

基于深度卷积生成对抗网络的图像生成

尹玉婷, 肖秦琨

(西安工业大学 电子信息工程学院, 陕西 西安 710021)

摘要: 图像生成是虚拟现实技术(virtual reality, VR)中的重要技术手段, 针对传统图片生成方法需要大量的数据集进行训练, 且生成的图片轮廓不清晰等问题, 采用基于深度卷积神经网络和生成对抗网络来实现图片的生成。为了保证生成图片的真实性和完整性, 在图片生成阶段引入变分自编码器, 通过编码器获取到输入图片数据的均值和方差, 将图片对应的隐藏变量转化为标准的高斯分布, 然后通过生成器生成新的图片; 在识别阶段, 采用深度卷积神经网络训练判别器, 将生成的新的图片输入到已经训练好的判别器中, 运用梯度下降法计算损失函数, 不断优化整体系统模型。通过对 MNIST 图像数据集的训练, 实验表明该方法能生成质量较高的图片, 它生成的图像无法用肉眼与真实数据区分开, 并且在不同网络条件下都有较高的识别率。该方法提高了 MNIST 生成模型的技术水平。

关键词: 生成对抗网络; 深度卷积网络; 变分自编码器; 图像生成; 梯度下降法

中图分类号: TP391.4

文献标识码: A

文章编号: 1673-629X(2021)04-0086-07

doi:10.3969/j.issn.1673-629X.2021.04.015

Image Generation Based on Deep Convolution Generative Adversarial Networks

YIN Yu-ting, XIAO Qin-kun

(School of Electronic Information Engineering, Xi'an Technological University, Xi'an 710021, China)

Abstract: Image generation is an important part of virtual reality (VR). In order to solve the problem that traditional image generation method needs a large number of data sets for training and the generated image contour is not clear, the deep convolutional neural network and generation adversarial network are used to realize the image generation. To ensure the authenticity and integrity of the generated image, a variational auto-encoder is introduced in the image generation stage. The mean value and variance of the input image data are obtained by the encoder, and the hidden variables corresponding to the image are transformed into the standard Gaussian distribution, then the new image is generated by the generator. In the recognition stage, the deep convolutional neural network is used to train the discriminator, and the generated new images are input into the trained discriminator. The gradient descent method is used to calculate the loss function and continuously optimize the overall system model. Through the training of MNIST image data set, the experiment shows that the proposed method can generate high-quality images which cannot be distinguished from the real data with the naked eye, with high recognition rate under different network conditions. It improves on the state of the art for generative models on MNIST.

Key words: generative adversarial network; deep convolutional network; variational auto-encoder; image generation; gradient descent method

0 引言

近年来传感器技术、互联网物联网通信技术以及计算机视觉技术迅速发展, 人们越来越希望通过监控系统和计算机视觉完成视觉事件的记录以及对感兴趣事件做出判断的任务, 从而实现人机交互等。各种新型虚拟增强现实应用技术在军事科学与工程、教育以及医疗、娱乐等各个方面得到了广泛应用。VR 的实际

应用以当前计算机网络视觉技术领域视频中视频图像图片生成后的相关视觉技术应用为理论基础, 将图像生成后的影像图片、视频与其他视觉技术进行结合, 构造显示出—个真实虚拟的“现实”并将世界展示出来。为了生成多样化、清晰化的图片, 展开了对计算机模拟图像生成方法的研究。

生成模型是通过应用最小—最大优化框架来训练

收稿日期: 2020-05-31

修回日期: 2020-09-31

基金项目: 国家自然科学基金资助项目(61671362, 6207010855)

作者简介: 尹玉婷(1994-), 女, 硕士研究生, 研究方向为人工智能、图像处理; 肖秦琨, 博士, 教授, 研究方向为人工智能与图像处理、手势识别与跟踪。

和生成数据。近年来提出的图像生成方法主要有两种,一种是 Goodfellow 等人提出的生成对抗网络^[1] (generative adversarial networks, GAN), 另一种是自编码器^[2] (auto encoder, AE)。GAN 是由生成网络和判别网络组成,这两种网络之间是一种对抗博弈关系,生成网络就像造“假币”的机器,根据已有的“纸币”(真实数据)模仿生产出近乎真实的纸币,判别网络作为一个分类器起到了类似“验钞机”的作用,判断纸币的真伪,也就是决策出生成数据是否真实。依据判别网络给出的结果,生成网络不断改进网络参数设置,提高自身的生成能力,使生成的结果更接近原始数据。从博弈论的角度来说,两种网络处于一种竞争关系,最终两个网络实现纳什均衡。原始的生成对抗网络在训练过程中生成的图像质量不太稳定,生成效果不理想。自动编码器主要使用编码—解码框架生成数据,学习编码图像和潜在对象代码之间的编码器和解码器映射,其生成的图像更加准确;但由于其使用均方误差等作为损失函数,生成的图片没有细节,画面不清晰,单纯为了判别器无法判断出图片是否是真实样本而过度发挥,造成图片失真。基于这两种方法各自的优缺点,该在 GAN 中引入深度卷积神经网络以及编码原理,提高网络提取特征的能力,该方法避免了生成图片质量不高、图像失真的缺点,使得生成的图像拥有更高的质量。

1 生成模型

机器学习的一个主要部分是生成模型与判别模型。在判别建模中,一个目标是根据给定的观测值学习预测变量,另一个目标是解决学习所有变量的联合分布这一更普遍的问题。生成模型可模拟在现实世界中如何生成数据。几乎每门科学都将“建模”理解为通过假设理论并通过观察测试这些理论来揭示这一生成过程。生成模型具有吸引力的原因有很多,将不知道或不关心的细节(即有害变量)视为噪声。

设图片的真实样本数据为 $\{x_1, x_2, \dots, x_n\}$, 其整体用 X 表示,在理想情况下,根据真实样本数据得到 X 的分布 $p(X)$, 对 $p(X)$ 进行采样就可以得到所需的 X , 即 $p(Z|X)$ 根据贝叶斯公式可得:

$$p(Z|X) = \frac{P(X|Z)P(Z)}{P(X)} \quad (1)$$

实际上, X 的分布 $p(X)$ 无法直接求得,此时引入隐变量 Z 希望在给定 X 的条件下可以推断出 Z 的分布,由全概率公式得:

$$p(X) = \int p(X|Z)p(Z) dZ \quad (2)$$

其中, $p(X|Z)$ 表示一个由 Z 生成 X 的模型,假设 Z 服

从标准正态分布,即 $p(Z) \sim N(0,1)$ 。首先从标准正态分布中采样一个隐变量 Z , 然后根据隐变量 Z 来不断拟合 X , 达到生成 X 的目的。

当 Z 为高维度变量时,那么

$$p(X) = \iiint \dots p(X|Z)p(Z) dZ \quad (3)$$

实际中大多数 Z 的 $p(X|Z)$ 都是 0, 对评估 $p(X)$ 几无用处。这需要一个对任意 X 能给出产生 X 的 Z 值分布的函数 $q(Z|X)$ 。 q 下产生 X 的 Z 值空间很有希望远小于先验 $p(Z)$ 的 Z 值空间,因而计算 $E_{Z \sim q} p(X|Z)$ 就更为简便。但现在 Z 来自任意分布的 $q(Z)$ 而非 $N(0,1)$, 要用于优化 $p(X)$ 需先将 $E_{Z \sim q} p(X|Z)$ 和 $p(X)$ 关联起来。 $E_{Z \sim q} p(X|Z)$ 和 $p(X)$ 之间的关系是变分贝叶斯方法的基础之一。定义 $p(Z|X)$ 和 $q(Z)$ 之间的 KL 散度, 希望其之间差异最小, 即求解 $\min \text{KL}[q(Z) \| p(Z|X)]$ 。

$$\begin{aligned} \text{KL}[q(Z) \| p(Z|X)] &= \int q(Z) \log \frac{q(Z)}{p(Z|X)} = \\ &= \log p(X) + \int q(Z) \log q(Z) dZ - \\ &= \int q(Z) \log [p(X|Z)p(Z)] dZ \end{aligned} \quad (4)$$

由于 X 是给定数据, 则 $p(X)$ 为定值, 那么最小化 $\min \text{KL}[q(Z) \| p(Z|X)]$ 即最小化下式:

$$\begin{aligned} \min L &= \int q(Z) \log q(Z) dZ - \\ &= \int q(Z) \log [p(X|Z)p(Z)] dZ = \\ &= \text{KL}(q(Z) \| p(Z)) - \\ &= \int q(Z) \log [p(X|Z)] dZ \end{aligned} \quad (5)$$

2 变分自编码器

变分自编码器 (variational auto-encoder, VAE) 是无监督深度学习中最重要生成模型之一^[3]。

由于自编码器输出的隐空间信息未知且不可测量, 无法对其进行数学建模, 由此衍生出一种新的生成模型, 即变分自编码器。在了解 VAE 之前, 首先了解自编码器的原理。自动编码器本质上是学习输入数据低维特征表示的神经网络结构, 它由两个相连接的网络组成, 分别是编码器 (encoder) 和解码器 (decoder)。为了提高编解码的效率, 编码器和解码器是由人工神经网络构成。自编码器 AE 用于图片生成, 将高维数据缩减为低维特征空间, 然后从低维表示中还原原始数据。如图 1 所示, 编码器 Z 获取输入数据 X_i 并编码产生特征映射集合 X'_i 。通常情况下, 为了减少输入数据冗余度, 特征映射 X'_i 大小小于输入数据 X_i 。编码器完成了原始数据从高到低的转换, 接下来, 解码器将原

始数据从低到高进行还原重构。为了提取输入数据的特征,使用解码器,它将存在解码器 f 中的编码数据集 X'_i 重构出原始数据 X_i 。

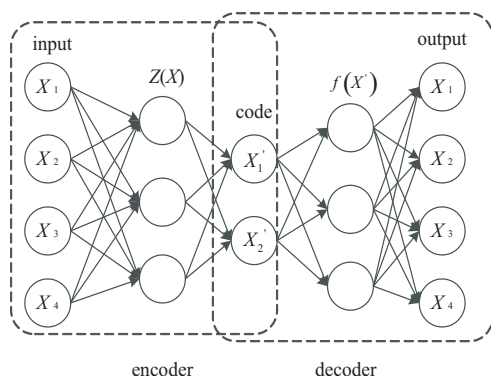


图1 自编码器原理

变分自编码器也就是在自编码器的基础上,构建

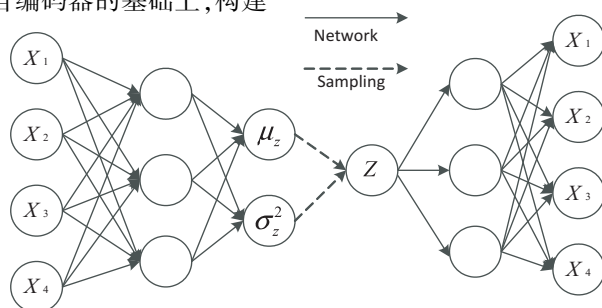


图2 变分自编码器原理

VAE 的目标是最大限度上降低数据对数似然的损失下限,根据式(4),可以推断出通过不断减小损失函数 $KL(q(Z) \parallel p(Z)) - \int q(Z) \log[p(X|Z)] dZ$ 使其无限逼近0,这意味着文中生成模型将变得更好,那么此时的分布 q 就可以近似代替原始数据 p 分布,而 q 是已知的可以求解的。可以将VAE视为两个耦合但独立参数化的模型:编码器或识别模型,以及解码器或生成模型。这两个模型相互支持。识别模型向生成模型提供其潜在随机变量后验的近似值^[5-6],该模型需要在“期望最大化”学习的迭代内更新其参数。相反,生成模型是用于识别模型的各种支架,以学习有意义

编码网络过程中对网络结构加以限制。根据贝叶斯理论,编码器把输入数据映射为在一个隐变量空间的后验分布,但是根据式(1),输入数据到隐变量的后验分布并不好求出其解。编码器就相当于一个变分推断网络^[4],VAE给编码器加了一个条件约束,让编码网络不再是每次产生一个隐含向量而是生成两个向量,其分布能够大致上遵循一个标准正态分布,即就是编码器一个用来计算均值,另一个计算标准差,隐向量从正态分布采样得出。解码器实现从隐向量到输入数据的恢复,将低维数据还原为高维,此过程可以用似然函数 $p(X|Z)$ 表示。由于解码器的输入为隐向量 Z ,其中丢失了原始数据的部分信息,因此对于图片生成重构出的原始图片会存在失真、模糊等问题。变分自编码器原理如图2所示。

的数据表示,包括可能的类别标签,识别模型是根据贝叶斯规则的生成模型的近似逆过程。

3 深度卷积生成对抗网络图像算法

该文采用深度卷积生成对抗网络来进行手写字体图像生成,图像生成模型总体框架由两部分组成,如图3所示,左边是判别网络,又称判别器(Discriminator),根据输入图片的真假输出结果“1”或者“0”,当输入的图片为真实图片时,输出为“1”,否则判别器输出为“0”;右边是生成网络,又称生成器(Generator),它可以看作是一个变分自编码器,经过编码获取输入图片的隐空间信息,经由解码器生成新的图片。

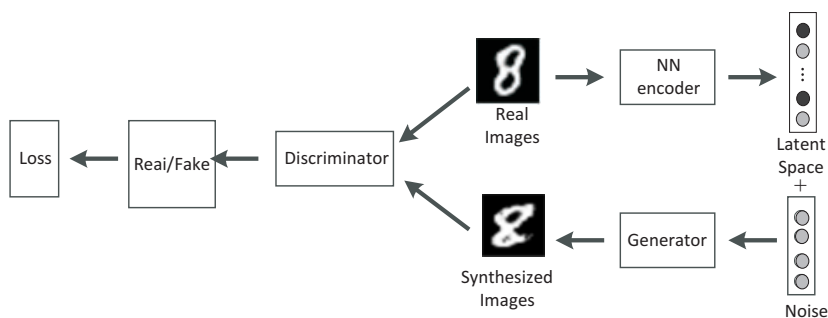


图3 图像生成模型总体框架

编码器网络通过将训练数据集中的真实图像进行编码,将真实图像映射出来到一个潜在空间,提取真实

训练图像的物理特征^[7]。生成器网络将得到的真实图像编码作为输入,以期生成能够让判别器不能正确判

别的图片。判别器网络对真实图片和生成图片进行判别,判定其属于真实图片还是生成器生成图片,为了避免前期网络判别效果与后期网络判别效果差距太大的缺点,网络加入了随机性技术,提高了判别器网络的稳定性与鲁棒性,得到了更好的判别效果,同时促使生成器生成更加真实的图像。

3.1 生成模型的优化

生成器网络将得到的真实图像编码作为输入,以

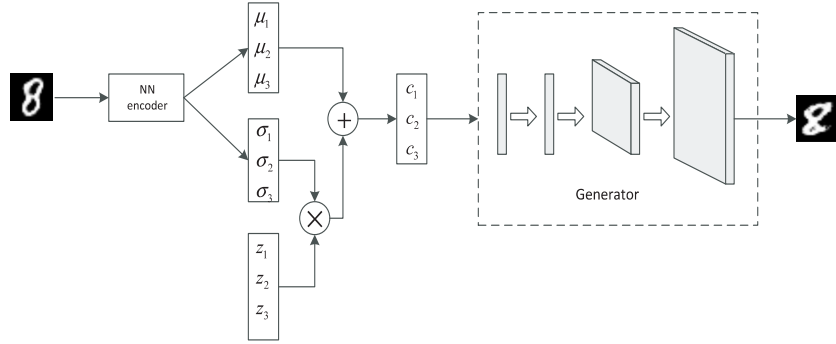


图4 生成器优化模型

通常情况下,一般的CNN网络包含输入层、隐藏层和输出层,中间的隐藏层是由卷积层、池化层和全连接层组成的^[8-9]。输入层负责接收数据并将数据传递给隐藏层,不进行计算。隐藏层的神经元不与外部数据直接接触,它将接收到的数据经过激活函数计算后输出传递给下一层隐藏神经元或者输出神经元,隐藏层的层数和每层含有的神经元数目属于模型的超参数,通常在训练前使用验证集寻找最佳的一组超参数集。输出层是输出网络的计算结果,在训练过程中,通常和训练集中该次输入对应的理想输出进行比较,它们的差异就是网络优化的指标。但是在生成网络的卷积结构的改进上取消了池化层,而是将卷积层替代了池化层。

设图像的真实样本为 x ,经过编码器编码后得到均值 μ 和方差 σ ,使用标准正态分布构造随机特征向量 c 。

$$c = \mu + \sigma^2 \odot N(0, I) \quad (6)$$

生成器得到的样本为 \tilde{x} ,判别器输出结果为 $y \in [0, 1]$ 。特征向量经过生成器,其过程可描述为:

$$G: \tilde{x} = g(z, \theta_g) \quad (7)$$

其中, $g(\cdot)$ 为生成函数, θ_g 为生成器参数,经过判别器输入的图片数据,原始图像和生成图像分别为 \tilde{X} 、 X ,其过程如式(8)表示:

$$\begin{cases} \tilde{X} = D_f(\tilde{x}, \theta_f) \\ X = D_f(x, \theta_f) \end{cases} \quad (8)$$

根据式(8),可得经过判别器后的输出:

$$y = \begin{pmatrix} D(x) \\ D(\tilde{x}) \end{pmatrix} = \begin{pmatrix} D(x) \\ D(G(z)) \end{pmatrix} \quad (9)$$

期生成能够让判别器不能正确判别的图片。为了能够生成高质量的图片,在输入端进行改进,将真实图片进行编码操作,提取真实的图像特征,编码器也由神经网络构成。对于每个得到的均值和方差,进入具有正态分布的噪声,使其向满足正态分布的特征向量靠近,然后将得到的特征向量输入生成器生成图像。生成模型的结构如图4所示。

其中, $D(x)$ 的含义是判别样本 x 正确的概率,那么 $(1 - D(x))$ 则为判别成错误样本的概率。GAN网络整个输入数据要么是真实图片要么是生成图片,当网络达到平衡时,判别器 $D(x)$ 的概率会接近0.5,此时根据式(9)的定义,建立误差函数,其表达式为:

$$\min_{\theta_d} \left\{ - \left[\sum_{x \in p(x)} \log(D(x, \theta_d)) + \sum_{\tilde{x} \in p(\tilde{x})} \log(1 - D(\tilde{x}, \theta_d)) \right] \right\} \quad (10)$$

其中, $D(x)$ 表示图像真实样本分布, $D(\tilde{x})$ 是服从于投到生成器中噪声的分布,判别器内部参数 θ_d 如下:

$$\theta_d = (\theta_f, \theta_c) \quad (11)$$

生成模型通过反向传播算法^[10-11]可以调节生成网络的权重大小,使伪图像(就是生成的样本图片)与真实的样本图片不断接近,判别模型的损失函数会不断由大变小,生成模型的损失函数不断由小变大,随着时间的推移,最终使得网络平衡的具体方法为最大化噪声分布的熵值,最大化式(12):

$$\max_{\theta_g} \sum_{\tilde{x} \in p(\tilde{x})} \log(D(\tilde{x})) \quad (12)$$

当 $\log(D(\tilde{x}))$ 值越大,说明生成网络的拟合能力越好,生成的图片可以以假乱真,判别器识别该图片的概率为0.5,此时网络图片的生成效果最好。

3.2 判别模型的优化

判别网络的改进设计如图5所示。它相较于生成网络没有做出巨大改变,如果将这一过程比作编码,将是一个编解码操作,生成网络是编码器,而判别器就是解码器。生成网络实现反卷积,那么判别网络完成卷积操作,实现生成网络的逆过程。与生成网络相同的是,同样取消了池化层的操作,用卷积层来代替池

化层^[12]。

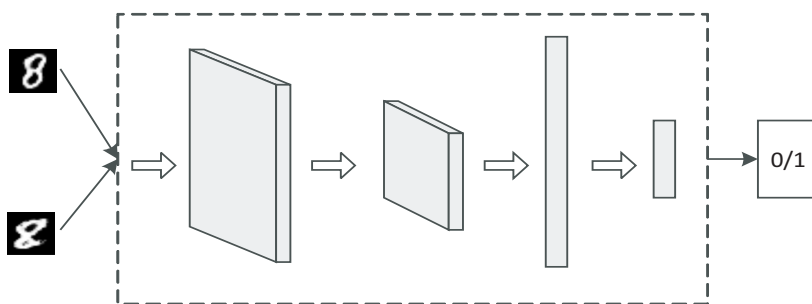


图 5 判别器优化模型

设 $p_{\text{data}}(x)$ 和 $p_g(x)$ 分别表示真实手写字体样本和优化后的生成模型输出的图片样本。判别网络和生成网络都有自己的损失函数,但是 GAN 由这两者组成,损失函数就依照下式进行修正:

$$V(D, G) = \int_x p_{\text{data}}(x) \log(D(x)) dx + p_g \log(1 - D(x)) dx \quad (13)$$

在训练过程中,需要 $p_{\text{data}}(x)$ 和 $p_g(x)$ 无限接近。在对式(13)进行优化的过程中,需要对生成网络和判别网络分别进行交互迭代^[13-14]。当训练判别网络时,将真实手写字体图片与生成的图片混合在一起,作为输入送入判别器,判别器会给出结果“0”或“1”,根据输出值调整网络参数;当训练生成网络时,将两个网络看成一个网络,随机输入一组数据,生成网络产生一张图片,判别器给出判别结果,将差异反馈给生成网络,调整参数,如此进行交互迭代,直到最终两个网络达到稳定。对判别网络进行优化^[15]的时候,也需要对其进行最小化交叉熵操作,其损失函数如下:

$$V(D, G) = -\frac{1}{2} \int_x p_{\text{data}}(x) \log(D(x)) dx - \frac{1}{2} p_g \log(1 - D(x)) dx = -\frac{1}{2} \int_x [p_{\text{data}}(x) \log(D(x)) dx + p_g \log(1 - D(x))] dx \quad (14)$$

式(15)是典型的统计分布问题,其最优解为:

$$D^*(x) = \frac{p_{\text{data}}(x)}{p_{\text{data}}(x) + p_g(x)} \quad (15)$$

由于 $p_{\text{data}}(x)$ 和 $p_g(x)$ 都表示概率大小,在实际训练时, $p_{\text{data}}(x)$ 和 $p_g(x)$ 取值不可能为 0,其取值一定为 (0,1] 之间的实数。根据上述理论,目标函数必须要在连续空间中处处可微,根据一阶求导规则,目标函数 $V(D, G)$ 在式(15)处取得最小值。当固定生成网络时,判别网络已经具有最优解,同时判别网络达到了最好的性能。同理,在判别器最优的条件下,将式(14)可以重写为下面的形式:

$$V(D, G) =$$

$$\begin{aligned} & E_{x \sim p_{\text{data}}} [\log D_G^*(x)] + E_{x \sim p_g} [\log(1 - D_G^*(z))] = \\ & E_{x \sim p_{\text{data}}} [\log D_G^*(x)] + E_{x \sim p_g} [\log(1 - D_G^*(x))] = \\ & E_{x \sim p_{\text{data}}} \left[\frac{p_{\text{data}}(x)}{p_{\text{data}}(x) + p_g(x)} \right] + \\ & E_{x \sim p_g} \left[\frac{p_g(x)}{p_{\text{data}}(x) + p_g(x)} \right] = \\ & -\log(4) + \text{KL}(p_{\text{data}}(x) \parallel \frac{p_{\text{data}}(x) + p_g(x)}{2}) + \\ & \text{KL}(p_g(x) \parallel \frac{p_{\text{data}}(x) + p_g(x)}{2}) = \\ & -\log(4) + 2\text{JS}(p_{\text{data}}(x) \parallel p_g(x)) \end{aligned}$$

当 $p_g(x) = p_{\text{data}}(x)$ 时,可以得到全局最小值 $-\log(4)$ 。

4 实验过程及结果分析

实验数据集使用深度学习领域最常用的公共数据集—MNIST 手写体数字数据集,该数据集共包含 70 000 张 28×28 像素大小的手写数字的灰度图片,每一张灰度图片都有对应的标签,即图片对应的数字。该实验将数据集分成两部分,其中训练数据集 images_num 有 60 000 张图片,剩下 10 000 张作测试数据集,实验中判别网络的学习率设为 0.001,生成网络的学习率设为 0.001,训练进程 epoch 为 2 000, batch_size 设为 60,共有 1 000 个 batch_num,每一个进程重复训练这 1 000 个 batch_num,由图 6 可以看出,当随着训练进程 epoch 的增大,损失函数发生突变。这是由于在每一个 epoch 中,网络前向传播输出值,神经网络的参数太多,发生错误时进行反馈,训练时间足够长时,网络会记住整个训练细节,造成训练过度梯度消失。

为了防止训练过度梯度消失,在生成网络和判别网络的每一个卷积层之后均使用批量归一化操作,批量归一化层对网络中的激活值和梯度传播进行归一化,使网络训练成为更简单的优化问题。卷积层之后使用批量归一化层,来加速网络训练并降低对网络初始化的敏感度,可以防止训练时梯度消失,随着训练进程的增加,网络的损失函数也在不断变化,训练进程与模型损失函数的关系如图 6、图 7 所示。

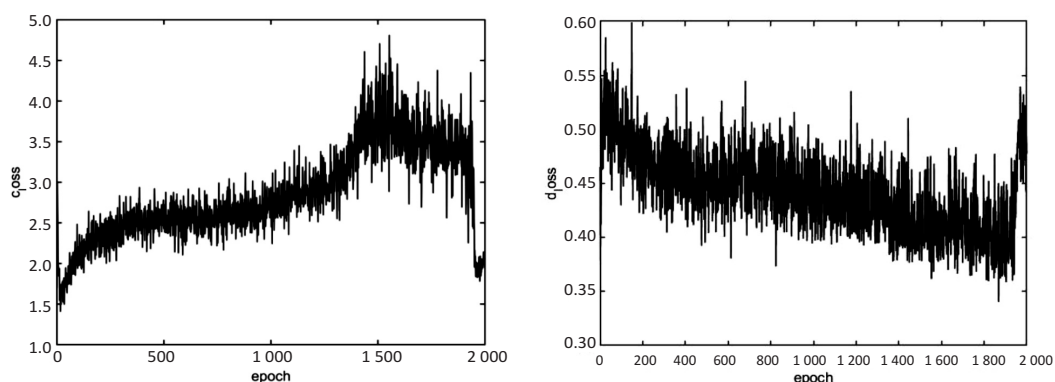


图6 判别器和生成器的损失函数曲线

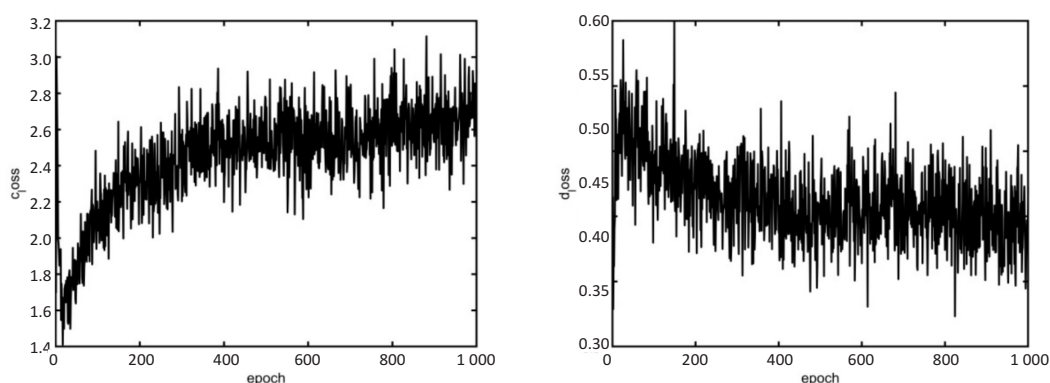


图7 改进后判别器和生成器的损失函数曲线

图7中表示判别网络中正确判断真实图片类别和生成图片类别的准确率,在加入卷积层之后均使用批量归一化操作,损失函数没有出现突变,没有发生梯度消失。另外适当减少训练时间,将训练进程减少到

1 000次,避免造成神经网络过度学习数据集的特征,过度拟合,使生成和判别效果大打折扣。最终判别器损失函数稳定在2.6左右,生成器的损失一直在减小,稳定在0.45左右。

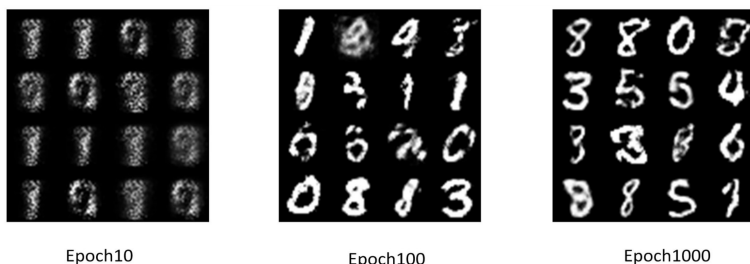


图8 不同迭代次数下生成的手写字体

图8表示在不同的训练迭代次数下生成网络生成的图片。可以看出,开始10个进程后,可以从随机噪声到隐约看出数字形状,再到100个进程字体雏形显

现,生成的字体没有模糊,边界清晰。实验证明,该方法能够完成手写字体图片的生成。



图9 图像生成效果对比

图 9 表示在不同网络下生成的图片,左侧是 VAE 生成图像效果图,中间是 GAN 生成图像效果图,右侧是文中网络生成图像效果图。能够明显看到,经过相同的训练轮数,VAE 生成的图像比较规范,基本没有噪点和乱点,但是图像模糊,缺少细节;GAN 生成的图像相比 VAE 细节明显,但是存在噪点和乱点;文中提出的网络生成的图像则更加清晰,且噪点和乱点相比 GAN 更少。

除此之外,实验还选取了基于 RNN、PCA、SVM 图像识别,对比这些方法和文中方法在图像识别的准确率。不同方法在 MNIST 数据集上的识别准确率对比如表 1 所示。

表 1 不同方法的识别率

图像识别方法	识别准确率/%
SVM	90.36
PCA	91.77
RNN	96.86
Our method	98.55

从表 1 可知,使用 SVM 和 PCA 的准确率可以达到 91% 左右,RNN 方法的准确率可以达到 96% 左右,文中方法图像识别的准确率可以稳定保持在 98% 左右。根据对比结果,文中方法有更好的性能,生成的图片更易被识别。

5 结束语

基于深度卷积神经网络和生成对抗网络,引入变分自编码器,深度卷积神经网络解决了传统人工神经网络特征提取不准确,造成生成图片质量不高的问题;变分自编码器的引入,充分发挥数据中的隐空间信息对图片生成的促进作用。同时模型的改进,大大提高了网络分类准确率。该模型的生成方法,弥补了训练数据不足带来的识别精度下降的缺点,除此之外,该模型的扩展还可以应用到其他图形。但是该方法只在 MNIST 数据集上进行验证,其稳定性和生成性能还需要在其他数据集上作进一步的研究。

参考文献:

- [1] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]//International conference on neural information processing systems. Massachusetts: MIT, 2014:2672-2680.
- [2] CHENG Gong, ZHOU Peicheng, HAN Junwei, et al. Auto-encoderbased shared mid-level visual dictionary learning for scene classification using very high resolution remote sensing image[J]. IET Computer Vision, 2015, 9(5): 639-647.
- [3] MIRZA M, OSINDERO S. Conditional generative adversarial nets[J]. Computer Science, 2014, 27(8): 2672-2680.
- [4] 李明宇. 基于变分自动编码器的特征表示学习研究及其应用[D]. 哈尔滨: 哈尔滨工业大学, 2018.
- [5] WANG Ya-si, YAO Hongxun, ZHAO Sicheng. Auto-encoder based dimensionality reduction[J]. Neurocomputing, 2016, 184: 232-242.
- [6] KINGMA D P, REZENDE D J, MOHAMED S, et al. Semi-supervised learning with deep generative models[C]//Advances in neural information processing systems. Montreal, Canada: MIT Press, 2014: 3581-3589.
- [7] 李 策, 陈海霞, 汉语, 等. 深度学习算法中卷积神经网络的概念综述[J]. 电子测试, 2018(23): 61-62.
- [8] 胡 硕, 赵银妹, 孙 翔. 基于卷积神经网络的目标跟踪算法综述[J]. 高技术通讯, 2018, 28(3): 207-213.
- [9] 刘志宏, 李玉峰. 基于特征融合卷积神经网络的 SAR 图像目标检测方法[J]. 微处理机, 2020, 41(2): 31-37.
- [10] 李彦冬, 郝宗波, 雷 航. 卷积神经网络研究综述[J]. 计算机应用, 2016, 36(9): 2508-2515.
- [11] 卢宏涛, 张秦川. 深度卷积神经网络在计算机视觉中的应用研究综述[J]. 数据采集与处理, 2016, 31(1): 1-17.
- [12] 唐贤伦, 杜一铭, 刘雨微, 等. 基于条件深度卷积生成对抗网络的图像识别方法[J]. 自动化学报, 2018, 44(5): 855-864.
- [13] 王坤峰, 左旺孟, 谭 营, 等. 生成式对抗网络: 从生成数据到创造智能[J]. 自动化学报, 2018, 44(5): 769-774.
- [14] 陈 霖, 段 巍, 刘立志. 基于 GAN 网络的面部表情识别[J]. 电子技术与软件工程, 2020(1): 3-4.
- [15] 黄利文. 一种判别模型的择优方法[J]. 江西理工大学学报, 2013, 34(1): 96-99.