$  代表图像的标签, 在本文中就是图片的类别。不同 WGAN-GP 的目标函数, CDCGAN-GP 的生成器输入数据融合了图片标签信息, 而标签信息可以指导生成器的样本生成。判别器 D 和生成器 G 的损失函数下:

$$loss(D) = E_{x \sim P_x} [D(x/y)] - E_{x \sim P_g} [D(x)] + \lambda E_{x \sim P_x} [(|| \nabla_x D(x) ||_2 - 1)^2] \quad (9)$$

判别器的损失函数由两部分构成: 1. 生成样本通过判别器得到的损失; 2. 真实样本通过判别器得到的损失, 并在前添加负号; 两部分加起来就是判别器的损失函数。

$$loss(G) = -E_{x \sim P_g} [D(x/y)] \quad (10)$$

生成器的损失函数定义为生成样本通过判别器得到的损失, 再添加负号。

### 2.2 生成器的网络结构

我们以训练集 CIFAR10 为例, 图像大小为  $32 * 32$ 。生成器 G 的输入  $z$  为服从正态分布的 100 维的随机噪声向量,

将  $z$  与类别标签  $y$  (10 维) 连接形成一个大小为  $[64, 110]$  的张量, 通过一个 linear 全连接层将  $z$  变换为维度是 2 048 的向量, 然后将其 reshape 为维度大小是  $[512, 2, 2]$  的张量, 进行非线性 relu 函数变换, 得到第一层网络的输出。然后通过四层卷积核大小为  $5 * 5$ , 步长为  $(2, 2)$  的转置卷积层<sup>[10]</sup>, 输出都用非线性 relu 函数激活, 得到大小为  $[32, 32, 3]$  的张量, 最后用 tanh 函数激活, 得到生成器的输出图片。

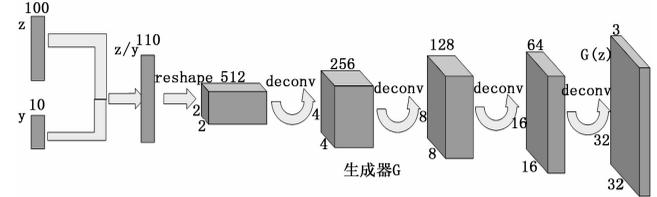


图 3 CDCWGAN-GP 生成器结构示意图

表 1 生成器网络层

类型	卷积核大小	步长	卷积核个数
concat	/	/	110
linear	/	/	2048
reshape	/	/	512
deconv2d(1)	$5 * 5$	$2 * 2$	256
deconv2d(2)	$5 * 5$	$2 * 2$	128
deconv2d(3)	$5 * 5$	$2 * 2$	64
deconv2d(4)	$5 * 5$	$2 * 2$	3

### 2.3 判别器网络结构

无论是 GAN 模型还是 DCGAN 模型中, 判别器的网络模型都是不断缩小特征图从而提取到一整张图片的全局特征, 全局特征可以代表一张图片的整体结构, 但是对于图片中的局部特征并不能很好地表达出来。为了加强模型中判别器提取图像特征的能力, 本文再设计一个局部判别器, 用于提取图片中局部特征, 最后将全局判别器和局部判别器结合起来一起打分。

以 CIFAR10 数据集为例。全局判别器的输入是一张维度为  $[3, 32, 32]$  大小的图像, 经过五层卷积核大小为  $5 * 5$ , 步长为  $2 * 2$  的卷积层, 每层输出都要进行 leakyrelu 函数变换, 得到大小为  $(512, 1, 1)$  的张量, 最后 reshape 成一个维度大小是  $[1, 512]$  的特征向量。

局部判别器的网络结构类似于全局判别器, 只不过局部判别器的输入是以原图中心点为中心, 大小是原图  $1/4$  的图像, 尺寸为  $16 * 16$ 。经过四层卷积核大小为  $5 * 5$ , 步长为  $2 * 2$  的卷积层, 每层输出都要进行 leakyrelu 函数变换, 最后 reshape 成一个大小为  $(1, 512)$  的特征向量。

最后我们将全局判别器和局部判别器的输出用 concat 函数直接连接起来, 形成一个维度大小为 1 024 的特征向量, 然后经过 linear 全连接层, 输出为 1 或者 0 代表输入图像的真或者假。

这里我们不同于 CDCGAN 的网络，由于我们之后还要将判别器提取出来进行图像分类，所以判别器 D 每层的输入没有加入标签  $y$ 。并且由于我们使用了梯度惩罚，所以最后的输出去掉 sigmoid 函数。

### 2.4 训练

本文对数据集进行批量化处理，batch 的大小设为 64。

表 2 全局判别器的网络层

类型	卷积核大小	步长	卷积核个数
conv2d	5 * 5	2 * 2	64
conv2d	5 * 5	2 * 2	128
conv2d	5 * 5	2 * 2	256
conv2d	5 * 5	2 * 2	512
conv2d	5 * 5	2 * 2	512
reshape	—	—	512

表 3 局部判别器的网络层

类型	卷积核大小	步长	卷积核个数
conv2d	5 * 5	2 * 2	64
conv2d	5 * 5	2 * 2	128
conv2d	5 * 5	2 * 2	256
conv2d	5 * 5	2 * 2	512
reshape	—	—	512

表 4 concat 连接层

类型	卷积核大小	步长	输出
concat	—	—	1024
linear	—	—	1

由于是对每个 batch 中的每一个样本都做了梯度惩罚，因此判别器中不能用 batchnormalization<sup>[11]</sup>，但是可以使用其他的 Normalization 方法<sup>[12]</sup>，比如 Layer Normalization、InstanceNormalization。在本文的 D 网络，再训练 G 网络，并且每训练 5 次 D 网络，再训练 1 次 G 网络，目的就是为了保证 D 网络大致满足 Wasserstein 距离的条件。在 WGAN 网络模型中训练网络并不建议用 Adam 优化器，但是在本文的网络模型中因为使用了 WGAN-GP 的训练方式，所以还是可以用 Adam 优化器，且参数设定为： $\beta_1 = 0.0005$ ， $\beta_2 = 0.9$ ，并且本文采用了适应性更强的根据训练次数可变的学習率，学习率初始化为 0.002，当训练的迭代次数 epoch 大于设定的训练次数的一半时，学习率变为原来的十分之一，当训练的迭代次数 epoch 大于设定的训练次数的五分之四时，学习率再变为上一次的十分之一。可变化的学习率可以让参数调整变得更加有效。

### 2.5 进行图像识别

当训练好 CDCWGAN-GP 网络模型后，把判别器 D 提取出来，将最后一层 linear 全连接层的输出改为  $n$ （将要分类的类别），最后加上 Sigmoid 分类器，构建一个新的卷积网络用于图像识别。卷积网络的损失函数是 Softmax<sup>[14]</sup> 的

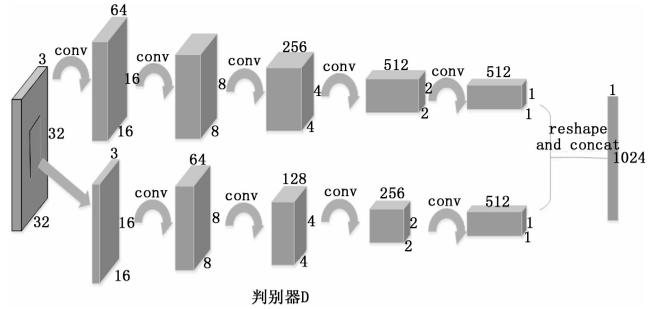


图 4 CDCWGAN-GP 判别器结构示意图

结果与标签  $y$  的交叉熵，使用 Adam 优化器进行优化 ( $\beta_1 = 0.0005$ ， $\beta_2 = 0.9$ )，并且同样对数据集进行批量化处理，batch 的大小设为 64。新的用于图像识别的网络只需要进行参数微调，就可以得到很好地分类效果了。在生成式对抗模型中，由于生成器可以生成高质量的样本，所以判别器不仅能通过原始数据集学习图像特征，而且还可以通过生成样本学习到更多的图像特征。并且本文设计的具有全局和局部一致性的判别器，能更有效的提取到图像的内部隐含特征，从而有效提高了图像识别的准确率。

## 3 实验结果和分析

本文分别在 FASHION-MNIST<sup>[15]</sup> 和 CIFAR-10 数据集上进行实例验证。实验环境为：Intel (R) Core (TM) i7-7700K CPU @4.20 GHz 四核处理器，金士顿 DDR4 2400 MHz 16 GB 运行内存 (RAM)，Nvidia GeForce GTX 1080 (8 GB/戴尔) GPU，Tensorflow (1.4.0) 平台。

### 3.1 FASHION-MNIST 的实验结果

许多网络模型在 MNIST 数据集上的分类测试结果都已经接近 100%，并且 MNIST 数据集的识别并不能代表真正的计算机视觉问题。所以，我们采用 FASHION-MNIST 数据集代替 MNIST 数据集。类似 MNIST 数据集，FASHION-MNIST 数据集包含了 7 万张商品照片，其中分为 10 个类别，训练集有 6 万张，测试集有 1 万张，每张图片都是灰度图片，且像素大小为  $28 * 28$ 。

首先我们训练对抗网络，训练初期，判别器的损失迅速增加，而生成器损失缓慢减少。随着训练对抗的不断进行，判别器损失的增幅减慢，而生成器损失的降幅加快，最终生成器和判别器的对抗达到一个比较稳定的范围，模型达到收敛。

在整个训练过程中，生成器 G 和判别器 D 一直在进行对抗，训练开始后，判别器强于生成器，意味着生成样本与判别样本的 Wasserstein 距离较大。随着训练次数的不断增加，对抗趋势最终将会稳定在一个范围内，此时意味着生成样本与判别样本的 Wasserstein 距离较小，生成器会略强于判别器。通过对抗网络得到的生成样本如图 5 所示。

图 5 表明，当 epoch 达到 50 的时候，生成器就可以生成质量不错的图片，随着 epoch 不断增加，生成样板的质量

也越来越高。

表 5 FASHION-MNIST 各种方法的分类准确率对比

识别方法	预处理	准确率/%
Linear classifier	/	87.6
Cnn(3 Conv with Maxpooling)	Normalization, randomtranslation	89.1
SVM,Gaussian Kernel	/	91.6
LeNet-5	huge distortions	92.1
DCGAN+softmax	/	93.4
Ours	/	94.5

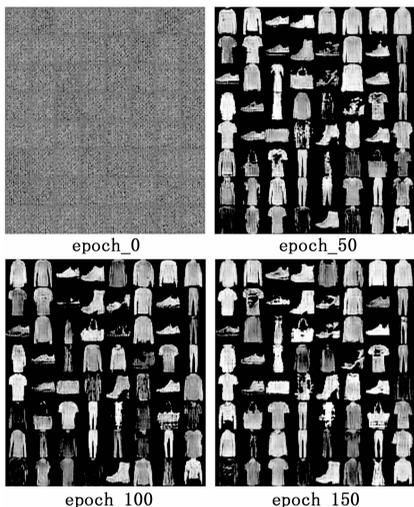


图 5 FASHION-MNIST 生成样本

当对抗网络训练好后, 我们将判别器抽取出来再进行图像分类实验。实验结果如图 6 所示。

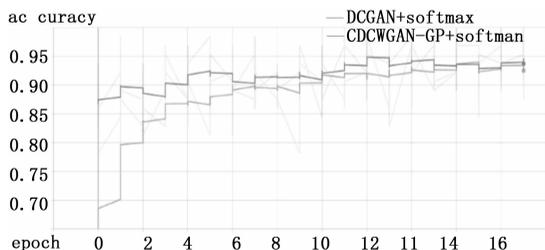


图 6 FASHION-MNIST 分类准确率的变化

通过与 DCGAN 进行的实验对比, 突显了本文模型的分​​类效果, 实验结果如图 6。其中蓝色曲线为 DCGAN+softmax 的实验结果, 橙色曲线为 CDCWGAN-GP+softmax 的实验结果。可以发现, 将判别器提取出来后, 再通过较少的训练次数进行微调就可以得到很好地分类效果。值得注意的是, 本文提出的 CDCWGAN-GP 相比 DCGAN, 收敛的速度更快, 最终准确率也更高, 可以达到 94.5%。

表 5 将本文模型与传统卷积神经网络及其各类变种神经网络模型<sup>[17]</sup>进行了实验对比, 在与本文网络结构相似的非大型模型中, 本文的分类准确率最高, 更加突出显示了本文模型分类能力的优越性。值得特别说明的是, 通过在 CDC-

WGAN-GP 训练对抗网络的过程中间接进行的数据增强效果优于对原始数据进行预处理的数据增强效果。

### 3.2 CIFAR-10 的实验结果

CIFAR-10 数据集中一共有 6 万张彩色图片, 包含了来自 10 个类别的自然图像, 每张图片大小是 32 \* 32。其中, 训练集有 5 万张图片, 5 等分地分为 5 个训练批, 测试集有 1 万张图片。训练中的损失如下: 在 CIFAR10 上的训练过程中, 整个网络的判别器损失  $d\_loss$  呈增加趋势, 且增幅不断减慢, 而生成器损失  $g\_$

表 6 CIFAR10 各种方法的分类准确率对比

识别方法	预处理	准确率
1layer k-means	/	80.6%
3 layer K-means learned BF	/	82.1%
View invariant	/	81.9%
Exemplar CNN	/	84.3%
DCGAN+L2SVM	/	82.8%
Ours	/	88.1%

loss 呈增加趋势, 且增幅不断加快, 最终判别器和生成器的对抗趋于稳定的状态, 模型达到收敛。

CIFAR-10 数据集的生成样本如图 7 所示, 可以发现, 随着 epoch 增加, 生成样本的效果越来越好, 当 epoch 达到 150 后, 生成样本的效果达到最优。

最后, 通过对判别器参数进行微调, 得到分类的实验结果如图 8。一开始分类的准确率就可以达到 80% 左右, 当 epoch 达到 20 的时候, 准确率趋于稳定, 可以达到 88% 左右。图 8 表明本文模型的分​​类损失随着 epoch 增加而逐渐降低。

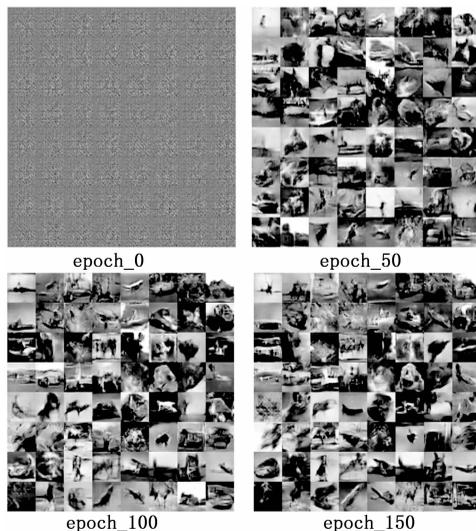


图 7 CIFAR10 生成样本

在 CIFAR-10 数据集上, 同样将本文模型与 DCGAN 模型进行了实验对比, 如图 8 通过实验结果可以证明, 相比 DCGAN 模型, 我们的模型收敛速度更快, 并且分类的

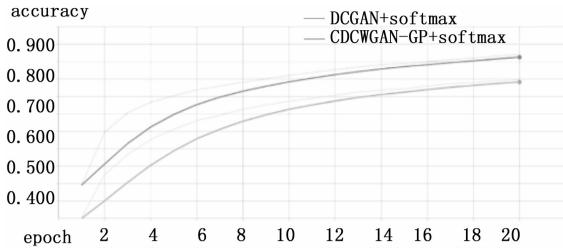


图 8 CIFAR10 的分类准确率的变化

结果准确度也更高。

表 6 是本文模型与其他模型<sup>[18-19]</sup>在 CIFAR-10 数据集上进行的实验结果比对。可以发现, 本文模型优于其他的比较模型且取得了最佳的效果。

### 4 结论

本文结合 WGAN-GP 网络模型和 CDCGAN 网络模型, 提出了混合模型—条件梯度 Wasserstein 生成对抗网络 CDCWGAN-GP。并且为了能更全面的提取图像的特征, 将判别器分为了全局判别器和局部判别器, 利用两个判别器一起打分, 从而更好地识别图像的全局和局部一致性。利用本文的模型来进行图像识别, 可以使准确率显著提升。本文提出的模型不仅能得到质量很高的生成样本, 而且相比传统的 GAN 训练方式, 收敛速度更快, 训练时也更加稳定, 再利用具有全局和局部一致性的判别器能更好地提取图像特征并进行特征表达。在 FASHION-MNIST 和 CIFAR10 数据集上进行的实验证明, 相比于其他类似的方法, 本文的模型分类效果更加出色, 证明了此模型的可行性。在以后的工作中, 作者将针对高分辨率图像该如何训练, 对抗模型什么时候达到最优效果, 以及不同参数(例如不同优化器的选择, batch 的大小不同)对精度的影响等问题进行更加深入的研究。

### 参考文献:

[1] Goodfellow I J, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets [A]. International Conference on Neural Information Processing Systems [C]. MIT Press, 2014: 2672-2680.

[2] Li J W, et al. Adversarial learning for neural dialogue generation [J]. [S. l.], 2017.

[11] 刘漳辉, 王晓莉. 云计算虚拟机群中带遗传算法的负载均衡算法 [J]. 福州大学学报, 2012, 40 (4): 453-458.

[12] Yaser J, Lo'ai T, Fadi A. Scalable Cloudlet-based Mobile Computing Model [J]. Procedia Computer Science, 2014, 34 (1): 434-441.

[13] Karadimce A, Danco D. Adaptive multimedia learning delivered in mobile cloud computing environment [A]. The 4th International Conference on Cloud Computing, Grid, and Virtualization [C]. 2013: 1012-1020.

[14] 陈伟宏, 吴宏斌, 肖卫初. 基于 WiFi/WiMAX 的车载移动网

[3] Yeh R A, Chen C, Lim T Y, et al. Semantic Image Inpainting with Deep Generative Models [J]. [S. l.], 2016.

[4] Mirza M, Osindero S. Conditional Generative Adversarial Nets [J]. Computer Science, 2014: 2672-2680.

[5] Radford, Alec, Metz, Luke, et al. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks [J]. Computer Science, 2015.

[6] Arjovsky M, Chintala S, Bottou L. Wasserstein GAN [J]. [S. l.], 2017.

[7] Salimans T, Goodfellow I, Zaremba W, et al. Improved Techniques for Training GANs [J]. [S. l.], 2016.

[8] 王坤峰, 苟超, 段艳杰, 等. 生成式对抗网络 GAN 的研究进展与展望 [J]. 自动化学报, 2017, 43 (3): 321-332.

[9] 滕少华, 孔棣睿, 基于生成式对抗网络的中文字体风格迁移 [J/OL]. 计算机应用研究, 2019, 36 (11).

[10] 刘海龙, 李宝安, 吕学强, 等. 基于深度卷积神经网络的图像检索算法研究 [J/OL]. 计算机应用研究, 2017, 34 (12).

[11] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift [J]. [S. l.], 2015: 448-456.

[12] Gulrajani I, Ahmed F, Arjovsky M, et al. Improved Training of Wasserstein GANs [J]. [S. l.], 2017.

[13] Ba J L, Kiros J R, Hinton G E. Layer Normalization [J]. [S. l.], 2016.

[14] 焦计晗, 张帆, 张良. 基于改进 AlexNet 模型的油菜种植面积遥感估测 [J]. 计算机测量与控制, 2018, 26 (2): 186-189.

[15] Peng K C, Wu Z, Ernst J. Zero-shot deep domain adaptation [J]. [S. l.], 2017.

[16] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [A]. International Conference on Neural Information Processing Systems [C]. Curran Associates Inc. 2012: 1097-1105.

[17] 邹冲, 蔡敦波, 赵娜, 等. 基于 SVM-LeNet 模型融合的行人检测算法 [J]. 计算机工程, 2017, 43 (5): 169-173.

[18] Li H, Liu H, Ji X, et al. CIFAR10-DVS: An Event-Stream Dataset for Object Classification [J]. Frontiers in Neuroscience, 2017, 11.

[15] Yaser J, Zakarea A, Moath J. Teachcloud: a cloud computing educational tool kit [J]. International Journal of Cloud Computing, 2013, 2 (2/3): 237-257.

[16] 王英强, 陈绥阳, 张首军. 保障云服务性能的组件服务副本增删方法研究 [J]. 控制工程, 2016, 23 (7): 1090-1096.

[17] 李晓凯, 吕宏伟. 一种路点受限的移动传感器网络移动模型 [J]. 传感器与微系统, 2014, 33 (8): 62-64.

[18] 张兵. 一种用于云计算数据库的数据挖掘方法研究 [J]. 控制工程, 2016, 23 (6): 956-961.

(上接第 156 页)