

文章编号: 2095-2163(2020)06-0055-04

中图分类号: TP391.4

文献标志码: A

基于卷积神经网络的中餐菜品名称成分识别

边 竞¹, 王艺璇¹, 代宇晖¹, 陈泽众², 黄靖淳³

(1 河北工业大学 理学院, 天津 300000; 2 河北工业大学 人工智能与数据科学学院, 天津 300000;

3 河北工业大学 经济管理学院, 天津 300000)

摘要: 针对食物图像识别领域中菜品图像识别精度低的问题, 提出了融合迁移学习与 Inception-V3 模型的中餐菜品名称及原料识别方法, 实现针对中餐菜品名称及原料的识别工作。该方法在预训练模型的基础上对瓶颈层进行训练, 并添加新的全连接层, 完成对包含 172 类食物的中餐菜品数据库 VIREO Food-172 的菜品图像标签分类工作。实验表明: 在迭代次数为 50 000 次时, 菜品名称识别准确率达到 70.85%, 原料识别率达到 56.26%, 同时对该模型进行轻量化压缩, 借助移动端架构设计一款中餐菜品识别 APP 是后续的研究方向。

关键词: 菜品识别; 图像标签分类; 迁移学习; Inception-V3; 卷积神经网络

Recognition of ingredients and dish names based on convolutional neural network

BIAN Jing¹, WANG Yixuan¹, DAI Yuhui¹, CHEN Zezhong², HUANG Jingchun³

(1 School of Science, Hebei University of Technology, Tianjin 300000, China;

2 School of Artificial Intelligence, Hebei University of Technology, Tianjin 300000, China;

3 School of Economics and Management, Hebei University of Technology, Tianjin 300000, China)

[Abstract] Aiming at the problem of difficult recognition of Chinese food dishes in the field of food image recognition, a method for recognizing Chinese food dish names and raw materials combining transfer learning and Inception-V3 model is proposed to realize the recognition of Chinese food dish names and raw materials. This method trains the bottleneck layer on the basis of the pre-trained model and adds a new fully connected layer to complete the classification of food image labels for the Chinese food dish database VIREO Food-172 containing 172 kinds of food. Experiments show that when the number of generations is 50000, the recognition rate of dish name recognition reaches 70.85%, and the recognition rate of raw materials reaches 56.26%. At the same time, the model is lightly compressed, and the design of a Chinese dish recognition APP with the help of mobile terminal architecture is the follow-up Research direction.

[Key words] dish recognition; image label classification; migration learning; inception-V3; convolutional neural network

0 引言

食物是供给机体活动所需能量的重要物质来源, 与人类日常生活息息相关。由于食物种类繁多, 所含营养成分也千差万别, 在普通条件下无法快速测定食物的种类及营养成分, 但随着深度学习的火热, 计算机视觉技术可以完成对食物图像的识别任务, 因此, 如何搭建一个泛化性较强、识别精度较高的食物图像识别分类算法是当前热门的研究领域之一。在当下的食物图像识别分类领域, 在国外, 学者 Yoshiyuki Kawano 等^[1] 提出了一种基于能移动设备端的食物识别技术; Joutou 等^[2] 提出采用多重核函数学习的方法集成颜色、纹理等图像特征用于食物图像的识别工作; Austin Myers 等^[3] 通过卷积神经网络实现对食物的识别, 并基于食物的常用烹饪方法对其营养成分进

行分析, Keiji Yanai 等人提出使用 Pre-training 和 Fine-training 来提高深度卷积神经网络对食物图像的识别精度^[4], Simon Mezgec 等^[5] 融合 4 种卷积神经网络形成深度神经网络进行食物识别。

在国内, Zhihui Fu 等^[6] 提出使用残差网络模型来进行中餐菜品食物图像的识别; 郭礼华等^[7] 使用 SIFT 特征、颜色直方图特征、梯度直方图、SURF 特征、LBP 特征和 Gabor 特征等对食材图像数据库进行了分类, 并与传统特征分类方法进行了分析与比较; Shulin Yang 等^[8] 对食物不同分成之间的特征关系进行两两统计形成食物识别分类器; Xin Chen 等^[9] 通过分别对不同的深度神经网络进行训练, 提出了一种简单集成训练好的深度神经网络模型用于中餐食物识别的研究方法。

作者简介: 边 竞(1999-), 女, 本科生, 主要研究方向: 图像处理、模式识别; 王艺璇(1999-), 女, 本科生, 主要研究方向: 图像分类; 代宇晖(1999-), 女, 本科生, 主要研究方向: 图像识别; 陈泽众(1999-), 男, 本科生, 主要研究方向: 物联网工程; 黄靖淳(1999-), 女, 本科生, 主要研究方向: 智能通信。

收稿日期: 2020-03-18

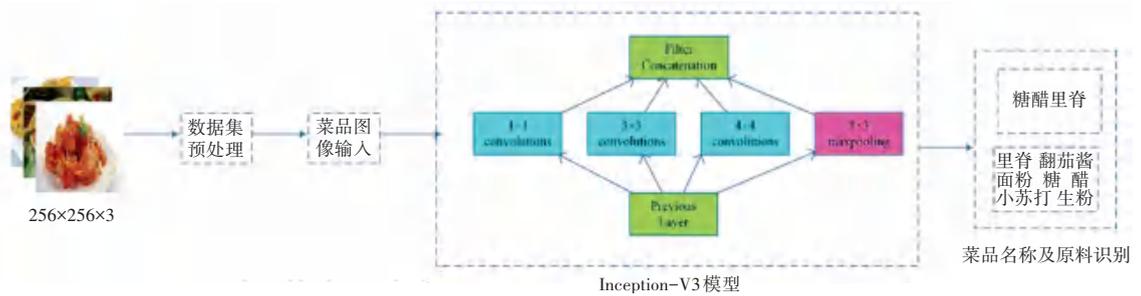


图1 基于 Inception-V3 和迁移学习的中餐菜品标签分类模型

Fig. 1 Chinese food dish label classification model based on Inception-V3 and transfer learning

此外,深度学习离不开海量训练数据的支持,由于一般的图像识别任务获取百万级别的标签数据样本相对困难,小样本数据集训练的模型往往会出现“过拟合”情况。基于上述存在的问题,迁移学习可以将具有较强泛化性能的预训练模型进行微调训练,从而可以检测到新对象类别^[10]。

综合深度学习与迁移学习的优势^[11],本文提出一种基于迁移学习和卷积神经网络的中餐菜品图像多标签分类方法。该方法首先采用 ImageNet 数据集上已完成预训练的 Inception-V3 模型初始化各层的网络参数,然后通过微调,在本文所用到的中餐菜品图像数据集上进行迁移学习,依次完成单标签、多标签分类的卷积神经网络模型的搭建;分别完成中

餐菜品图像的名称及原料的识别工作。测试实验结果表明,该方法具有一定的准确可靠性。

1 基于 Inception-V3 和迁移学习的中餐菜品标签分类模型

本文提出的基于 Inception-V3 模型的中餐菜品标签分类模型如图 1 所示,该模型选取经过预训练的 Inception-V3 作为基础网络架构,用于获取菜品图像的高层次视觉特征,从而实现中餐菜品图像的名称及原料识别工作。该模型主要分为 3 个部分:(1)中餐菜品图像数据集的预处理。(2)对经过预训练的 Inception-V3 进行微调,进行菜品图像视觉特征的提取。(3)将提取的菜品图像特征应用至名称及原料的标签分类任务。



图2 VIREO Food-172 数据集样例

Fig. 2 Sample VIREO Food-172 data set

1.1 中餐菜品图像数据集的预处理

VIREO Food-172 中餐图像数据库是 Jingjing Chen 等^[12]为进行中餐成分识别而建立的数据库,该数据库共包含 172 种中餐菜品,353 种原料标签,图片总数共达 110241 张。本文以该数据库为蓝本,在每一类菜品图像集中随机抽取 100 张图片。由于 Inception-V3 模型的输入图像像素大小至少为 224,因此首先将抽取到的图片进行归一化处理,统一尺寸为(256,256,3);其次进行数据分离,将图像进行

重命名,手工对每一种中餐菜品图像的实际位置进行标注,并注明类别标签,将对应的保存有中餐菜品位置信息及类别标签的 xml 文件进行保存;最后仿照 VOC 数据集的格式,将用于训练的图片放置到对应的目录下,同时将保存着图片中菜品信息的 xml 文件也保存至相应目录,构建中餐菜品训练数据集,作为 Inception-V3 模型的输入。

1.2 Inception-V3 网络训练和迁移学习

Inception-V3 是由 47 层网络组成的深度卷积

神经网络, 假设模型输入为 M , 网络对应的输出为 N , 则该模型通过训练得到一个参数为 Θ 的非线性映射 $F(\cdot | \Theta)^{[13]}$, 即

$$N = F(M | \Theta) = f_l(\dots f_2(f_1(M | \Theta_1) \Theta_2) | \Theta_l). \quad (1)$$

式中, $f_l(\cdot | \Theta_l)$ 表示模型的第 l 层。对 Inception-V3 的迁移训练过程如下: (1) 将 ImageNet 数据集上已完成预训练的 Inception-V3 模型参数作为初始化网络参数。(2) 将 Inception-V3 模型中 Softmax 输出层的类别数, 改成中餐菜品名称的类别数 172、原料的类别数 353。(3) 在进行监督分类的网络训练模式时, 采用后向传播算法, 在中餐菜品训练数据集上进行迁移学习, 并对模型参数进行更新。

1.3 中餐菜品标签分类任务

当完成 Inception-V3 的训练之后, 提取其最后一层 (全连接层) 产生的长度分别为 172 和 353 的特征向量作为模型最终学习到的高层次视觉特征, 然后输入到线性支持向量机分类器, 分别用于实现中餐菜品名称及原料组成的分类任务。

2 结果与分析

本实验在 Python3.7 (64 bit) 平台上完成, 具体测试环境见表 1。为测试本文方法的准确性, 在图 2 中经过预处理的中餐菜品图像数据集上进行实验的测试。该数据集数据规模相对较小, 共包含 172 种中餐菜品, 图 3 是部分中餐菜品图像样例。

表 1 实验测试环境配置

Tab. 1 Experimental test environment configuration

软硬件	指标	指标参数
	内存	128G/16G
硬件	处理器	Intel(R) Core(TM) i7-9750 CPU@ 4.50GHz
	图形显卡	NVIDIA GEFORCE GTX 1680Ti
	操作系统	Win10 家庭中文版
软件	软件框架	Tensorflow
	语言	Python

在对模型进行训练之前, 首先需要使用预训练模型 Inception-V3 进行参数设置, 部分参数设置如下: 每步使用的 Batch Size 为 64、epoch 为 50 000、迭代轮数为 5、学习率为 0.001; 其次将 VIREO Food-172 数据集划分成两部分: 10% 作为测试集; 90% 作为训练集。其中测试集 1 000 张, 训练集 10 239 张; 然后将彩色图像进行灰度化处理, 进行各层卷积特征的提取, 以便于进行菜品名称及原料类别标签的匹配, 完成中餐菜品名称及原料的标签分类任务。



图 3 测试数据集样例

Fig. 3 Sample VIREO Food-172 data set

此外, 由于损失函数值对模型精确度影响最大, 因此降低该数值可以提高模型的精确度, 即损失函数值越小, 模型精度越高, 应用和研究的价值也就越大^[14]。因此在保持其他参数值不变的前提下, 保证损失函数值相对较低可保证模型精度。本文利用 Inception-V3 模型在数据集 VIREO Food-172 上迭代训练至 50 000 次, 损失函数值整体呈现下降趋势, 表明模型的收敛效果较好。

使用上述模型对 172 类中餐菜品进行名称的单标签分类, 得出准确率为 80.85%; 对 353 种菜品原料进行多目标分类, 得出的准确率为 56.26%。部分中餐菜品的标签分类结果如图 4 所示。

		
Hot and sour potato silk	Steamed egg custard	Fried pork
Hob blocks of potato 0.531 14	Eggs 0.510 37	Scrambled pork 0.505 73
Minced green onion 0.484 92	Water 0.447 64	Minced green onion 0.328 56
Crushed hot and dry chili 0.214 41	Minced green onion 0.377 98	Chinese Parsleycoriander 0.215 01
Crushed pepper 0.134 02	Chinese Parsleycoriander 0.155 07	Crushed pepper 0.161 52
Pepper slices 0.129 99	Chili oil 0.141 07	Chili oil 0.119 11
Hot and dry pepper 0.090 44		Crushed hot and dry chili 0.115 26

图 4 部分中餐菜品名称及原料分类结果

Fig. 4 Part Chinese food dish names and raw material classification results

由图4可见,本文提出的中餐菜品名称及原料识别模型对菜品名称的识别率可达到75%以上,对菜品原料的识别率可达50%以上,表明本文提出的模型具有较高的识别性能。

3 结束语

本文针对基于卷积神经网络的中餐菜品名称、成分识别技术进行了研究,在迁移学习的基础上使用Inception-V3模型对VIREO Food-172数据集进行了标签分类,完成了中餐菜品名称及成分的识别工作,菜品名称准确率达到80.85%,菜品原料识别率达到了56.26%,取得了较好的识别性能;与传统的CNN以及SVM-HOG方法相比,本文的方法可以更快地对中餐菜品的名称及原料进行识别,且具有较好的可移植性,因此将训练好的模型与移动端架构相结合,实现智能移动设备上的APP模块化操作也将成为后续工作的重点研究方向^[15]。

参考文献

[1] KAWANO Y, YANAI K. A real-time food recognition system on a smartphone[J]. *Multimed Tools Appl*, 2014, 74(14): 1-25.

[2] JOUTOU T, YANAI K. A food image recognition system with multiple kernel learning [C]//2009 16th IEEE International Conference on Image Processing (ICIP). IEEE, 2009; 285-288.

[3] MEYERS A, JOHNSTON N, RATHOD V, et al. Im2Calories: towards an automated mobile vision food diary[C]//Proceedings of the IEEE International Conference on Computer Vision.2015:1233-1241.

[4] YANAI K, KAWANO Y. Food image recognition using deep convolutional network with pre-training and fine-tuning [C]//

2015 IEEE International Conference on Multimedia & Expo Workshops (ICMEW). IEEE, 2015; 1-6.

[5] MEZGEC S, KOROUŠIĆ SELJAK B. NutriNet: a deep learning food and drink image recognition system for dietary assessment [J]. *Nutrients*, 2017, 9(7): 657.

[6] FU Z, CHEN D, LI H. Chinfood1000: A large benchmark dataset for chinese food recognition [C]//International Conference on Intelligent Computing. Springer, Cham, 2017; 273-281.

[7] 郭礼华,罗材. 食材数据库统计与对比实验性能分析[J]. *中国图象图形学报*,2017,22(8):1079.

[8] YANG S, CHEN M, POMERLEAU D, et al. Food recognition using statistics of pairwise local features [C]//2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2010; 2249-2256.

[9] CHEN X, ZHU Y, ZHOU H, et al. ChineseFoodNet: A large-scale image dataset for chinese food recognition[J]. *arXiv preprint arXiv:1705.02743*, 2017.

[10] PAN S J, YANG Q. A survey on transfer learning [J]. *IEEE Transactions on knowledge and data engineering*, 2009, 22(10): 1345-1359.

[11] 廉小亲,成开元,安飒,等. 基于深度学习和迁移学习的水果图像分类[J]. *测控技术*,2019, 38(6):15.

[12] CHEN J, NGO C W. Deep-based ingredient recognition for cooking recipe retrieval [C]//Proceedings of the 24th ACM international conference on Multimedia. 2016; 32-41.

[13] RAMCHARAN A, BARANOWSKI K, MCCLOSKEY P, et al. Deep learning for image-based cassava disease detection [J]. *Frontiers in plant science*, 2017, 8: 1852.

[14] 贺智超. 基于深度学习和迁移学习的多任务图像分类[D]. 广州:华南理工大学,2017.

[15] 庄宁. 基于Tensorflow的Android端本地图像分类的应用设计与实现[D]. 杭州:浙江工业大学,2019.

(上接第54页)

[5] RAMANATHAN C, JIA P, GHANEM R, et al. Activation and repolarization of the normal human heart under complete physiological conditions[J]. *PNAS*,2006,103(16): 6309-6314.

[6] 景永明,王智华,潘运萍. 心电瀑布图的原理及应用[J]. *中国心脏起搏与心电生理杂志*, 2018, 032(3):291-296.

[7] 向晋涛. 心电散点图的临床应用[J]. *中国心脏起搏与心电生理杂志*, 2019(3):197-201.

[8] 刘兵,邱鹏,李桥. 心电图心肌缺血可视化显示与输出[J]. *中国医学物理学杂志*,2007, 2007(6): 450-453.

[9] ZAIDI AMA, AHMED M J, BAKIBILLAH ASM. Feature extraction and characterization of cardiovascular arrhythmia and normal sinus rhythm from ECG signals using LabVIEW [C]//IEEE International Conference on Imaging. 2017.

[10] KURANGKAR K V, NANDGAONKAR A B, NALBALWAR S L. ECG Analysis and Abnormality Detection [C]//2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS). IEEE, 2018; 1761-1764.

[11] KALIDAS V, TAMIL L. Real-time QRS detector using Stationary Wavelet Transform for Automated ECG Analysis [C]//IEEE International Conference on Bioinformatics & Bioengineering. IEEE, 2017.

[12] SEDGHAMIZ H, SANTONOCITO D. Unsupervised Detection and Classification of Motor Unit Action Potentials in Intramuscular

Electromyography Signals [C]// The 5th IEEE International Conference on E-Health and Bioengineering-EHB 2015. IEEE, 2015.

[13] KHAMIS H, WEISS R, XIE Y, et al. QRS Detection Algorithm for Telehealth Electrocardiogram Recordings[J]. *IEEE Transactions on Biomedical Engineering*, 2016, 63(7):1377-1388.

[14] JAMBUKIA S H, DABHI V K, PRAJAPATI H B. Classification of ECG signals using machine learning techniques: A survey [C]. *International Conference on Computer Engineering & Applications*.2015.

[15] AGRAWAL A, GAWALI D H. Comparative study of ECG feature extraction methods [C]// IEEE International Conference on Recent Trends in Electronics. IEEE, 2018.

[16] BHYRI C, HAMDE S T, WAGHMARE L M. ECG feature extraction and disease diagnosis [J]. *Journal of Medical Engineering & Technology*, 2011, 35(6-7):354-361.

[17] MINGJING Y. Correlation dimensions and entropy of series in electrocardiogram. [C]// International Conference of the Engineering in Medicine & Biology Society. *Conf Proc IEEE Eng Med Biol Soc*, 2005.

[18] RIZAL A, WIJAYANTO I. Classification of Premature Ventricular Contraction based on ECG Signal using Multiorder Rényi Entropy [C]//2019 International Conference of Artificial Intelligence and Information Technology (ICAIIIT), Yogyakarta, Indonesia, 2019.