

# 基于卷积神经网络的食物图像识别

王 博<sup>1,2</sup>, 刘俊康<sup>3</sup>, 陆逢贵<sup>1</sup>, 刘登勇<sup>1,4\*</sup>, 曹振霞<sup>1</sup>

(1. 渤海大学食品科学与工程学院, 生鲜农产品贮藏加工及安全控制技术国家地方联合工程研究中心, 锦州 121013; 2. 渤海大学化学化工学院, 锦州 121013; 3. 大连民族大学信息与通信工程学院, 通辽 116000; 4. 江苏省肉类生产与加工质量安全控制协同创新中心, 南京 210095)

**摘要:** **目的** 探究一种基于 Inception\_V3-CNN 模型的食物图像识别和分类方法。**方法** 选取包含 20 类食品和 19609 张的食物图像建立数据集 Food-101, 通过卷积神经网络(convolutional neural networks, CNN)提取图像候选区域的视觉特征, 并自动进行分类, 使其具有较高的识别率; 此外, 采集来自 kaggle 库中已标注的食物图像集做预测对比实验, 以保证 Inception\_V3-CNN 模型检测的准确度。**结果** 该方法能够去除背景噪音, 且仅需部分提取视觉特征, 可以有效解决网络食物图像的分类问题, 与多视图支持向量机(support vector machine, SVM)+方向梯度直方图(histogram of oriented gradients, HOG)和传统 CNN 方法相比, 在测试时间相近、设备计算能力相同的条件下, 该方法识别率更高, 迭代次数为 15000 次时, Loss 值降至 4.92, 准确率可达 93.89%。**结论** 此方法可以快速识别食物图像, 在实际网络图片中能有较好的可移植性。将算法移植到移动设备中实现 APP 的模块化操作也将成为后续工作探索的重点方向。

**关键词:** 食物图像; 图像检索; 图像分类; Inception\_V3-CNN; 卷积神经网络

## Application of convolutional neural network in image retrieval and classification of food image

WANG Bo<sup>1,2</sup>, LIU Jun-Kang<sup>3</sup>, LU Feng-Gui<sup>1</sup>, LIU Deng-Yong<sup>1,4\*</sup>, CAO Zhen-Xia<sup>1</sup>

(1. College of Food Science and Technology, Bohai University, Food Safety Key Lab of Liaoning Province, National and Local Joint Engineering Research Centre of Storage, Processing and Safety Control Technology for Fresh Agricultural and Aquatic Products, Jinzhou 121013, China; 2. College of Chemistry and Chemical Engineering, Bohai University, Jinzhou 121013, China; 3. College of Information and Communication Engineering, Dalian University for Nationalities, Dalian 116000, China; 4. National Center of Meat Quality and Safety Control, College of Food Science and Technology, Nanjing Agricultural University, Nanjing 210095, China)

**ABSTRACT: Objective** To investigate a food image recognition and classification method based on Inception\_V3-CNN model. **Methods** The food image dataset Food-101 containing 20 foods and 19609 food images was established, and the visual features of the candidate regions of the image were extracted by convolutional neural network (CNN), and automatically classified to have a higher recognition rate. The food image set from the kaggle library was collected for predictive comparison experiments to ensure the accuracy of the Inception\_V3-CNN model. **Results** This method could remove background noise and only needed to extract visual features in part, which could

基金项目: 辽宁省高等学校产业技术研究院重大应用研究项目(041804)、辽宁省重点研发计划指导计划项目(2017205003)

Fund: Supported by Major Applied Research Projects of Liaoning Institute of Industrial Technology (041804), and Key R&D Program of Liaoning Province (2017205003)

\*通讯作者: 刘登勇, 博士, 教授, 主要研究方向为肉品加工与质量安全控制。E-mail: jz\_dyliu@126.com

\*Corresponding author: LIU Deng-Yong, Ph.D, Professor, Bohai University, No.19, Keji Road, New Songshan District, Jinzhou 121013, China. E-mail: jz\_dyliu@126.com

effectively solve the classification problem of network food images. Compared with support vector machine+ histogram of oriented gradients (SVM+HOG) and traditional CNN methods, under the conditions of similar test time and the same computing power of equipment, the recognition rate of this method was higher. When the number of iterations was 15000 times, the Loss value dropped to 4.92, and the accuracy rate could reach 93.89%. **Conclusion** This method can quickly recognize food images and has good portability in real network images. Transplantation of the algorithm to mobile devices to realize the modular operation of APP will also become the focus of future work.

**KEY WORDS:** food image; image retrieval; image classification; inception\_V3-CNN; convolutional neural network

## 1 引言

食品行业的迅速兴起离不开电子商务的快速发展, 电子商务在实践中表现出来的时空优势使人们看到了它在食品领域中的无限潜力。在大数据时代, 运用图像分类识别方法可以有效的检索食品图像, 帮助消费者了解其营养成分, 从而定制个性化服务, 同时也可以帮助食品监管部门追溯食品信息、自动分拣过期食品, 从而加大食品安全的监察力度<sup>[1,2]</sup>。鉴于食品图像识别具有广泛的潜在应用价值, 但存在泛化性不强和识别精度较差等问题<sup>[3]</sup>, 因此需要寻找适合面向食品领域的图像识别和分类算法。

在食品图像识别领域, 基于传统图像识别服务器和复杂外部设备的识别方法较为常见, 而最新的研究方法是通过智能手机实现在线实时识别图像<sup>[4]</sup>, 并围绕其开展不同算法的建模优化, 如支持向量机方法, 可基于云计算或者使用多视图支持向量机(support vector machine, SVM)从食品的原料和类别进行学习分类, 也可结合神经网络特性与监督机器学习等算法对不同食品图像的数据集进行分类, 但该方法要求识别目标具有完整性和形状分割的准确性, 而阴影可能会导致目标分割缺乏完整性, 最终影响分类算法的性能<sup>[5-9]</sup>。随着机器视觉技术的发展, 多层卷积神经网络(convolutional neural networks, CNN)法已经成为提取图像视觉特征的主流方法, 与其他机器学习方法相比, 基于深度学习 CNN 框架(网络训练、调整使用预处理模型和融合滤波器), 可以确保输出结果达到更高的精度。Tahir 等<sup>[10]</sup>利用 CNN 模型对 5 个不同种类的真菌进行分类, 准确率达 94.8%, 效果较好; Philipsen 等<sup>[11]</sup>利用模型对不同种类家禽的不同部位内脏进行识别, 肌肉分类准确率高达 91.58%。但是传统 CNN 法并不适合大规模和高像素细粒度的图像识别任务, 算法模型的进步会优化分类过程, 提高准确率。Paritosh 等<sup>[12]</sup>对传统 CNN 模型、残留神经网络模型(residual neural network, ResNet)模型和 Inception\_V3-CNN 进行对比研究, 结果表明 Inception\_V3-CNN 准确率最高, 对某种食品识别准确率最高达 97.60%。Inception\_V3-CNN 是谷歌公司 2014 年提出的 CNN 分类模型<sup>[13]</sup>, 与传统 CNN 模型相比, 在保证分类网络分类准确率的同时降低了模型对内容的消耗, 从而提取的视觉特征

更具有判别性, 在食品图像检索和分类任务中, 能够有效提升识别效率。鉴于此, 本文采用 Inception\_V3-CNN 模型进行实验, 并将全面评估其实验结果与 SVM+方向梯度直方图(histogram of oriented gradients, HOG)和 CNN 方法比较, 探究其食品图像检索能力和分类性能, 以期拓展食品领域的图像识别研究奠定基础。

## 2 食品图像识别

### 2.1 卷积计算

卷积神经网络图像在分类识别过程中, 卷积核会按照步长所设定的大小遍历特征图像中的像素, 并进行点积运算<sup>[14]</sup>。其中卷积核(算子)是用来做图像处理时的矩阵, 在图像处理时也被称为掩膜, 作为参与原图像做运算的一个参数。卷积计算的过程如公式 1 所示, 演示过程如图 1 所示, 当使用模板和图像进行卷积时, 对于图像上的某一个点, 首先通过过滤器将模板的原点和该点重合, 其次将模板上的点和图像上对应的点相乘, 最后各点的积相加, 得到的结果即为该点的卷积值。

$$g = f \times h$$

$$g(i,j) = \sum_{k,l} f(i+k,j+l)h(k,l) \quad \text{式(1)}$$

式中,  $h$ --为相关核(Kernel);

$k, l$ --卷积核的大小;

$i, j$ --是图像大小。

### 2.2 特征提取及传递

基于 Inception\_V3-CNN 模型的食品图像检索和分类方法可以更好地检测并提取原数据中的二维特征, 如图 2 所示。该方法主要分为 3 个部分: 首先微调(fine-tune), 采用经过训练的 Inception\_V3-CNN 模型检测食品图像的食品区域; 然后基于检测的食品区域, 利用 CNN 深度神经网络提取该候选区域的视觉特征<sup>[15]</sup>; 最后将提取的食物图像视觉特征应用到食品图像检索和分类任务。

### 2.3 网络训练

网络训练方式为监督分类, 该训练过程主要分为 2 个阶段。

第一阶段: 向前传播阶段。输入的样本均为灰度图时, 即样本图片没有红黄蓝 3 个通道, 输入层  $X$  即为一个二维

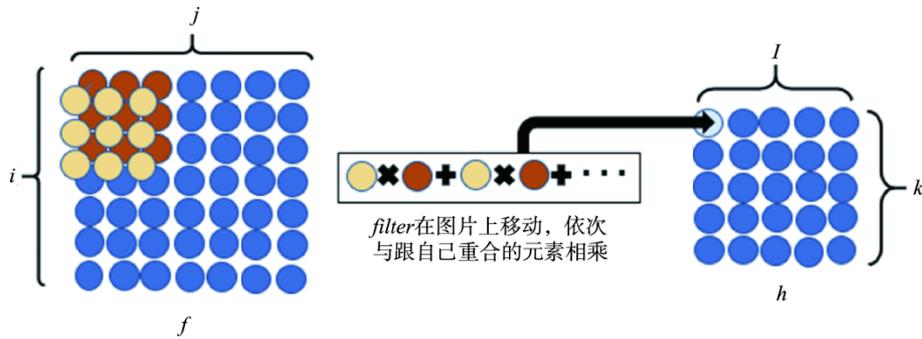


图 1 算法框架概述

Fig.1 Overview of the algorithm framework

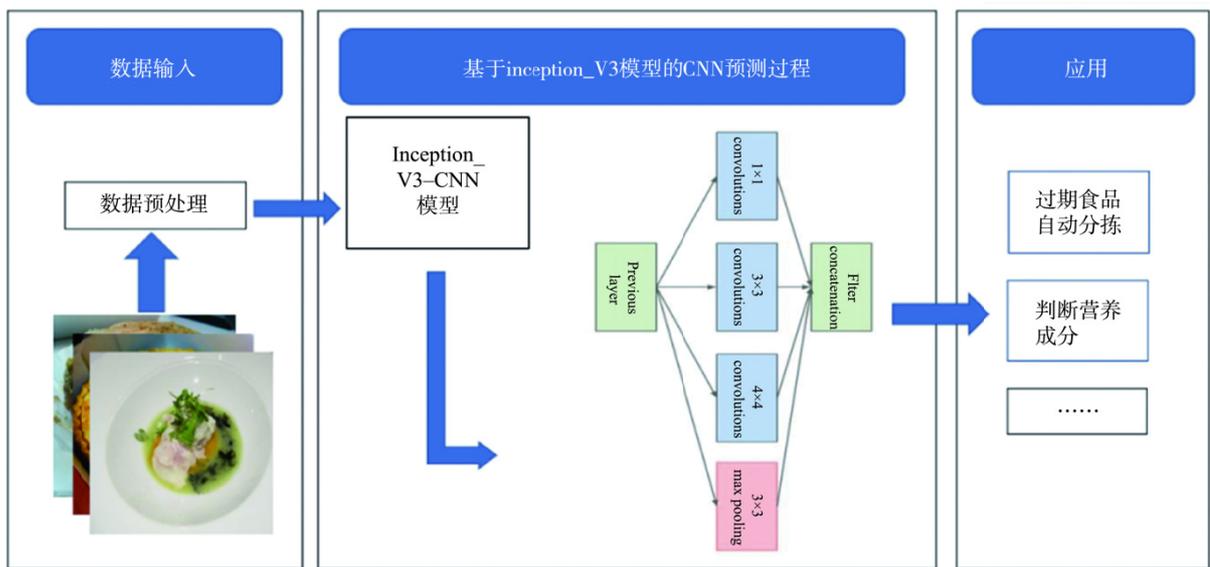


图 2 基于 Inception\_V3-CNN 的食品图像检索和分类方法

Fig.2 Food image retrieval and classification method based on Inception\_V3-CNN

矩阵, 矩阵的值与图片的各个像素位置的值相等, 和卷积层相连的卷积核  $W$  也是矩阵; 样本是有 RGB 的彩色图片时, 输入  $X$  即为 3 个矩阵, 即分别对应 R、G 和 B 的 3 个矩阵(张量), 此时与卷积层相连的卷积核  $W$  也是张量, 对应的最后一维的维度为 3, 即每个卷积核都是由 3 个子矩阵组成。以此类推, 对于 3D 的彩色图片样本, 输入的变量  $X$  可以是 4 维、5 维的张量, 那么对应的卷积核  $W$  也应是高维的张量。无论其维度达到多高, 图片训练的向前传播过程都可以用公式 2 表示。

$$a^2 = \partial(z^2) = \partial(a^1 * w^2 + b^2) \tag{2}$$

式中,  $z$  是经历了激活函数的  $a$ ;  $w$  是前层卷积结果;

$b$ -偏移的大小;  $\sigma$ -激活函数, 一般用 ReLU。

第二阶段: 向后传播阶段。将矩阵问题转变为最优化判断问题, 通过不断迭代地使用梯度下降算法找到使代价函数最小的权重值和偏置值。学习算法的性能表现由学习率决定, 而衡量神经网络预测结果与正确答案之间的误差距离则由损失函数表示。损失函数(loss function, loss 值)是

用来估量你模型的预测值与真实值  $Y$  的不一致程度表示, 损失率越低则准确率越高。基础损失函数分为 2 种, 分别是用于解决分类任务的交叉熵损失函数(cross entropy)和用于解决回归任务的均方误差损失函数(mean-square error, MSE)<sup>[16]</sup>。具体过程如公式 3 所示。

$$(w, b)^2 = \frac{1}{2n} \sum_x \|y(x) - a\|^2 \tag{3}$$

式中,  $w$ -神经元权重集合;  $b$ -偏置量的集合;  $n$ -训练集中的样本量;

$a$ -输入为  $x$  时, 神经网络的输出向量, 计算与真实标签之间的差距;

$\|y(x) - a\|^2$ -向量的模, 也称为二次代价函数或者均方误差(MSE)。

### 3 结果及分析

#### 3.1 实验数据

实验在 Python3.5(64bit)平台上完成, 具体测试开发

环境如表 1 所示。所采用的图像由 Kaggle 平台提供(图像分辨率统一格式化为  $384 \times 384 \times 3$ (RGB, uint8)), 该平台是开发商和数据科学家用于举办机器学习竞赛、托管数据库、编写和分享代码的功能性平台, 数据一般为企业或政府为解决面临问题而提供, 可信度和质量较高。本研究利用该平台分享的图像数据集来验证基于卷积神经网络方法识别食品图像的有效性(数据集地址为 [https://www.vision.ee.ethz.ch/datasets\\_extra/food-101/](https://www.vision.ee.ethz.ch/datasets_extra/food-101/))。原始采集得到的食品数据集包含 101000 张图像和 101 种食品类别, 从中抽取 20 种食品图片, 并用脚本统一成  $256 \times 256$  的分辨率, 在此过程中脚本自动删除部分比例严重失调的图片, 最终形成包含 20 种食品类别和 19609 张图片的新数据集, 将其定义为数据集 Food-101。实验数据集如表 2 所示, 图 3 展示了 Food-101 数据集中的部分实验样本。

表 1 测试环境  
Table 1 Testing environment

硬件		框架
Ideapad Y700-151SK 笔记本		Tensorflow-1.8.0.0
Win10 家庭中文版系统		CUDA 9.0
CPU	Intel(R)I7- 6700HQ 处理器	CUDNN 7.4.2
GPU	NVIDIA_GTX 960M - 4G	

表 2 Food-101-01 数据集示例  
Table 2 Food-101-01 dataset sample

	训 练 集	测 试 集		训 练 集	测 试 集
苹果派 (apple_pie)	792	198	意式特色面包 (bruschetta)	779	195
小排骨 (baby_back_ribs)	784	196	凯萨沙拉 (caesar_salad)	781	195
巴克拉瓦 (baklava)	780	195	奶油甜馅煎饼卷 (cannoli)	786	196
生牛肉片 (beef_carpaccio)	786	196	卡普列塞沙拉 (caprese_salad)	784	196
牛肉塔塔 (beef_tartare)	786	197	胡萝卜蛋糕 (carrot_cake)	788	197
甜菜头沙拉 (beet_salad)	789	197	酸橘汁腌鱼 (ceviche)	780	195
面包圈 (beignets)	784	196	奶酪 (cheese_plate)	786	196
石锅拌饭 (bibimbap)	780	195	奶酪蛋糕 (cheesecake)	787	196
面包布丁 (bread_pudding)	788	197	咖喱鸡 (chicken_curry)	785	197
墨西哥卷 (breakfast_burrito)	779	195	墨西哥烤鸡煎饼 (chicken_quesadilla)	784	196
合计	7848	1962	合计	7840	1959



图 3 Food-101 数据集部分食品图像  
Fig.3 Food-101 dataset partial food images

### 3.2 图像预处理结果

利用训练集的图像对深度学习模型进行训练之前, 首先需要利用预训练模型 Inception\_V3.ckpt 对 CNN 模型进行参数设置, 且为了适应数据集, 数据恢复过程中选择不恢复连接层 Inception\_V3/Logits 和 InceptionV3/AuxLogits 的默认参数, 而使用随机参数[最大执行步数(max-batch)为 50000、每步使用的 batch 数量为 32、学习率为 0.001、设置固定学习率 fixed、优化器使用 rmsprop、模型中所有参数的二次正则化超参数为 0.00004]; 其次需要对数据集进行预处理以扩充数据量、预防数据过拟合现象和提高深度学习模型的鲁棒性, 在计算机视觉方面, 其主要方法是随机裁剪、水平翻转和亮度调节。由于食品图片的特殊性, 在保证标签不变的情况下, 可以任意对食品图片进行随机裁剪和水平翻转。图 4 展示了 bread\_pudding(面包布丁)预处理后的各层特征图。

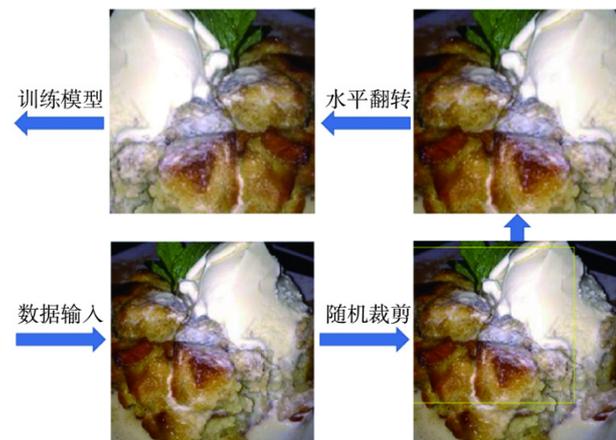


图 4 影像预处理  
Fig.4 Pretreatment of image

### 3.3 图像特征可视化结果

食品图片经过网络预处理后, 可以有效排除背景噪音的干扰, 使得提取到的视觉特征更具有明显的判别性。具体处理过程如下: 将 Food-101 图像集划分成 2 个部分, 80%作为训练集, 20%作为测试集。首先将彩色图像转化为灰度图像, 进行灰度化处理; 其次借助于算法的强大计算

能力对数据集的食品图像各区域分别进行检测, 提取各层的不同特征; 最后, 将数据输入模型, 模型会自动基于最合理的特征将图片按类别进行匹配, 匹配过程所需图片的特征筛选并不需要人为的操作, 只需设定识别标签即可。图 5 是对所有的图像进行灰度化、区域检测和视觉二维特征提取过程的简单描述, 由图 5 可知经过 Inception\_V3-CNN 算法提取得到的图像特征更加细微化。

最后将细微化的特征应用到食品图像检索和分类任务, 且模型每层卷积采用 Max\_pooling 算法避免过拟合现象, 网络深度设为 50 Layer<sup>[17]</sup>, 该深度下模型部分层级(1层)的结构如图 6 所示。其中 input1 为输入层, Conv2d\_1 为卷积层一层, Max\_pooling\_1 为池化层一层。

### 3.4 识别结果评估及分析

Loss 值对模型精确度影响最大, 降低该数值对提高精确度具有积极作用, 即 Loss 值越小证明模型精确度越高, 研究的价值也就越大<sup>[18]</sup>, 因此在其他参数恒定的前提下, 追求较低 Loss 值是保证模型精确度的重要方式。本文利用 Inception\_V3-CNN 模型在数据集上迭代训练至 45000 次, 得到的 Loss 趋势如图 7 所示, 其中, X 轴为迭代次数, Y 轴为 Loss 值, 在神经网络的训练过程中, Loss 值整体呈现下降的趋势, 说明算法选定的各个神经元的权重 W 和偏移量 B、filter 的大小等参数, 在不断的拟合样本集, 并使全局的损失函数减小, 收敛效果较好。Loss 趋势基本经历 3 个阶

段, 0~15000 次为缓慢递减阶段, 15000~35000 次为平稳变化阶段, 35000~47000 次为波动变化阶段。在第一阶段, Loss 值最低降至 4.92, 在第二阶段, Loss 最低值回升至 4.93, 在第三阶段, 经过 45000 次迭代后, Loss 值由 4.92 缓慢下降至 4.81 后又缓慢上升, 且相比较 15000 次, 损失值仅降低了 2.36%, 说明 15000 次和 45000 次的迭代训练差别对 Loss 值影响较小。虽然可以继续对模型进行训练, 以进一步提升模型的精度和加强深度神经网络的鲁棒性, 但迭代训练次数增加也会消耗大量的资源, 仅通过增加迭代次数提高性能的意义较低, 应利用扩充数据集的方式提升数据多样性。基于上述原因, 文章将迭代次数定为 15000 次。

如图 8 所示, 分别使用 SVM+HOG(方向梯度直方图)、传统 CNN(卷积层+池化层)和 Inception\_V3-CNN 3 种算法对 Food-101 数据集进行图像分类识别实验, 其中, X 轴为迭代次数, Y 轴为准确度。由图 8 可知迭代次数对 3 种算法的提升效果不同, 综合考虑运算速度和参数修正速度, 将迭代次数设定为 10000、15000、20000、30000、40000 和 45000 次。传统 CNN 和 Inception\_V3-CNN 算法随着迭代次数的提高, 准确率分别从 86.72%和 92.13%提高至 89.81%和 94.81%, 分别提升了 3.06%和 2.64%, 提升效果较好, 说明迭代次数越大越适合这两种算法, 但迭代 15000 次与 45000 次相比准确率仅分别提高 0.41%和

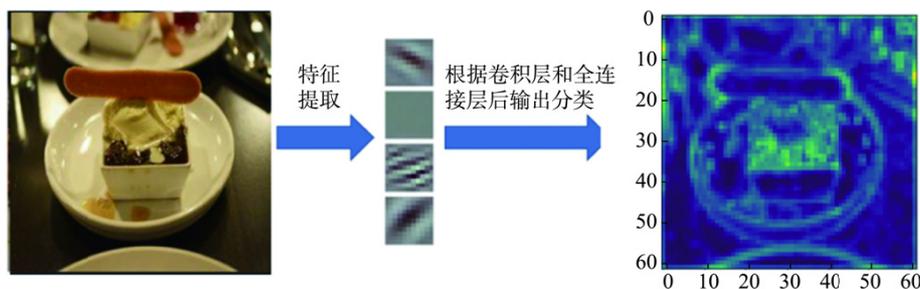


图 5 基于 Inception\_V3-CNN 进行特征提取  
Fig.5 Feature extraction based on Inception\_V3-CNN

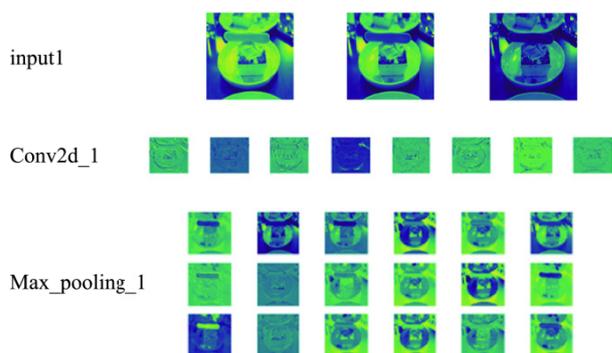


图 6 卷积神经网络的各层可视化图  
Fig.6 Visualization of different layers of convolutional neural networks

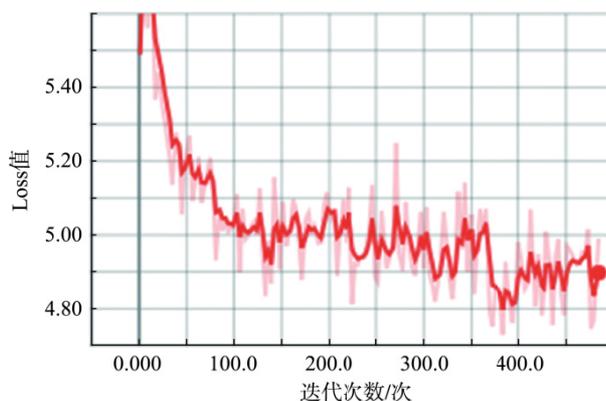


图 7 Inception\_V3-CNN 模型在 Food-101 上的训练趋势  
Fig.7 Training trend of Inception\_V3-CNN model on Food-101

0.88%，差别较小，因此 15000 次更为适合这 2 种算法。SVM+HOG 算法迭代对于次数较小时收敛效果较差，只有大于 10000 步，经过较大的 epoch 才能收敛，但 10000 次与 45000 次相比准确率仅提高 0.02%，对电脑负荷较高，因此迭代次数为 10000 次时较适合 SVM+HOG 算法。综上所述，在识别性能方面，3 种算法中 Inception\_V3-CNN 的准确度最高，原因在于 Inception\_V3-CNN 通过优化算法减少了人工抽取图像特征的步骤，选取了合适的卷积核，因此相较于普通的卷积神经网络模型，在相同情况下精度的提升具有一定的优势。传统 CNN 算法和 Inception\_V3-CNN 算法在识别准确度上都优于 SVM+HOG 算法，说明卷积神经网络可以有效减少食品图像背景信息产生的干扰，进而提高其分类识别性能。

图 9 展示了 max-Batch 为 40000 次时，基于 SVM+HOG、CNN 和 Inception\_V3-CNN 3 种算法对 10 类食品的识别结果，由图可知在所有样本中 CNN 和 Inception\_V3-CNN 算法的实验效果均优于 SVM+HOG 算法，准确率平均提高了 23.00%和 31.90%。而优化后的 Inception\_V3-CNN 算法较传统 CNN 识别大多数食品过程中的性能表现稍占优势，且进一步验证了其有效性。

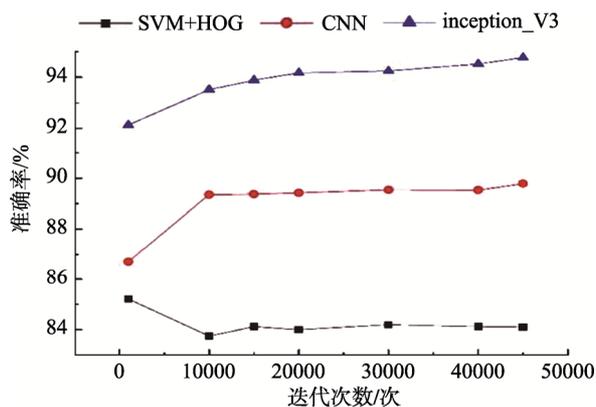


图 8 不同方法的识别性能比较

Fig.8 Retrieval performance comparison between the proposed and other methods

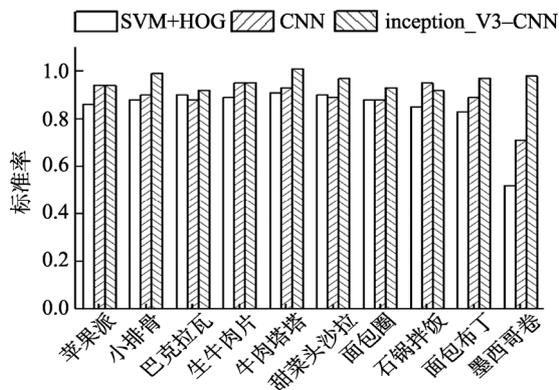


图 9 10 类食品在不同方法下的识别结果

Fig.9 Classification performance of different methods on randomly selected 10 food categories

## 4 结论与讨论

本研究将卷积神经网络引入到食品图像的识别领域中，基于 Inception\_V3 -CNN 模型在采集得到的 Food-101 图像集进行实验。结果如下：

(1) 利用 Kaggle 平台分享的 20 种食品类别和 19609 张图片数据集，来验证 Inception\_V3 -CNN 模型方法识别食品图像的有效性，其中训练集和预测集样本数量比为 8:2，分别是 7840 张和 1962 张。在识别分类之前首先需对模型进行参数设置，max-batch 为 50000、batch 为 32、学习率为 0.00。二次正则化超参数为 0.00004；其次对图片进行随机裁剪、水平翻转和亮度调节的网络与处理；最后将网络深度设为 50 层进行特征提取。

(2) 当迭代次数设定为 10000、15000、20000、30000、40000 和 45000 次时比较 Inception\_V3-CNN 与传统 CNN 和 SVM+HOG 3 种方法识别性能，考虑到 Loss 值的变化规律和提高识别效率的前提，本研究将 3 种方法迭代次数分别设定为 15000 次、15000 次和 10000 次，在此情况下准确率分别为 93.89%、89.37%和 84.12%，Inception\_V3-CNN 算法优势明显。

(3) 为了进一步验证结果的准确性，比较上述 3 种算法对 10 类食品的识别结果，95.80%、89.20%和 84.20%，Inception\_V3-CNN 算法表现优异。

(4) 本研究提出的基于 Inception\_V3 -CNN 模型的快速识别方法，与传统 CNN 和 SVM+HOG 方法相比，可以更准确定位到各类食品的图片特征，减少人工抽取图像特征的步骤，并且模型可以根据样本实际情况自行选取合适的卷积核大小，从而适应已被标记的不同食品图片类别。该算法下 15000 次迭代 Loss 值降至 4.92，准确率达到 93.89%，说明此方法可以快速识别食品图像，在实际网络图片中能较好的可移植性。此外，将算法移植到移动设备中实现 APP 的模块化操作也将成为后续工作探索的重点方向。

## 参考文献

- [1] 安进, 张丹. 基于 RFID 和图像识别的食品追溯信息查询算法[J]. 科技通报, 2017, 33(5): 134-137.  
An J, Zhang D. Food traceability information query algorithm based on RFID and image recognition [J]. Bull Sci Technol, 2017, 33(5): 134-137.
- [2] 顾理琴. 基于图像识别技术的食品种类检测方法[J]. 食品研究与开发, 2017, 38(2): 160-175.  
Gu LQ. Image recognition technology of the food species detection method [J]. Food Res Dev, 2017, 38(2): 160-175.
- [3] Yang SL, Chen M, Pomerleau D, et al. Food recognition using statistics of pairwise local features [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2010: 2249-2256.
- [4] Kawano, Yoshiyuki, Yanai, et al. FoodCam: A real-time food recognition system on a smartphone [J]. Multimed Tools Appl, 2015, 74(14): 5263-5287.

- [5] Kawano Y, Yanai K. Real-time mobile food recognition system [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2013: 1–7.
- [6] McAllister P, Zheng H, Bond R, *et al.* Combining deep residual neural network features with supervised machine learning algorithms to classify diverse food image datasets [J]. *Comput Biol Med*, 2018, (95): 217–233.
- [7] Tatsuma A, Aono M. Food image recognition using covariance of convolutional layer feature maps [J]. *Ictice Trans Inf System*, 2016, (6): 1711–1715.
- [8] He HS, Kong FY, Tan JD. DietCam: Multiview food recognition using a multikernel SVM [J]. *IEEE J Biomed Health Inf*, 2015, 20(3): 848–855.
- [9] Pouladzadeh P, Shirmohammadi S, Bakirov A, *et al.* Cloud-based SVM for food categorization [J]. *Multimed Tools Appl*, 2015, 74(14): 5243–5260.
- [10] Tahir MW, Zaidi NA, Rao AA, *et al.* A fungus spores dataset and a convolutional neural network based approach for fungus detection [J]. *IEEE Trans Nanobiosci*, 2018, 17(3): 281–290.
- [11] Philipsen MP, Ducholm JV, Jorgensen A, *et al.* Organ segmentation in poultry viscera using RGB-D [J]. *Sensors*, 2018, 18(1): 1–15.
- [12] Pandey P, Deepthi A, Mandal B, *et al.* FoodNet: Recognizing foods using ensemble of deep networks [J]. *IEEE Singal Process Lett*, 2017, 24(12): 1758–1762.
- [13] Szegedy C, Vanhoucke V, Ioffe S, *et al.* Rethinking the inception architecture for computer vision [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 2818–2826.
- [14] Yu D, Wang H, Chen P, *et al.* Mixed pooling for convolutional neural networks [M]. *Rough Sets and Knowledge Technology*. Springer International Publishing, 2014: 364–375.
- [15] Min WQ, Bao BK, Mei SH, *et al.* You are what you eat: Exploring rich recipe information for cross-region food analysis [C]. *IEEE Transactions on Multimedia*, 2017: 1–15.
- [16] Meyers A, Johnston N, Rathod V, *et al.* Im2Calories: Towards an automated mobile vision food diar [C]. *IEEE International Conference on Computer Vision*, 2015: 1233–1241.
- [17] He K, Zhang X, Ren S, *et al.* Deep residual learning for image recognition [C]. *Proc of the 29th IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 770–778.
- [18] Ramcharan A, Baranowski K, McCloskey P, *et al.* Deep learning for image-based cassava disease detection [J]. *Front Plant Sci*, 2017, (8): 1–7.

(责任编辑: 武英华)

### 作者简介



王 博, 讲师, 主要研究方向为食品大数据分析。

E-mail: daqingwb@163.com



刘登勇, 教授, 主要研究方向为肉品加工与质量安全控制。

E-mail: jz\_dyliu@126.com