

基于 DCGAN 数据增强的水产动物分类方法

王德兴,秦恩倩,袁红春
(上海海洋大学信息学院,上海 201306)

摘要:针对公开大规模水产动物数据集少、人为采集数据工作量大以及传统数据增强方法对数据的特征提升有限的问题,提出一种基于深度卷积生成对抗网络的数据增强方法用于水产动物图像识别。首先,使用深度卷积生成对抗网络(DCGAN)对样本数据进行增强,然后分别使用 VGG16、InceptionV3、ResNet50 这三个训练模型,以微调的方式,对样本进行训练、识别。结果显示,所提出的方法在水产动物数据集上,与非生成式的数据增强方法相比,在3种模型上分类的准确率可分别提高9.8%、2.7%、1.2%。试验证实,DCGAN可有效增强水产动物图像数据,提高深度神经网络模型对水产动物图像分类的准确率。

关键词:深度卷积生成对抗网络;数据增强;水产动物分类

中图分类号:S917.4

文献标志码:A

文章编号:1007-9580(2019)06-068-08

中国是水产养殖大国,多年来,水产养殖业主要依赖粗放经营的传统方式,不仅效率低、缺乏科学性,且存在较强的主观性,随着社会经济的发展,水产养殖信息化已成为必然趋势,信息化技术对现代化渔业的可持续发展具有重要意义^[1-2]。为制订相应的信息化管理措施,水产养殖需要进行高效的图像识别工作,随着深度学习的发展,目前深度卷积神经网络(Deep Convolutional Neural Network, DCNN)在图像分类领域的应用已经取得巨大成功。如 Krizhevsky 等^[3]训练一个大型 DCNN,将 LSVRC-2010 ImageNet 训练集中的 130 万个高分辨率图像分类为 1 000 个不同的类别;He 等^[4]针对更深层次神经网络难以训练的问题,提出一种深度残差学习用于图像识别,在 ImageNet 测试集上的误差只有 3.57%。基于卷积神经网络的模型要取得较高的识别率或好的分类效果,往往取决于大规模带有标签的数据集,而利用较小规模数据集训练卷积神经网络(CNN)模型通常会出现过拟合现象^[5]。针对水产动物图像、数据集的采集需要耗费大量的人力物力,获得合适的训练数据往往是有限的,严重限制了这一技术在水产养殖领域的应用。

针对这一问题,需要对样本数据进行增强,人为扩大数据集。在数据增强研究方面, Bjerrum

等^[6]利用仿射变换生成新样本,再与原始样本混合,作为新的训练集输入到神经网络中进行分类。近年来,出现了很多更复杂的数据增强方法^[7],如: Allken 等^[8]为了弥补训练数据的不足,以实现一个可以生成合成图像作为训练数据的模拟器,图像由真实图像的分量随机组合而成,处理速度快、直观; Goodfellow 等^[9]提出的生成对抗网络(Generative Adversarial Network, GAN)是一种生成式模型,通过生成器与判别器的不断对抗,达到一种纳什均衡^[10],最终生成符合真实样本分布的数据,具有比传统深度学习更强大的特征学习和表达能力,但模型本身容易出现训练不稳定、难以训练等问题。

提出了一种基于深度卷积生成对抗网络的数据增强方法,在水产动物的分类中引入该网络模型,将传统的卷积神经网络与新兴的生成对抗网络结合,通过对已有的少量样本特征提取并实现样本重建,解决训练样本数据不足以及特征不平衡的问题,生成更多可用的样本提供给深度学习分类网络训练。

1 相关数据增强方法

1.1 非生成式数据增强方法

在图像分类任务中,图像数据增强是很常见

收稿日期:2019-07-25

基金项目:国家自然科学基金资助项目(41776142);上海市科学技术委员会支撑项目(1439190400)

作者简介:王德兴(1968—),男,副教授,博士,研究方向:人工智能、数据挖掘等。E-mail:dxwang@shou.edu.cn

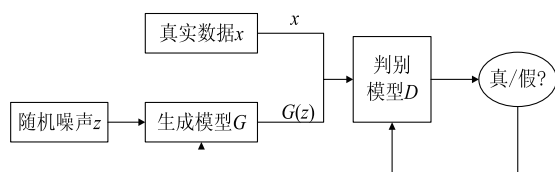
的方法,这是由于深度学习对数据集的规模有一定的要求,如果原始数据集较小,则网络模型易出现过拟合现象,很难训练出良好的结果,直接影响模型的性能。图像数据增强通过对原始图像进行一定的处理来扩充数据集,能够在一定程度上提升模型的性能。根据生成样本的方式,将数据增强方法分为两类:非生成式数据增强方法和生成式数据增强方法。常用的非生成式数据增强方法有以下几种:

(1)几何变换。包括对原图片的平移、缩放、翻转或者镜像^[12]、旋转^[13]和随机裁剪。同时对图片做平移、缩放、翻转、旋转、裁剪等多重操作称为仿射变换。

(2)颜色抖动。指对颜色的数据增强,包括对原图片的亮度、饱和度、对比度等变化。

(3)随机噪声。在原来图像上加一些随机生成的“误差”,就能构成新的图片,常见的方法有:高斯噪声、椒盐噪声。其中高斯噪声是指其概率分布函数服从高斯分布的噪声。

常用的生成式数据增强方法即生成对抗网络以及它的衍生模型,如深度卷积生成对抗网络、条件生成对抗网络等。其中生成对抗网络模型的基本结构如图1所示^[11]。



注: z 为输入的随机噪声, G 为生成器, $G(z)$ 表示生成的图片, 即假样本; x 为真实数据, D 为判别器

图1 生成对抗网络基本结构

Fig.1 Framework of GAN

1.2 深度卷积生成对抗网络

深度卷积生成对抗网络(DCGAN)是GAN的一种变体^[14],是在GAN的基础上,将生成器与判别器中的网络由改进的卷积神经网络替换,并结合带泄露线性整流(Leaky rectified liner unit, Leaky ReLU)激活函数和双曲正切(tanh)激活函数的一种网络结构。DCGAN中的 D 、 G 的含义以及目标函数都和原始GAN中完全一致,但是它在

D 和 G 中采用了较为特殊的结构,以便对图片进行有效建模。

生成模型的网络结构是一个类似反卷积的神经网络,其输入层是一个服从均匀分布的100维随机向量,而不是普通卷积神经网络的一张图;其次,卷积操作使用4个微步幅卷积(或称分数步长卷积),类似于用反卷积来代替池化层,学习自己的空间上采样。设置微步幅卷积的步长为2,那么生成网络经过每一次微步幅卷积,通道数减半,图像尺寸加倍。在除了输出层以外的每一层加上批量规范化(BN)^[15]的操作,能够解决糟糕的初始化问题,避免生成模型崩溃(这是GAN经常遇到的失败现象)。同时,将特征层的输出归一化能够让训练更加稳定,此外还能让梯度向更深层次的网络传播。最后输出层的激活函数采用tanh,其他层设为线性整流(Rectified Linear Unit, ReLU)激活函数,输出一个 64×64 像素的3通道RGB图像。整个网络结构没有池化层,在卷积特征之上没有全连接层(表1)。

表1 DCGAN 生成模型

Tab.1 Generation model of DCGAN

网络层	卷积核	步长	填充值	激活函数	是否加入 BN
输入层	4×4	0	0	ReLU	是
Conv 1	5×5	2	1	ReLU	是
Conv 2	5×5	2	1	ReLU	是
Conv 3	5×5	2	1	ReLU	是
Conv 4	5×5	2	1	ReLU	是
输出层	5×5	2	1	tanh	否

注: Conv 代表卷积层,单位均为像素,下同

判别模型网络结构与卷积神经网络类似(表2)。卷积层使用带步长的卷积取代空间池化,实现空间上采样。与生成模型类似,对除了输入层以外的每一层数据进行批量规范化操作,使其均值变为0,方差变为1,缓解训练中的梯度消失,加快模型的训练速度。为了防止梯度稀疏,激活函数使用 Leaky Relu 函数替代 Relu 函数。最后,通过一个全连接层输出一个标量,范围为0到1,用来表示输入数据属于真实数据而不是生成样本的概率。

表 2 DCGAN 判别模型

Tab. 2 Discriminator model of DCGAN

网络层	卷积核	步长	填充值	激活函数	是否加入 BN
输入层	5×5	2	1	LeakyReLU	否
Conv1	5×5	2	1	LeakyReLU	是
Conv2	5×5	2	1	LeakyReLU	是
Conv3	5×5	2	1	LeakyReLU	是
Conv4	5×5	2	1	LeakyReLU	是
输出层	4×4	0	0	LeakyReLU	是

深度卷积生成对抗网络与生成对抗网络的目标函数是一致的,如公式(1)所示^[8]:

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

式中: $p_{data}(x)$ 表示真实样本 x 的分布; $p_z(z)$ 表示随机噪声 z 的分布, $D(G(z))$ 表示 D 网络判断 G 生成的图片是否真实的概率,生成器 G 和判别器 D 通过对抗学习,不断更新参数,最终达到纳什均衡。

本研究将水产动物图像作为真实样本,利用 DCGAN 训练生成新的样本,其网络结构如图 2 所示,DCGAN 的训练是一个交替进行的过程,本质

上是最大最小优化问题,分为两个部分:

1)生成网络的训练。固定判别网络 D ,优化生成网络 G 的参数。由于 $D(G(z))$ 表示 D 网络判断 G 生成的图片是否真实的概率,那么目标函数应最大化 $D(G(z))$,等同于最小化 $1 - D(G(z))$,也就是最小化目标函数 $V(D, G)$,如公式(2)所示^[8]。经过“判断”,判别器会将它的梯度传回 G 来更新生成网络的参数。

$$\min_G V(D, G) = E_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (2)$$

2)判别网络的训练。固定生成网络 G ,优化判别网络 D 的参数。要使得判别网络具有较好的判别能力,那么对于真样本,得到的结果越大越好,即最大化;对于假样本,需要其结果越小越好,也就是 $D(G(z))$ 最小化,即 $1 - D(G(z))$ 最大化。因此训练判别网络时,应最大化目标函数 $V(D, G)$,如公式(3)所示^[8]。通过误差的反向传播,更新判别网络的参数。

$$\max_D V(D, G) = E_{x \sim p_{data}(x)} [\log(D(x))] + E_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (3)$$

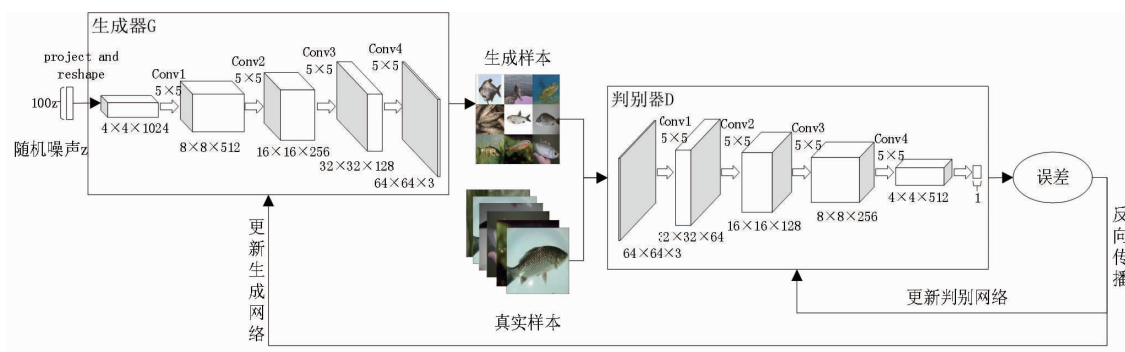


图 2 DCGAN 结构图

Fig. 2 Structure diagram of DCGAN

2 非生成式数据增强方法

2.1 试验环境与数据集

为评估基于 DCGAN 的数据增强方法对水产动物图像识别准确率的提升,采用鱼虾蟹贝数据集进行试验。硬件环境:GPU 为 NVIDIA GTX1080Ti,通过 CUDA8.0 进行加速运算,CPU 为 AMD Ryzen Threadripper 1950X。软件环境:Windows10 专业版,Python3.5,TensorFlow-gpu1.6 深度学习框架^[16]和 Keras2.1 库。

试验所采用的鱼虾蟹贝数据集均来自于本研究背景项目的上海海洋大学滨海养殖基地,每一类各 300 张共计 1 200 张图片作为试验训练集,图 3 为其中部分样本。

2.2 图像识别模型

试验采用的模型为 VGG16^[17]、InceptionV3^[18]和 ResNet50^[4]这 3 个采用 ImageNet 训练集开发的训练模型,为提高模型整体的识别能力,再通过微调的方式使模型针对目标样本进行调整。训练的初始化权值设置 epoch 为 100,每个批次 32 个

样本, Learning rate 为 0.0001, Momentum 为 0.9, 其中 VGG-16 和 ResNet50 模型在训练之前需将图片调整为 224×224 像素大小, InceptionV3 模型需将原始图片全部归一化为 299×299 像素再进行训练, InceptionV3 和 ResNet50 模型在中间层的每一层加上 Batch-Normalization 操作(表 3)。

2.3 样本增强效果

采用多种非生成方式进行训练数据的增强, 通过翻转、旋转、裁剪、亮度调节、添加高斯噪声等方法, 将每张原始图片随机变换为 32 张不同的变形图。一方面增加了训练的数据量, 可以提高模型的泛化能力, 另一方面增加一些噪声数据, 可提高模型的鲁棒性。图 4 给出了一组对比示例, 分别为原始图和增强后的效果图。

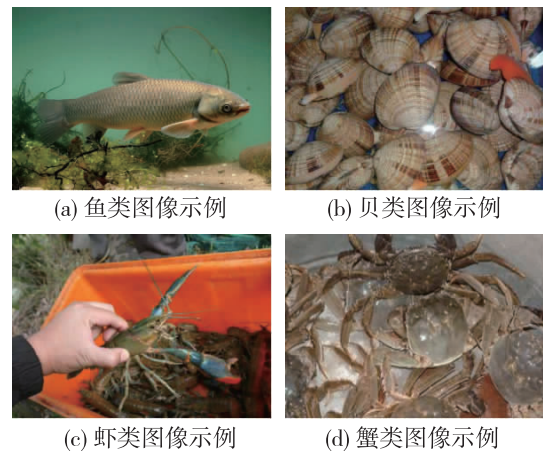


图 3 水产动物图像示例

Fig. 3 Examples of aquatic animal images

表 3 各网络模型训练参数

Tab. 3 Training parameters of each network model

网络模型	Ecoph	Batch-size	Learning rate	Momentum	Input-tensor	BN
VGG-16	100	32	0.0001	0.9	$224 \times 224 \times 3$	—
ResNet50	100	32	0.0001	0.9	$224 \times 224 \times 3$	✓
InceptionV3	100	32	0.0001	0.9	$299 \times 299 \times 3$	✓

注: Ecoph 为迭代次数, Batch-size 为批处理大小, Learning rate 为学习率, Momentum 为动量, Input-tensor 为对应模型所接受原始图像的维度, BN 为批量规范化操作

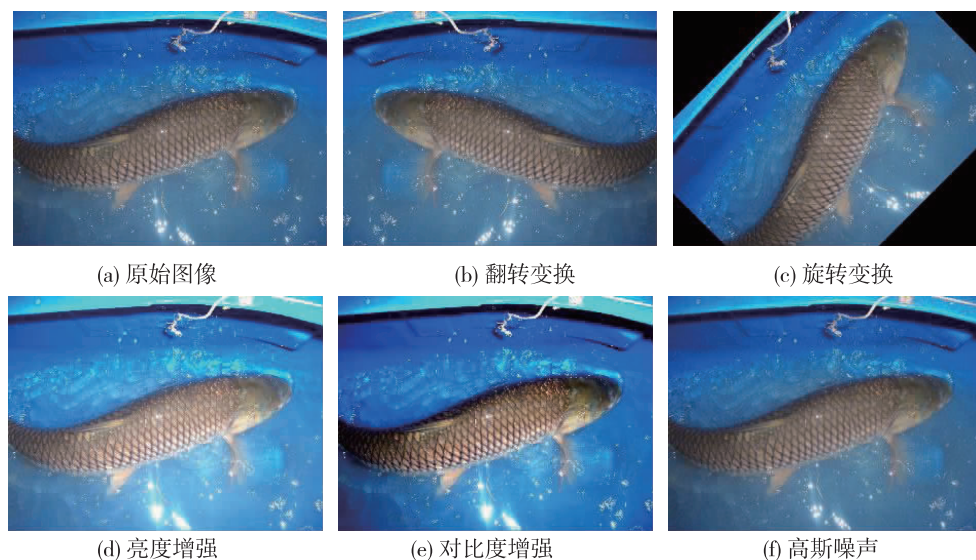


图 4 非生成式数据增强方法示例

Fig. 4 Examples of non-generative data enhancement methods

3 基于 DCGAN 的数据增强方法

3.1 数据预处理

基于 DCGAN 的数据增强方法与非生成式数

据增强方法进行水产动物图像识别的试验环境、原始数据集以及训练使用的模型完全一致。为提高模型整体的识别能力, 同样采用微调方式进行调整。由于拍摄角度、拍摄距离变化、拍摄设备性

能好坏等因素,水产动物图像中可能会有许多的背景噪声,同时,待识别物体在拍摄图像中可能只占小部分,而用 DCGAN 生成尽可能类似真实样本的水产动物图像,需要原始数据集中目标物体特征尽可能明显,背景噪声的存在将会对生成图像有负面影响。因此,提取水产动物图像的感兴趣区域(ROI)是非常有必要的^[19]。

由于拍摄的水产动物图片中,目标物体几乎在图片的中心,所以以中心为原点,对图片进行剪切,大小为 108×108 像素,再根据模型所接受的原始图片的像素要求,将所有图片重新调整尺寸为 64×64 像素。图 5 是以数据集中一张原始的图片为例进行 ROI 图像提取的结果。

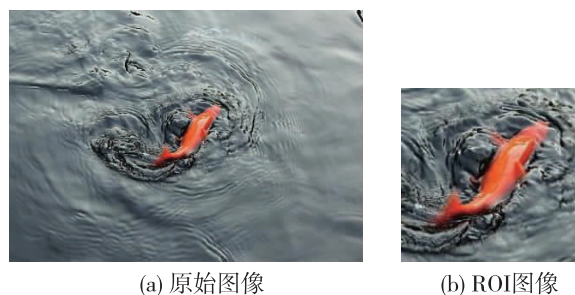


图 5 ROI 图像示例

Fig. 5 Examples of ROI image

3.2 参数配置

适当地初始化权值可以更有效地训练神经网络。DCGAN 模型的学习率为 0.0002,动量设为 0.5,批处理大小设置为 64,生成器的输入为 $N(0,1)$ 分布的噪声。使用 Adam 优化器,学习率设置为 $1e-4$ 。其余参数使用 Tensor Flow 的默认设置。经过多次试验对比,发现在迭代 500 次以后损失函数震荡幅度较大,模型易崩溃,极不稳定,所以 DCGAN 训练迭代次数设置为 500。

3.3 生成样本效果

训练迭代次数与损失函数的变化如图 6 所示。图中的(a)、(b)分别表示 DCGAN 在鱼虾蟹贝数据集上判别器判断样本是真的损失函数(d_loss_real)和判断样本是假的损失函数(d_loss_fake)随着训练次数增加而变化的情况,(c)表示 DCGAN 在鱼虾蟹贝数据集上判别器的损失函数(d_loss)随着训练次数增加而变化的情况,(d)表示 DCGAN 在鱼虾蟹贝数据集上生成器的损失函数(g_loss)变化情况。

从图 6(a)、(b)可以看出判别器判断样本是真的损失函数在缓慢上升,判别样本是假的损失函数在缓慢下降,随着生成图像与真实样本越来越相似,判别器判别能力减弱,生成器生成的样本被判别器识别为真的概率越来越大,总体震荡较大,(a)、(b)的损失函数的和即为(c)中的损失函数,即 $d_loss = d_loss_real + d_loss_fake$ 。从(c)可以看出判别器的损失函数随着迭代次数变化在初期较为稳定,随着迭代次数增加,在一段缓慢地上升之后,整体呈现缓慢的下降趋势,判别器的判别能力增强。从(d)可以看出生成器的损失函数在训练初期值较大且变化范围大,生成器生成的图片判别器很容易判别为假,迭代 150 次以后,损失函数值下降到较小值且趋于稳定,这个时期生成器能够生成较为清晰的图片。总体来看,两个函数震荡变化,生成器与判别器这两个网络结构在相互制约、平衡。

以数据集中鱼类为例,图 7 为数据集前 500 个生成样本在 3 个阶段(第 0、200、500 Epoch)生成样本的情况。可以看出,随着 Epoch 次数的增加,生成样本的轮廓越来越清晰,特征也越来越明显。

4 结果与分析

经过非生成式数据增强方法的扩增,每一类数据扩增至 3 200 张,分别使用 VGG16、InceptionV3 和 ResNet50 模型进行训练,记录每次训练完成后图像分类的准确率,经过多次试验,取其中准确率最高的一次作为评估模型识别性能的依据;经过 DCGAN 数据增强方法的扩增,采用同样的方法进行训练,试验结果见表 4。两种方法对比,可以看出经过 DCGAN 模型的数据增强^[14]后,比用传统的旋转、剪切、翻转变化的数据增强方法^[6]在 3 个不同的模型上用于水产动物图像识别的准确率都有所提高。在 VGG16 模型上准确率提高 9.8%,在 InceptionV3 模型上准确率提高 2.7%,在 ResNet50 模型上准确率提高 1.2%。相比较而言,对于用非生成式的数据增强方法扩增样本后准确率较低的模型,准确率提升较高,而在 InceptionV3、ResNet50 这两个用非生成式的数据增强方法扩增样本后准确率达到 95% 以上的模型,准确率的提升空间较小,总体均有所提高。

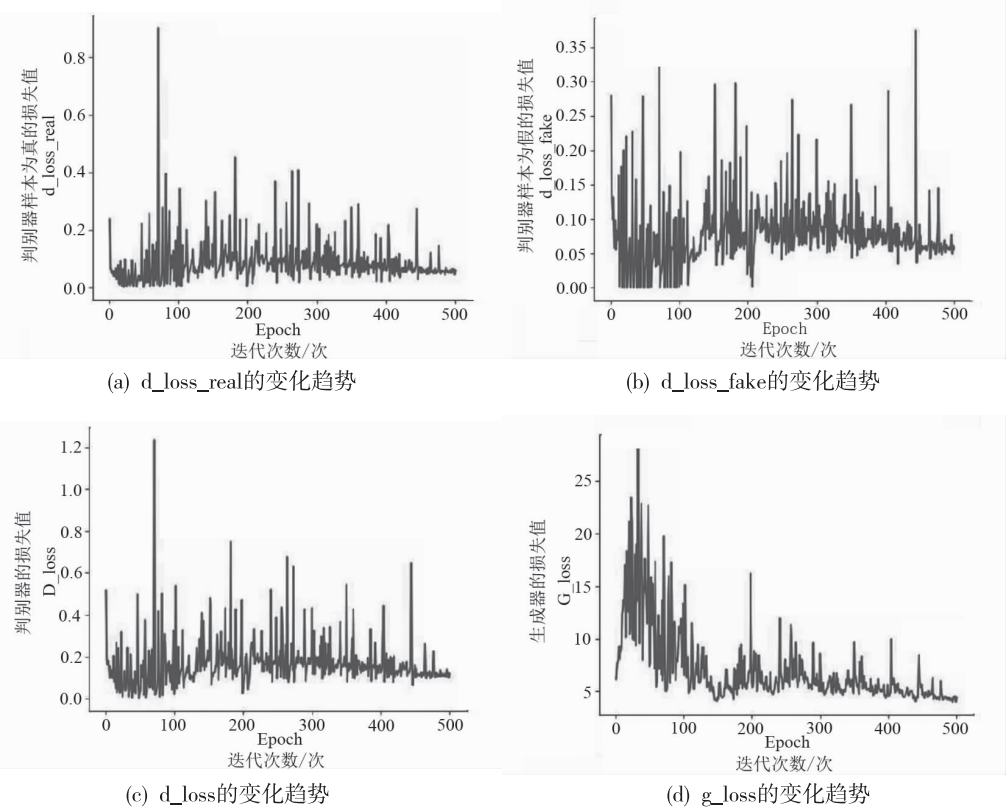


图 6 生成器和判别器的损失函数变化趋势

Fig. 6 Trend of loss function of generator and discriminator

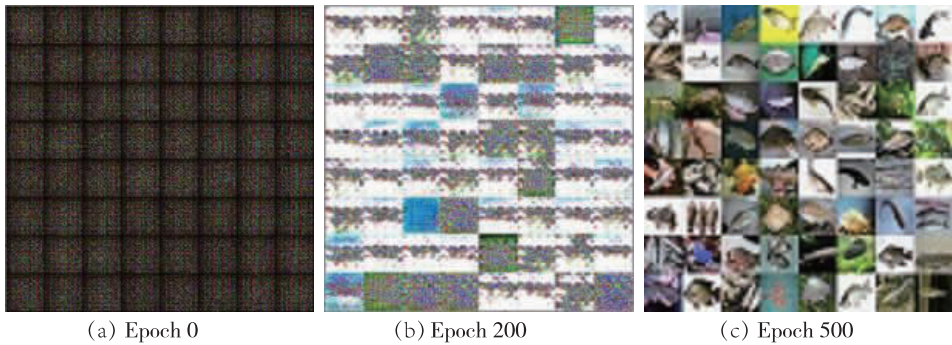


图 7 DCGAN 生成样本

Fig. 7 Samples generated by DCGAN

表 4 各模型准确率对比

Tab. 4 Accuracy comparison of different models

数据增强方法	网络结构	准确率
非生成式	VGG16	73.3%
	InceptionV3	95.7%
	ResNet50	98.1%
DCGAN	VGG16	83.1%
	InceptionV3	98.4%
	ResNet50	99.3%

5 结论

从数据增强的角度,提出一种基于深度卷积生成对抗网络的数据增强方法用于水产动物识别,能解决样本不足、特征不明显的问题。通过对比传统的非生成式数据增强方法,该方法适用于对水产动物图像的增强,有助于提高不同模型识别的准确率;还有助于提高图片生成的速度以及生成图片的质量。 □

参考文献

- [1] 胡金有,王靖杰,张小栓,等.水产养殖信息化关键技术研究现状与趋势[J].农业机械学报,2015,46(7):251-263.
- [2] 杨宁生,袁永明,孙英泽.物联网技术在我国水产养殖上的应用发展对策[J].中国工程科学,2016,18(3):57-61.
- [3] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks [C]. Advances in neural information processing systems. New York: Curran Associates, 2012: 1097-1105.
- [4] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [C]. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016. 770-778.
- [5] 陈文兵,管正雄,陈允杰.基于条件生成式对抗网络的数据增强方法[J].计算机应用,2018,38(11):3305-3311.
- [6] BJERRUM E J. Smiles enumeration as data augmentation for neural network modeling of molecules [J/OL]. Arxiv Preprint, 2017(3):1-7. (2017-03-21) [2019-11-16]. <https://arxiv.org/pdf/1703.07076.pdf>.
- [7] TAYLOR L, NITSCHKE G. Improving deep learning using generic data augmentation [J/OL]. Arxiv Preprint, 2017(8):1-6. (2017-08-20) [2019-11-16]. <https://arxiv.org/pdf/1708.06020.pdf>.
- [8] ALLKEN V, HANDEGARD N O, ROSEN S, et al. Fish species identification using a convolutional neural network trained on synthetic data [J]. ICES Journal of Marine Science, 2018, 76(1): 342-349.
- [9] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets [C]. Proceedings of the 27th International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2014: 2672-2680.
- [10] RATLIFF L J, BURDEN S A, SASTRY S S. Characterization and computation of local Nash equilibria in continuous games [C]. Proceedings of the 51st Communication, Control, and Computing (Allerton). Monticello, IL, USA: IEEE, 2013. 917-924.
- [11] 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.
- [12] WANG Y C, LI M T, PAN Z C, et al. Pulsar candidate classification with deep convolutional neural networks [J]. Research in Astronomy and Astrophysics, 2019, 19(9): 119-128.
- [13] WAN L, ZEILER M, ZHANG S, et al. Regularization of neural networks using dropconnect [C]. International conference on machine learning. Atlanta: IMLS, 2013: 1058-1066.
- [14] RADFORD A, METZ L, CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks [J/OL]. Arxiv Preprint, 2016(11):1-6. (2015-11-19) [2019-11-16]. <https://arxiv.org/pdf/1511.06434.pdf>.
- [15] IOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift [J/OL]. Arxiv Preprint, 2015(2):1-11. (2015-02-11) [2019-11-16]. <https://arxiv.org/pdf/1502.03167.pdf>.
- [16] ABADI M, AGARWAL A, BARHAM P, et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems [J/OL]. Arxiv Preprint, 2016(3):1-19. (2016-03-14) [2019-11-16]. <https://arxiv.org/pdf/1603.04467.pdf>.
- [17] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J/OL]. Arxiv Preprint, 2014(9):1-14. (2014-09-04) [2019-11-16]. <https://arxiv.org/pdf/1409.1556.pdf>.
- [18] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the Inception Architecture for Computer Vision [C]. Computer Vision and Pattern Recognition. IEEE, 2016: 2818-2826.
- [19] 肖志涛,王红,张芳,等.复杂自然环境下感兴趣区域检测[J].中国图象图形学报,2015,20(5):625-632.

Classification of aquatic animals based on DCGAN-based data enhancement

WANG Dexing, QIN Enqian, YUAN Hongchun

(College of Information Technology, Shanghai Ocean University, Shanghai 201306, China)

Abstract: In order to solve the problems such as lack of large-scale aquatic animal data set, heavy workload of artificial data collection and limited feature enhancement of data by traditional data enhancement methods, a data enhancement method based on deep convolutional generative adversarial network was proposed for aquatic animal image recognition. First of all, deep convolutional generative adversarial network (DCGAN) was used to enhance the sample data, and then three training models: VGG16, InceptionV3 and ResNet50 were used to train and identify the samples by fine tuning. The results show that the classification accuracy of the proposed method on the aquatic animal data set can be improved by 9.8%, 2.7% and 1.2% respectively compared with the non-generated data enhancement method. The experiment proves that DCGAN can effectively enhance the image data of aquatic animals and improve the accuracy of deep neural network model in classifying aquatic animal images.

Key words: deep convolutional generative adversarial network; data enhancement; classification of aquatic animals