

基于FTVGG16卷积神经网络的鱼类识别方法

陈英义^{1,2} 龚川洋^{1,3} 刘烨琦^{1,3} 方晓敏^{1,3}

(1. 中国农业大学信息与电气工程学院, 北京 100083; 2. 农业农村部信息获取技术重点实验室, 北京 100083;
3. 北京农业物联网工程技术研究中心, 北京 100083)

摘要: 针对大多数应用场景中,大多数鱼类呈不规则条状,鱼类目标小,受他物遮挡和光线干扰,且一些基于颜色、形状、纹理特征的传统鱼类识别方法在提取图像特征方面存在计算复杂、特征提取具有盲目和不确定性,最终导致识别准确率低、分类效果差等问题,本文在分析已有的VGG16卷积神经网络良好的图像特征提取器的基础上,使用ImageNet大规模数据集上预训练的VGG16权重作为新模型的初始化权重,通过增加批规范层(Batch normalization, BN)、池化层、Dropout层、全连接层(Fully connected, FC)、softmax层,采用带有约束的正则权重项作为模型的损失函数,并使用Adam优化算法对模型的参数进行更新,汲取深度学习中迁移学习理论,构建了FTVGG16卷积神经网络(Fine-tuning VGG16 convolutional neural network, FTVGG16)。测试结果表明:FTVGG16模型在很大程度上能够克服训练的过拟合,收敛速度明显加快,训练时间明显减少,针对鱼类目标很小、背景干扰很强的图像,FTVGG16模型平均准确率为97.66%,对部分鱼的平均识别准确率达到了99.43%。

关键词: 鱼类识别; 卷积神经网络; 迁移学习; 模式识别

中图分类号: TP391 文献标识码: A 文章编号: 1000-1298(2019)05-0223-09

Fish Identification Method Based on FTVGG16 Convolutional Neural Network

CHEN Yingyi^{1,2} GONG Chuanyang^{1,3} LIU Yeqi^{1,3} FANG Xiaomin^{1,3}

(1. College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China

2. Key Laboratory of Agricultural Information Acquisition Technology, Ministry of Agriculture and Rural Affairs, Beijing 100083, China

3. Beijing Engineering and Technology Research Center for Internet of Things in Agriculture, Beijing 100083, China)

Abstract: Computer vision technology is widely applied in fish individual identification. Nevertheless, there are some problems such as small fish targets, occlusion of objects and light interference in videos and images. Some fish identification methods based on color, shape and texture also exit complicated calculations in feature extraction, such as non-migration of features will result in low recognition accuracy and poor classification. With the help of analysis of image feature extraction of the existing VGG16 convolutional neural network model, the FTVGG16 convolutional neural network (Fine-tuning VGG16 convolutional neural network) was designed. As it was known, the basic deep learning tool used in this work was convolutional neural networks. The FTVGG16 convolutional neural network was composed of convolutional layers, batch normalization layers, pooling layers, Dropout layers, fully connected layers and softmax layers. The experimental results showed that the average recognition accuracy of the FTVGG16 model for fish was about 97.66%, and the average recognition rate of some fishes could reach 99.43%. It had high recognition accuracy and robustness in pictures with small fish targets and strong background interference. It could be operated through an appropriate, easy-to-use, and user-friendly web application for the specific case of fish identification.

Key words: fish identification; convolutional neural network; transfer learning; pattern recognition

0 引言

鱼类的计算机视觉识别一直是该领域的研究热

点,多年来,许多学者对其进行了深入研究,尤其是基于颜色、形状、纹理等图像内容的鱼类识别研究较多,并取得了较大进展。STRACHAN等^[1]利用鱼的

形状进行分类。LARSEN 等^[2]使用线性判别分析方法(Latent dirichlet allocation, LDA)对鱼种形状和纹理特征进行分类,准确率为 76%。HUANG 等^[3]提出 BGOT(Balance guaranteed optimized tree)算法,对采集的 10 个鱼种 3 179 幅图像数据进行分类,准确率为 95%。吴一全等^[4]提出基于最小二乘支持向量机(Least squares support vector machine, LSSVM)的识别方法,准确率为 95.83% 以上。杜伟东等^[5]提出了基于支持向量机(Support vector machine, SVM)的决策融合鱼类识别方法,准确率在 90% 以上。

但是,随着机器视觉技术的发展,数据量剧增,基于内容的识别方法在特征提取方面存在的计算复杂、特征不可迁移性^[6]等问题愈发突出。特别是对图像中存在目标小、他物遮挡、光线干扰等情况的图像识别上,存在识别准确率低或运算效率低等问题,不能满足实际需求。深度学习^[7]中卷积神经网络^[8,11](Convolutional neural network, CNN)的出现已经彻底改变了图像识别^[8-9]、语音识别^[10]以及大数据处理在农业中的应用方式。刘德营等^[12]提出了一种基于卷积神经网络的白背飞虱自动识别方法;傅隆生等^[13]提出了一种基于 Let-Net5^[14]卷积神经网络的田间多簇猕猴桃果实图像识别深度学习模型;王璨等^[15]提出了一种基于卷积神经网络提取多尺度分层特征的田间玉米杂草识别方法,为本研究的开展提供了新思路。

本文在上述研究成果的基础上,通过分析已有的 VGG16^[16]卷积神经网络模型良好的图像特征提取功能,设计 FTVGG16 卷积神经网络,提高复杂应用场景鱼类目标的识别精度,为构建可靠、精准、快

速的鱼类识别系统提供理论基础。

1 材料与方法

1.1 图像获取

图像数据来自大自然保护协会提供的远洋捕捞船上摄像头拍摄的场景^[17]。该协会提供了 3 777 幅带标注的鱼类图像和 1 000 幅不带标注的鱼类图像。部分海鱼类图像示例如图 1 所示,图像包含 8 类目标,其中 7 类是不同种类的海鱼(目标)和 1 类不包括鱼类目标的背景,每幅图像只属于某一类别,构建的 FTVGG16 模型需要识别出每幅图像中包含的鱼类目标。

1.2 图像预处理

在深度学习领域,通过对图像进行预处理可以增强数据样本,进而提高模型性能。本文图像预处理主要包括:

(1)图像像素值处理,即将 JPG 格式的图像转换为 RGB 三通道的矩阵向量存储。映射图像像素值从 $[0, 255]$ 到 $[0, 1]$,这样处理可以使模型在使用激活函数时,避免神经元进入“死区”而导致神经元失活,最终使卷积神经网络获得比较快的收敛速度。

(2)图像空间尺度变换,即将图像进行随机缩放、平移、旋转、镜像、随机裁剪。

(3)图像均值与方差均衡,即样本归一化。图像预处理的结果示例如图 2 所示。预处理后图像尺寸为 600 像素 \times 600 像素。

1.3 试验方法设计

模型训练采用的环境是 Ubuntu 16.04 LTS 操作系统、运行内存为 16GB、NVIDIA 1080 Ti 显卡 GPU、并行 CUDA^[18]编程平台、Keras^[19]的深度学习框架



图 1 鱼类图像样本

Fig. 1 Fish image samples

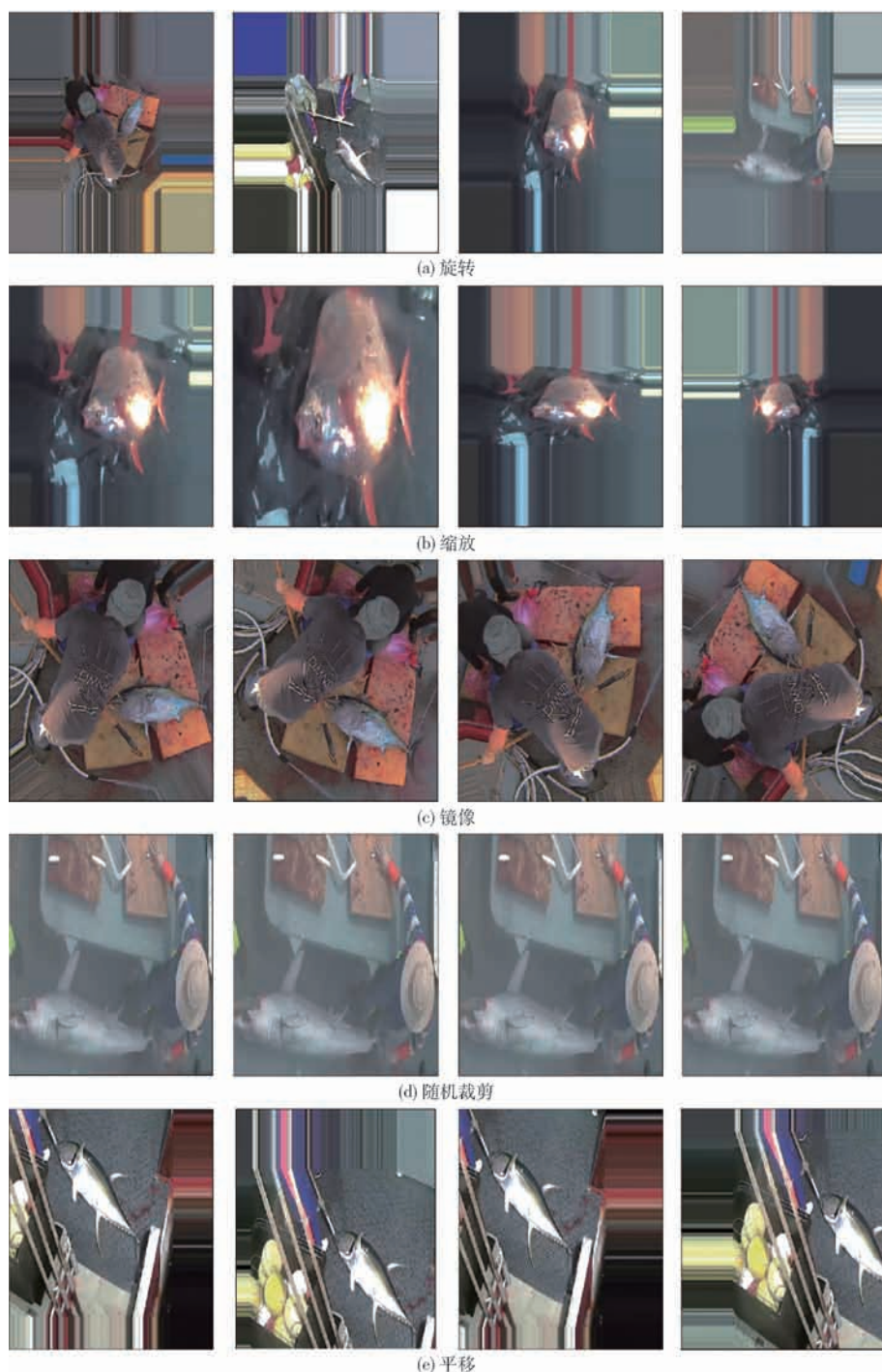


图2 鱼类图像预处理示例

Fig. 2 Illustration of fish image processing

和 Python 3.6.5 编程语言。试验方法设计包括模型训练、测试、对比验证。

(1) 模型训练: 经过预处理后的图像输入到卷积神经网络进行训练, 保存训练后的模型参数与结构。模型训练采用 k 折交叉验证的方式, k 取 5, 将整个数据集的 $1/k$ 作为测试集, 每个模型训练 k 次, 测试 k 次, 误差取 k 次的平均值, 选择误差最小的模型。5 折交叉验证示意图如图 3 所示。数据集的划分使用 Python 脚本程序, 提供划分数据集的比例, 程序自动切分数据集以及随机打乱样本。模型每次

训练都会随机地从数据集中划分 80% 的训练集和 20% 的验证集。2 953 幅图像作为训练集, 用来训练模型, 741 幅图像作为验证集, 用于调优模型参数、验证模型, 1 000 幅不带标注的图像作为测试集, 用来评价模型在未知数据集上的泛化性能, 数据集划分示例如表 1 所示。

(2) 模型测试: 将测试图像输入到模型, 获取输出分类结果, 选取概率最大的目标。基于 5 折交叉验证方法, 最后的预测结果为 5 次预测结果的平均值。

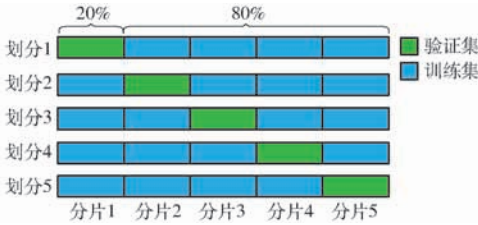


图3 5折交叉验证数据集划分

Fig. 3 Five cross validation data set split

表1 训练和验证数据集

Tab. 1 Summarization training set and validation set

鱼的种类	训练集数目	验证集数目
长鳍金枪鱼	1 375	344
大眼金枪鱼	168	40
海豚鱼	89	23
月光鱼	53	14
鲨鱼	140	36
黄鳍金枪鱼	532	134
带鱼	239	60
背景(负样本)	357	90

(3)对比验证:为验证本模型的识别效果,本文分别选择 VGG16、VGG19^[16]、ResNet50^[20] 进行对比。

2 FTVGG16 卷积神经网络的鱼类个体识别

2.1 FTVGG16 网络模型的构建

FTVGG16 网络由 5 个卷积块、3 个全连接层、1 个 softmax 层构成。每个卷积块由卷积层 Conv、批规范层 (Batch normalization, BN)^[21]、池化层 Pool 组成,每个卷积块中的卷积层数和卷积核的参数不同。例如,在第 1 个卷积块中,“Conv1_64×2”代表第 1 个卷积层中有 64 个卷积核,“×2”代表对上一层的输出进行两次相同的卷积操作运算,卷积后的输出进入 BN 层,BN 层的输出进入激活层,激活层的输出进入最大池化层 Pool,后续卷积块进行相同的操作。图像作为卷积神经网络的输入,依次经过卷积层、批规范层、池化层、全连接层,最后的输出经过分类的 softmax 层进行概率空间映射,输出每种鱼的置信度,从而给出分类结果,FTVGG16 网络架构示例如图 4 所示。

2.1.1 卷积层设计

FTVGG16 卷积神经网络有 13 个卷积层,每个卷积层中的卷积核尺寸是 3×3,卷积核的移动步幅为 1 个像素,像素填充为“0”值填充边界,这样处理能够使卷积前后图像的尺寸保持不变,尽量让卷积核感受图像的边缘信息,每个卷积层中的卷积核数量不固定,卷积核的数量依次为 64、128、256、512。

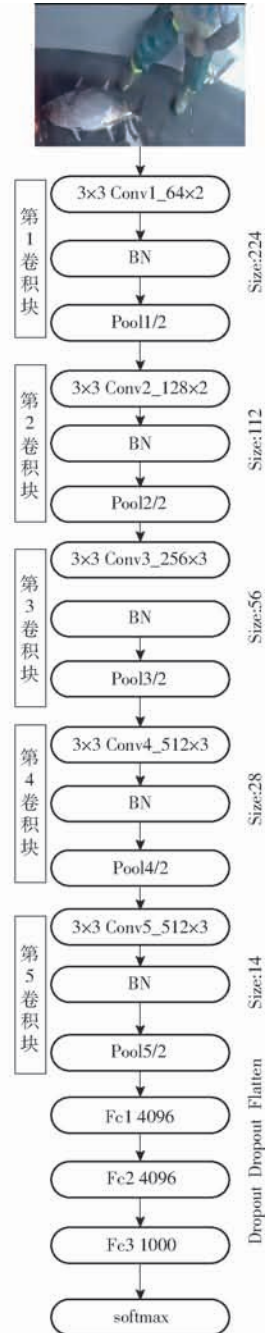


图4 FTVGG16 网络架构

Fig. 4 Neural network framework of FTVGG16

2.1.2 批规范层设计

FTVGG16 卷积神经网络在训练过程中每层输入分布一直在改变,通过在 FTVGG16 网络每个卷积层后面添加批规范层,使得激活值输出的各个维度的均值为 0,方差为 1,尺度缩放和平移操作为了还原最初的输入,批规范化有助于规范化模型,使卷积神经网络更容易学习到数据的分布规律,计算公式为

$$\mu_B = \frac{1}{m} \sum_{i=1}^m x_i \quad (1)$$

$$\sigma_B = \frac{1}{m-1} \sum_{i=1}^m (x_i - \mu_B)^2 \quad (2)$$

$$\mathbf{x}'_i = \frac{\mathbf{x}_i - \boldsymbol{\mu}_B}{\sqrt{\boldsymbol{\sigma}_B^2 + \boldsymbol{\varepsilon}}} \quad (3)$$

$$\mathbf{y}_i = \gamma \mathbf{x}'_i + \beta = \text{BN}_{\gamma, \beta}(\mathbf{x}_i) \quad (4)$$

式中 \mathbf{x}_i ——第 i 个训练样本

m ——批量化数目

$\boldsymbol{\mu}_B$ ——样本均值 $\boldsymbol{\sigma}_B$ ——样本方差

\mathbf{x}'_i ——规范化样本

$\boldsymbol{\varepsilon}$ ——误差项

\mathbf{y}_i ——线性映射以后的值

γ, β ——模型学习参数

$\text{BN}_{\gamma, \beta}(\mathbf{x}_i)$ ——输出

2.1.3 池化层设计

FTVGG16 卷积神经网络有 5 个池化层, 这里使用最大池化层 max-pooling, 池化层主要用于减少图像维度的特征值。池化层的卷积核尺寸是 2×2 , 卷积核的移动步长为 2 像素, 像素不填充, 图像经过 max-pooling 后尺寸会减小到原来的一半。

2.1.4 全连接层设计

FTVGG16 网络的全连接层为 Fc1、Fc2、Fc3、softmax, 全连接层输出的神经元个数分别为 4 096、4 096、1 000、8。Fc3 全连接层的输出经过 softmax 层的映射, 将实值映射为 0 到 1 之间的数值, 即等价于输出每个类别的概率。

2.1.5 Dropout 层设计

Dropout^[22]是指在模型训练过程中, 对于神经元按照一定的概率暂时从网络中丢弃。对于随机梯度下降来说, 由于是随机丢弃, 每批次梯度下降 (Batch gradient descent, BGD) 都在训练不同的网络。Dropout 在神经网络训练的过程中随机抛弃隐藏层的神经元, 断开神经元的连接, 在神经网络反向传播更新参数时, 断开的神经元不参与参数的更新。该技术能够有效地防止神经网络过拟合。分别在全连接层 Fc1、Fc2 后面加入 Dropout 层, Dropout 参数设置为 0.5, 表示有一半的神经元不参与训练。

综上, 设计 FTVGG16 卷积层网络参数如表 2 所示。

2.2 FTVGG16 网络的优化

2.2.1 FTVGG16 网络参数的优化

FTVGG16 采用 Leaky-ReLU^[23] 激活函数, 它的随机梯度下降^[24] (Stochastic gradient descent, SGD) 收敛速度比 sigmoid、tanh 快, 因为是线性的, 而且只需要一个阈值就可以得到激活值, 不需要进行复杂的运算。

2.2.2 损失函数优化

由于识别的是多类别鱼类, 故使用多类别交叉熵^[25]作为损失函数, 其表达式为

表 2 FTVGG16 卷积层网络参数

Tab.2 Layers parameters summarization of FTVGG16's

convolutional networks			
网络层	输出尺寸	参数	通道数
Conv1_1	224	3 × 3/1	64
Conv1_2	224	3 × 3/1	64
BN1	224	0	64
Pool1	112	2 × 2/2	64
Conv2_1	112	3 × 3/1	128
Conv2_2	112	3 × 3/1	128
BN2	112	0	128
Pool2	56	2 × 2/2	128
Conv3_1	56	3 × 3/1	256
Conv3_2	56	3 × 3/1	256
Conv3_3	56	3 × 3/1	256
BN3	56	0	256
Pool3	28	2 × 2/2	512
Conv4_1	28	3 × 3/1	512
Conv4_2	28	3 × 3/1	512
Conv4_3	28	3 × 3/1	512
BN4	28	0	512
Pool4	14	2 × 2/2	512
Conv5_1	14	3 × 3/1	512
Conv5_2	14	3 × 3/1	512
Conv5_3	14	3 × 3/1	512
BN5	14	0	512
Pool5	7	2 × 2/2	512

$$l = -\frac{1}{N} \sum_n^i \sum_k^j t_{nk} \lg y_{nk} \quad (5)$$

式中 l ——损失函数 N ——样本总数

t_{nk} ——标注的真实值第 n 个样本的第 k 个属性的值, 采用 one-hot 编码

y_{nk} ——卷积神经网络预测值第 n 个样本的第 k 个属性的值

防止模型过拟合, 在损失函数加上正则化项, 最终的损失函数表达式为

$$l \leftarrow l + \sum_i \sum_j \alpha \| \mathbf{w}_{i,j} \| + (1 - \alpha) \mathbf{w}_{i,j}^2 \quad (6)$$

式中 $\mathbf{w}_{i,j}$ ——第 i 层神经元与第 j 层神经元全连接权重

α ——调节因子 \leftarrow ——赋值操作

2.3 FTVGG16 网络模型的训练过程

FTVGG16 卷积神经网络采用 Adam^[26] (Adaptive moment estimation) 优化算法进行训练。Adam 利用梯度的一阶矩估计和二阶矩估计动态调整每个参数的学习率, 模型训练过程如图 5 所示。

系统训练过程步骤如下:

(1) 利用 Keras 框架中的生成器函数, 设置批尺寸 batch_size 为 32, 通过数据流的形式, 从训练样本中每次提取 32 个样本参与训练, 初始学习率

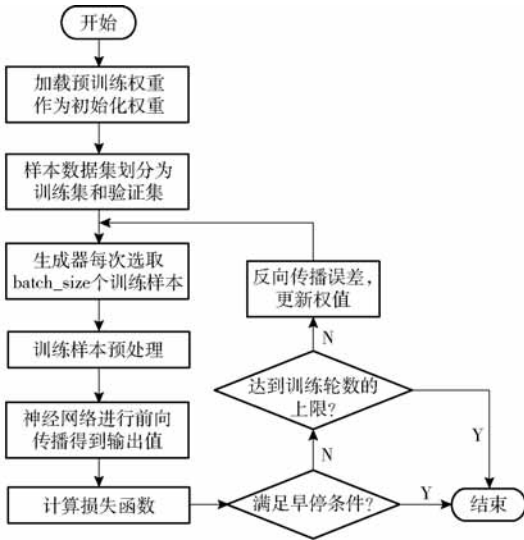


图5 FTVGG16 训练流程图

Fig. 5 Training flow chart of FTVGG16

learning_rate 设置为 0.000 01, 步长 step 设置为 0.001, 训练总轮数 epoch 设置为 50, 每次通过生成器从内存中取出 batch_size 个样本参与一次梯度下降的参数更新, 交叉熵损失函数 l 中的调节因子设置为 0.4, 模型训练参数值如表 3 所示。

表 3 FTVGG16 模型训练参数

Tab. 3 FTVGG16 model training parameters values

参数	数值
轮数	50
批尺寸	32
步长	0.001
学习率	0.000 01
指数衰减率 β_1, β_2	0.9, 0.999
除零误差项	1×10^{-8}

(2) 在训练集上进行训练, 并且在验证集上获取测试结果。模型训练使用了“早停”技术^[27]防止过拟合, 随着迭代轮数的增加, 如果发现验证误差上升, 则停止训练, 将停止之后的权重作为模型的最终参数。

(3) 神经网络通过前向传播得到预测值, 计算预测值和实际值误差, 如果误差满足系统所设阈值, 终止训练迭代; 否则继续进行训练, 直至达到系统设置训练轮数上限, 停止训练。

(4) 返回步骤(1), 循环训练。

3 测试结果与分析

3.1 FTVGG16 识别结果分析

FTVGG16 模型经过 24 轮训练, 在训练集上的准确率为 99.43%, 在验证集上的准确率为 95.89%, 模型训练到第 30 轮时, 训练集上的误差几乎降至零, 验证集上的误差也接近 0.2, 说明设计的

网络结构和优化算法可以使模型快速收敛, 从而达到较高的准确率和较低的误差。此外, 模型在验证集上的误差没有上升, 说明设计的 FTVGG16 模型没有出现过拟合, 能够较优进行图像识别, FTVGG16 模型在训练集和验证集上的准确率、误差如图 6 所示。FTVGG16 模型在训练集、验证集上的测试结果如表 4 所示。测试结果包括模型的准确率、误差、训练轮数和每轮训练时间。

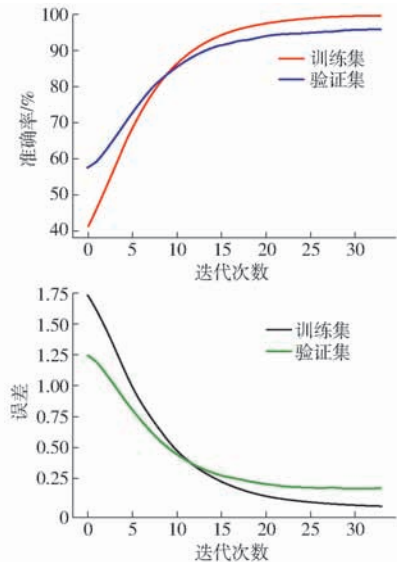


图6 FTVGG16 在训练集和验证集上训练的准确率和误差

Fig. 6 Accuracy and error of FTVGG16 on training set and validation set

表 4 FTVGG16 模型在训练集、验证集上的测试结果
Tab. 4 FTVGG16 model test results on training set and validation set

数据集	准确率/ %	误差	训练 轮数	每轮训练 时间/s
训练集	99.43	0.034 6	24	99.40
验证集	95.89	0.147 2	24	99.40

3.2 FTVGG16 与其他模型对比

为了验证 FTVGG16 模型的识别效果, 对比了其他卷积神经网络模型。对比方法均采用 5 折交叉验证法, 比较的预测结果为 5 次预测结果取平均。不同模型 (FTVGG16、VGG16、VGG19、ResNet50) 在训练集、验证集上的测试结果如表 5 所示。结果表明: FTVGG16 模型在训练集上的准确率为 99.43%, 误差为 0.034 6; 在验证集上的准确率为 95.89%, 误差为 0.147 2, 在训练集和验证集上的平均准确率为 97.66%, 明显高于其他模型的准确率。

3.3 FTVGG16 卷积神经网络可视化分析

3.3.1 FTVGG16 卷积神经网络特征图可视化

卷积神经网络中的卷积操作能够有效提取图像的特征^[28], 如图 7 所示。不同卷积层输出的特征由

表5 不同模型在训练集、验证集上的测试结果

Tab.5 Test results of different models on training set and validation set

数据集	模型	准确率/ %	误差	训练 轮数	每轮训练 时间/s
训练集	FTVGG16	99.43	0.034 6	24	99.40
	VGG16	97.45	0.075 3	24	58.40
	VGG19	98.10	0.061 4	24	58.80
	ResNet50	99.28	0.036 7	24	59.60
验证集	FTVGG16	95.89	0.147 2	24	99.40
	VGG16	95.03	0.231 7	24	58.40
	VGG19	94.61	0.234 7	24	58.40
	ResNet50	95.27	0.186 5	24	59.60

具体到抽象,Conv1_2和Conv2_2层的激活值输出仍保留大部分原始图像信息,靠后的Conv3_3、Conv4_3层输出更多地表达图像的线条、轮廓等抽

象特征,最后的Conv5_3仅仅保留了图像的高阶像素信息。通过可视化卷积层的输出,有助于理解卷积神经网络的工作原理,从而更好地改善模型性能。

3.3.2 FTVGG16卷积神经网络卷积核特征可视化

卷积神经网络中的卷积核可以提取图像不同层次的特征,如图8所示。Conv1_2层中的卷积核可以提取图像像素特征,Conv2_2层中的卷积核可以提取图像的边缘特征,这种边缘特征经过组合可以很好地描述图像的边、角、线条、轮廓特征,靠后的Conv3_3、Conv4_3层中的卷积核更多地提取图像的纹理特征,最后的Conv5_1、Conv5_2中的卷积核通过组合靠前的卷积核提取的特征,重构出图像的形状特征。通过可视化卷积核的输出,可以更好地研究卷积核如何从图像中提取各个层次的抽象特征,进而更好地理解卷积神经网络的工作原理。

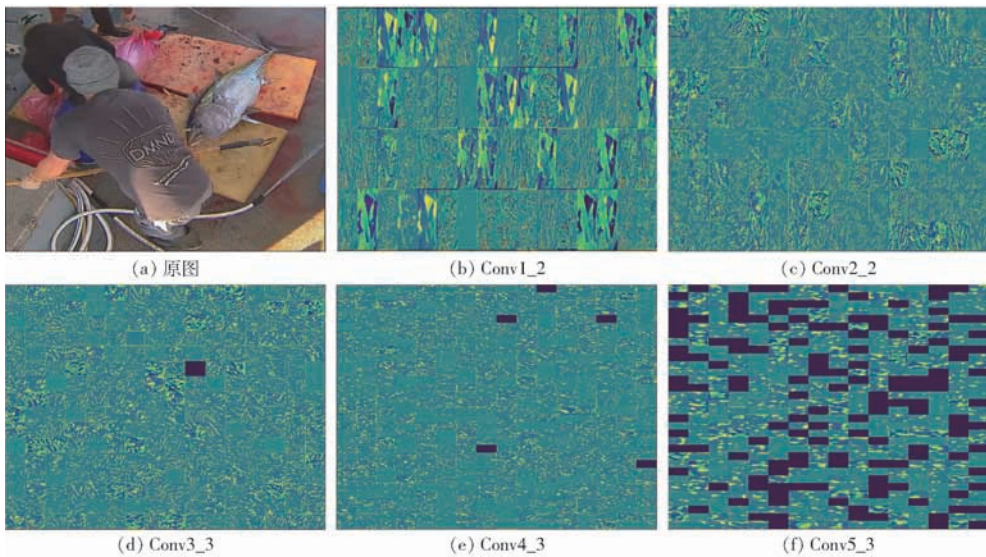


图7 FTVGG16部分卷积层输出可视化

Fig.7 FTVGG16 partial convolution layers output visualization

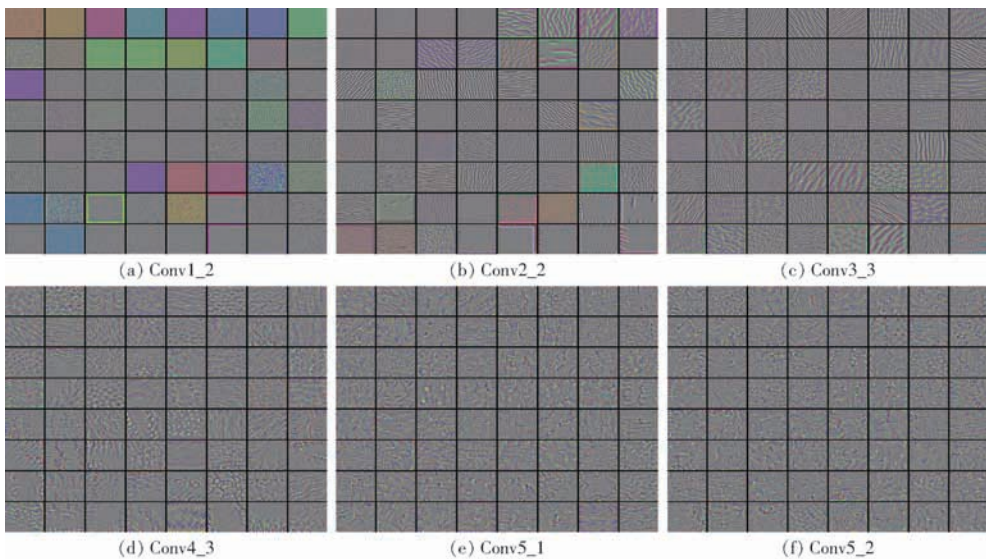


图8 FTVGG16部分卷积核输出可视化

Fig.8 FTVGG16 partial convolution kernels output visualization

3.4 FTVGG16 卷积神经网络在 Web 中的部署应用

本文使用 Python 语言编写的轻量级 Web 应用框架 Flask^[29],将训练的深度学习模型 FTVGG16 迁移到实际生产环境,Flask 集成了 WSGI(服务器网关接口),为 Python 应用程序或框架和 Web 服务器之间提供接口,同时 Flask 也集成了 Jinja2 模板引擎,能够快速搭建一个基于 Web 的应用。Fish identify 鱼类识别系统在部分图像上的测试结果如图 9 所示,Fish identify 的鱼类平均识别精度达到了 98% 以上,表明系统能够满足实际使用的要求。

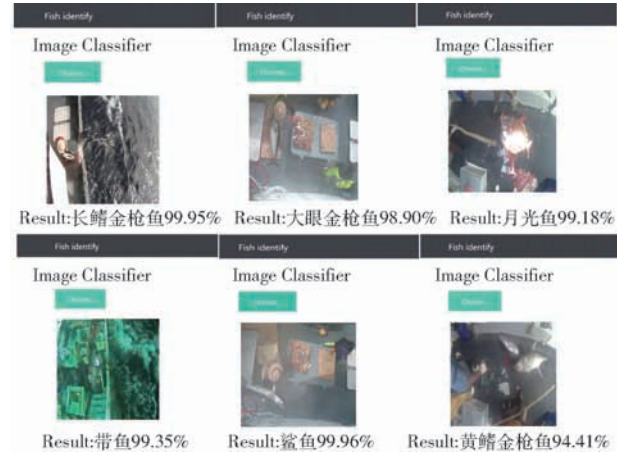


图 9 鱼类图像的识别示例

Fig. 9 Illustration of fish image recognition

4 结论

(1)从深度学习的角度,探讨了卷积神经网络在鱼类识别中的应用,通过分析已有的 VGG16 卷积神经网络模型良好的图像特征提取能力,设计了 FTVGG16 卷积神经网络。FTVGG16 卷积神经网络由卷积层、批规范层、池化层、Dropout 层、全连接层、softmax 层组成,采用带有约束的正则权重项作为模型的损失函数,并使用 Adam 优化算法对模型的参数进行更新。FTVGG16 网络在很大程度上能够克服训练的过拟合,收敛速度明显加快,训练时间明显减少,针对鱼类目标很小、背景干扰很强的图像,FTVGG16 模型平均准确率为 97.66%,对部分鱼的平均识别准确率达到 99.43%,模型具有较高的识别精度和鲁棒性。

(2)虽然 FTVGG16 模型在识别精度上达到了预期的效果,但是,实际运算过程中需要保存层与层之间的连接权重,会造成内存的负担,增大了计算成本,限制了系统响应。可以进一步从深度模型压缩加速方面进行改进,主要从模型参数修剪和共享、低秩分解、迁移/压缩卷积滤波器和知识精炼^[30]方面压缩模型,进而训练一个更加紧凑的神经网络,以再现大型网络的输出结果。

参 考 文 献

- [1] STRACHAN N J C, NESVADBA P, ALLEN A R. Fish species recognition by shape analysis of images [J]. Pattern Recognition, 1990, 23(5):539-544.
- [2] LARSEN R, OLAFSDOTTIR H, ERSBØLL B K. Shape and texture based classification of fish species [C] // Scandinavian Conference on Image Analysis, 2009, 5575:745-749.
- [3] HUANG P X, BOOM B J, FISHER R B. Underwater live fish recognition using a balance-guaranteed optimized tree [C] // Asian Conference on Computer Vision, 2013, 7724:422-433.
- [4] 吴一全, 殷骏, 戴一冕, 等. 基于蜂群优化多核支持向量机的淡水鱼种类识别 [J]. 农业工程学报, 2014, 30(16):312-319. WU Yiquan, YIN Jun, DAI Yimian, et al. Identification method of freshwater fish species using multi-kernel support vector machine with bee colony optimization [J]. Transactions of the CSAE, 2014, 30(16):312-319. (in Chinese)
- [5] 杜伟东, 李海森, 魏玉阔, 等. 基于 SVM 的决策融合鱼类识别方法 [J]. 哈尔滨工程大学学报, 2015, 36(5):623-627. DU Weidong, LI Haisen, WEI Yukuo, et al. Decision fusion fish identification using SVM and its experimental study [J]. Journal of Harbin Engineering University, 2015, 36(5):623-627. (in Chinese)
- [6] PAN S J, YANG Q. A survey on transfer learning [J]. IEEE Transactions on Knowledge & Data Engineering, 2010, 22(10):1345-1359.
- [7] DENG L, YU D. Deep learning: methods and applications [J]. Foundations & Trends in Signal Processing, 2014, 7(3):197-387.
- [8] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition [J]. Proceedings of the IEEE, 1998, 86(11):2278-2324.
- [9] AN D C, MEIER U, MASCI J, et al. Flexible, high performance convolutional neural networks for image classification [C] // Proceedings of the 22nd International Joint Conference on Artificial Intelligence, Barcelona, Catalonia, Spain, DBLP, 2011:1237-1242.
- [10] GEOFFREY H, LI Deng, DONG Yu, et al. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups [J]. IEEE Signal Processing Magazine, 2012, 29(6):82-97.
- [11] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [C] // International Conference on Neural Information Processing Systems. Curran Associates Inc, 2012, 12:1097-1105.
- [12] 刘德营, 王家亮, 林相泽, 等. 基于卷积神经网络的白背飞虱识别方法 [J/OL]. 农业机械学报, 2018, 49(5):51-56.

- LIU Deying, WANG Jialiang, LIN Xiangze, et al. Automatic identification method for *Sogatella furcifera* based on convolution neural network[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(5):51–56. http://www.jcsam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20180506&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2018.05.006. (in Chinese)
- [13] 傅隆生, 冯亚利, ELKAMIL T, 等. 基于卷积神经网络的田间多簇猕猴桃图像识别方法[J]. 农业工程学报, 2018, 34(2):205–211.
- FU Longsheng, FENG Yali, ELKAMIL T, et al. Image recognition method of multi-cluster kiwifruit in field based on convolutional neural networks[J]. Transactions of the CSAE, 2018, 34(2):205–211. (in Chinese)
- [14] 张力超, 马蓉, 张垚鑫. 改进的 LeNet-5 模型在苹果图像识别中的应用[J]. 计算机工程与设计, 2018, 39(11):3570–3575.
- ZHANG Lichao, MA Rong, ZHANG Yaoxin. Application of improved LeNet-5 model in apple image recognition[J]. Computer Engineering and Design, 2018, 39(11):3570–3575. (in Chinese)
- [15] 王璨, 武新慧, 李志伟. 基于卷积神经网络提取多尺度分层特征识别玉米杂草[J]. 农业工程学报, 2018, 34(5):144–151.
- WANG Can, WU Xinhui, LI Zhiwei. Recognition of maize and weed based on multi-scale hierarchical features extracted by convolutional neural network[J]. Transactions of the CSAE, 2018, 34(5):144–151. (in Chinese)
- [16] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]//Computer Science, 2014:1–14.
- [17] The nature conservancy fisheries monitoring [EB/OL]. <https://www.kaggle.com/c/the-nature-conservancy-fisheries-monitoring>.
- [18] BAKHODA A, YUAN G L, FUNG W W L, et al. Analyzing cuda workloads using a detailed GPU simulator[C]//IEEE Intl Symp Performance Analysis of Systems & Software, 2009:163–174.
- [19] ABADI M, AGARWAL A, BARHAM P, et al. TensorFlow: large-scale machine learning on heterogeneous systems[EB/OL]. <http://download.tensorflow.org/paper/whitepaper2015.pdf>.
- [20] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2015:770–778.
- [21] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift[C]//International Conference on Machine Learning. JMLR.org, 2015:448–456.
- [22] SRIVASTAVA N, HINTON G, KRIZHEVSKY A, et al. Dropout: a simple way to prevent neural networks from overfitting[J]. Journal of Machine Learning Research, 2014, 15(1):1929–1958.
- [23] HAHNLOSER R H, SARPESHKAR R, MAHOWALD M A, et al. Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit[J]. Nature, 2000, 405(6789):947–951.
- [24] LI M, ZHANG T, CHEN Y, et al. Efficient mini-batch training for stochastic optimization[C]//ACM Sigkdd International Conference on Knowledge Discovery & Data Mining. ACM, 2014:661–670.
- [25] WONG W S, QIN A. Method and apparatus for establishing topic word classes based on an entropy cost function to retrieve documents represented by the topic words: US6128613[P]. 2000–10–03.
- [26] KINGMA D P, BA J. Adam: a method for stochastic optimization[C]//The Third International Conference for Learning Representations, San Diego, 2014:1–15.
- [27] GENÇAY R, QI M. Pricing and hedging derivative securities with neural networks: Bayesian regularization, early stopping, and bagging[J]. IEEE Trans. Neural Networks, 2001, 12(4):726–734.
- [28] ZEILER M D, FERGUS R. Visualizing and understanding convolutional networks[C]//European Conference on Computer Vision, 2014, 8689:818–833.
- [29] Flask[EB/OL]. <http://flask.pocoo.org/>.
- [30] CHENG Y, WANG D, ZHOU P, et al. Model compression and acceleration for deep neural networks: the principles, progress, and challenges[J]. IEEE Signal Processing Magazine, 2018, 35(1):126–136.