

DOI: 10.16210/j.cnki.1007-7561.2024.01.012

杜慧江, 崔潇以, 王艺蒙, 等. 基于 CBAM-InceptionV3 迁移学习的食品图像分类[J]. 粮油食品科技, 2024, 32(1): 91-98.

DU H J, CUI X Y, WANG Y M, et al. Food image classification based on CBAM-Inception V3 transfer learning[J]. Science and Technology of Cereals, Oils and Foods, 2024, 32(1): 91-98.

# 基于 CBAM-InceptionV3 迁移学习的食品图像分类

杜慧江, 崔潇以, 王艺蒙, 孙丽萍

(上海健康医学院 医疗器械学院, 上海 201318)

**摘要:** 为提高食品图像自动识别分类的准确率, 提出一种嵌入通道注意力机制和空间注意力机制的卷积块注意力模块 (CBAM) 的“开端”第三版 (InceptionV3) 分类模型。将带有图像网络 (ImageNet) 预训练权重参数的 InceptionV3 模型拆分后, 在每个 Inception 块后嵌入 CBAM 模块, 再重新组装成新模型, 共嵌入 11 个 CBAM 模块。将此模型用于经过填充和缩放到  $299 \times 299$  像素的 Food-101 食品图像数据集进行迁移学习, 最高准确率达到 82.01%。与原始的 InceptionV3 模型相比, CBAM 模块能够有效提升模型的特征提取和分类能力; 同时迁移学习与从头开始训练相比也可以大幅提高准确率、缩短训练时间。与其它几类主流卷积神经网络模型进行对比实验, 结果表明该模型具有较高的识别准确率, 可为食品图像分类识别提供有力支撑。

**关键词:** 食品图像分类; 通道注意力; 空间注意力; CBAM; InceptionV3; 迁移学习

中图分类号: TP183 文献标识码: A 文章编号: 1007-7561(2024)01-0091-08

网络首发时间: 2023-12-27 17:47:09

网络首发地址: <https://link.cnki.net/urlid/11.3863.ts.20231226.1725.010>

## Food Image Classification Based on CBAM-Inception V3 Transfer Learning

DU Hui-jiang, CUI Xiao-yi, WANG Yi-meng, SUN Li-ping

(School of Medical Instruments, Shanghai University of Medicine and Health Sciences, Shanghai 201318, China)

**Abstract:** To improve the accuracy of automatic recognition and classification of food images, a classification model CBAM-InceptionV3 is proposed, which embeds the Convolutional Block Attention Module. The specific method is to split the Inception V3 model with ImageNet pre-trained weight parameters into blocks, embed CBAM modules after each Inception block, and reassemble them into a new model, embedding a total of 11 CBAM modules. This new model is used for transfer learning of Food-101 food image dataset padded and scaled to 299 pixels in both length and width, with the highest accuracy of 82.01%. Compared with the original Inception V3 model, the CBAM module can effectively improve the model's feature extraction and classification capabilities. At the same time, transfer learning can

收稿日期: 2023-08-09

基金项目: 国家重点研发计划 (2018YFB1307700)

Supported by: National Key Research and Development Project of China (No. 2018YFB1307700)

作者简介: 杜慧江, 男, 1975 年出生, 硕士, 工程师, 研究方向为医疗大数据。E-mail: duhj@sumhs.edu.cn

significantly improve the accuracy rate and shorten the training time compared with the training from scratch. Compared with several other mainstream convolutional neural network models, the results show that this new model has higher recognition accuracy and can provide strong support for food image classification and recognition.

**Key words:** food image classification; channel attention; spatial attention; CBAM; InceptionV3; transfer learning

近年来的医学研究证实,限制饮食摄入的热量对抑制胰岛素抵抗、延缓衰老等方面具有积极意义<sup>[1-2]</sup>。主动控制从食品中摄取的热量,制定并且执行有针对性的饮食健康计划是今后越来越多人的日常生活场景。此外,也更加需要对日常饮食所含的热量、是否容易引起过敏、某些疾病患者饮食营养成分等方面进行评估。因此,对各种食品进行分类、识别不可或缺。目前此类工作主要依靠手工检查分类,这种方法很容易出错,也不容易在用户中推行。基于深度学习的食品评价方案可以自动从图像中识别食品并分类,给全自动化的食物成份和营养评估带来可能。此外,以快餐店为例,自动食品识别和分类对于加快服务速度、帮助解决收银员在高需求时产生的瓶颈问题也很有意义<sup>[3]</sup>。

早期的食品分类大多基于经典的计算机图形学方法对图像中的色彩、形状轮廓、纹理等特征进行提取分析<sup>[4]</sup>。BETTADAPURA 等使用支持向量机 (Support Vector Machines, SVM), 在包含 75 个类别的食品分类实验中达到 63.33% 的准确率<sup>[5]</sup>。经典机器学习的食物图像分类技术受到特征提取算法的设计者水平的制约,食物图像分类的准确率不高,也难以批量部署。

深度学习是使用卷积计算的方式搭建神经网络模型,对大量输入的图像进行特征学习,然后将学习到的特征提取方法运用于分类器,解决了人工设计提取特征和分类方法的不足,不再需要人工设计提取图像特征的方法,将特征提取和分类识别合二为一,非常适合用于图像领域的分类<sup>[6]</sup>。CHEN J 等基于 VGG16 模型改进提出的 Arch-D 模型在 UEC Food-100 数据集上实现了 82.12% 的分类精度<sup>[7]</sup>。MEZGEC S 等设计了一个名为 NutriNet 的食品分类模型,在自建的含有 520 种

食品和饮料的数据集上的分类准确率达到 86.72%<sup>[8]</sup>。

注意力机制在卷积运算提取特征的过程中可以关注目标重点区域,获取重点区域的细节特征。本文提出了一种嵌入通道注意力机制和空间注意力机制 (Convolutional Block Attention Module, CBAM) 模块的 InceptionV3 深度学习模型,使用此模型进行迁移学习,对日常食品图像进行快速分类,可以为食物热量评估、健康饮食等应用提供强有力的支撑。

## 1 数据集及图像预处理

### 1.1 数据集

深度学习分类模型的训练过程需要大量的图像数据,如果数据规模不够,很容易造成过拟合,即在训练数据上表现良好,但是在实际数据上准确率不如训练时那么高的情况。本文选取 Food-101 作为实验的数据集,是因为这个数据集比较普及,在很多的食品分类实验中都作为基准数据集<sup>[9]</sup>。Food-101 是斯坦福大学建立的主要用于食品图像分类的数据集,包含 101 种食品类别的图像,共有 101 000 张图像。其中训练图像未经数据清洗,图像的长宽比不固定,最大边长为 512 像素。按照数据集原始的划分,每个类别的训练图像和验证图像分别有 750 张和 250 张。该数据集中的部分图片如图 1 所示。

### 1.2 短边填充和缩放处理

数据集中的图像有相当比例并不是正方形,如果直接改变纵横比例缩放到正方形,会造成图像纵横比变形,对于一些可能会被图像纵横比影响的识别任务而言,可能会影响识别的准确率。InceptionV3 模型的标准图像输入要求为 299 像素×299 像素,因此需要把 Food-101 数据集中的



图 1 Food-101 数据集中随机抽取的 12 类样本实例  
Fig.1 Twelve randomly selected sample instances from the Food-101 dataset

101 000 张图像统一做调整，匹配模型输入<sup>[10]</sup>。在输入图像的长和宽中找出较长的边，将图片按原始长宽比将长边缩放为 299 像素。然后计算出在短边上需要填充的像素数，在短边进行左右对称的单色背景填充。短边填充和缩放处理前后的对比样例如图 2 所示。

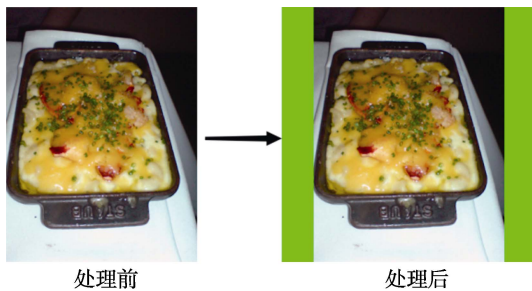


图 2 短边填充和缩放处理前后对比样例  
Fig.2 Comparison before and after short edge padding and scaling processing

和残差网络 50 (ResNet50)<sup>[14]</sup>等。本文采用了经典模型 InceptionV3，其主干结构如表 1 所示。

表 1 InceptionV3 模型主干结构  
Table 1 Stem of InceptionV3 model

层次类型	卷积核大小/步长 (像素×像素)或模块名	输入尺寸/(像素× 像素×通道数)
Conv	3×3/2	299×299×3
Conv	3×3/1	149×149×32
Conv padded	3×3/1	147×147×32
Pool	3×3/2	147×147×32
Conv	3×3/1	73×73×64
Conv	3×3/2	71×71×80
Conv	3×3/1	35×35×192
3×Inception	Inception 模块	35×35×288
5×Inception	Inception 模块	17×17×768
2×Inception	Inception 模块	8×8×1 280
Pool	8×8	8×8×2 048
Linear	logits 函数	1×1×2 048
Softmax	分类器	1×1×1 000

## 2 构建模型

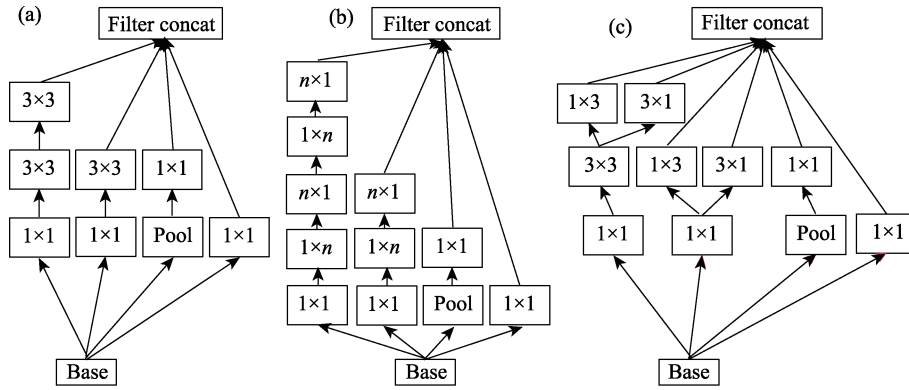
### 2.1 InceptionV3 网络模型主干

在名为“张量流”(Tensorflow)的深度学习框架中包括了一些典型的预训练模型，如 AlexNet<sup>[11]</sup>、视觉几何组(VGG)<sup>[12]</sup>、GoogLeNet<sup>[13]</sup>

InceptionV3 模型的核心是 Inception 块，共有 3 种 Inception 块。这三种 Inception 块分别命名为 Inception A、Inception B、Inception C，它们的结构如图 3 所示。Inception A 是将 5×5 卷积核替换成 2 个 3×3 卷积核，在保持同等特征提取能力的

同时，降低了计算量。Inception B 在中度大小的特征图上使用非对称卷积，用  $1 \times n$  卷积操作拼接  $n \times 1$  卷积来替换  $n \times n$  卷积，可以减少计算量提升计算效率。Inception C 使用  $1 \times 3$ 、 $3 \times 1$ 、 $3 \times 3$  等卷积多次拼接，加特征提取视野宽。GoogLeNet 的 InceptionV3 预训练模型深度为 189 层，其中共堆

叠了 10 个 Inception 块，已经在 ImageNet 数据集的一个子集上完成了训练，可以对图像进行 1 000 个类别的分类任务<sup>[10]</sup>。本文选用 Tensorflow 深度学习框架的 Keras 库中预置的 InceptionV3 模型，模型实例化时可以选择是否加载在 ImageNet 数据集训练得到的权重参数。



(a) Inception A; (b) Inception B; (c) Inception C

图 3 3 种 Inception 块结构

Fig.3 Three Inception Block Structures

## 2.2 CBAM 模块

CBAM 将通道注意力机制和空间注意力机制融合在一起，成为一个单独的模块，可以方便地

嵌入到现有模型中，给原本没有注意力机制的模型同时加入通道与空间上的注意力机制。CBAM 结构如图 4 所示：

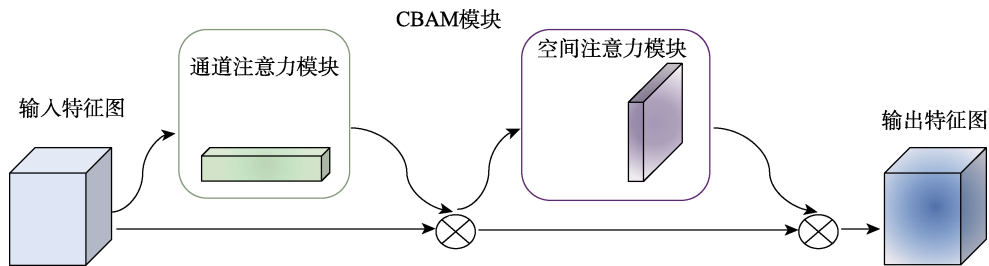


图 4 CBAM 模块结构图

Fig.4 Structure diagram of CBAM module

通道注意力机制是对特征图中的多个通道之间可能存在的关联信息进行关注。为有效计算通道注意力，需要对输入特征图的空间维度进行压缩，常用的方法是平均池化和最大池化。平均池化关注特征的普遍性，最大池化关注特征的独特性，二者同时使用，可以更好地提取通道间隐含的关联信息。如公式 (1) 所示， $F_{avg}^c$  和  $F_{max}^c$  分别表示平均池化特征和最大池化特征。最后，这两组特征被转入一个共享的卷积网络(MLP)，经过  $\sigma$  表示的 Sigmoid 函数，产生通道注意力的输出特征图  $M_c(F)$ 。

$$M_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) = \sigma(W_1(W_0(F_{avg}^c)) + W_1(W_0(F_{max}^c))) \quad \text{式 (1)}$$

空间注意力机制对特征图中值得关注或压制的区域进行聚焦，是对通道注意力机制的补充。将通道注意力模块输出的特征图  $F$  应用平均池化和最大池操作，然后将两个操作结果连接起来，应用卷积操作生成大小为  $R \times H \times W$  的空间注意力特征图  $M_s(F)$ ，如公式(2)所示。 $F_{avg}^s$  大小为  $1 \times H \times W$ ， $F_{max}^s$  大小也为  $1 \times H \times W$ 。 $\sigma$  表示 sigmoid 函数， $f_{7 \times 7}$  表示一个大小为  $7 \times 7$  的卷积运算。

$$\begin{aligned}
 M_s(F) &= \sigma(f^{7 \times 7}([\text{AvgPool}(F); \text{MaxPool}(F)])) \\
 &= \sigma(f^{7 \times 7}([F_{\text{avg}}^s; F_{\text{max}}^s])) \quad \text{式 (2)}
 \end{aligned}$$

CBAM 模块将通道注意力机制和空间注意力机制结合, 融入一个模块中, 在不大幅度增加模型参数和规模的前提下, 快速拓展原有模型能力, 体现了 CBAM 模块即插即用的特点<sup>[15]</sup>。

### 2.3 迁移学习

与传统的机器学习方法相比, 深度学习需要大量数据进行训练, 通过计算损失值和反向传播来不断优化权重参数, 因此极其依赖大规模训练数据。许多深度学习算法的一个主要假设是, 训练和未来的数据必须是在相同的特征空间中, 具有相同的分布。然而, 在许多实际应用中, 这一假设可能并不成立。在一些特殊领域, 训练数据不足不可避免, 构建大规模、高质量的带标注数据集非常困难。此外, 使用大型数据集对模型从头开始训练, 需要花费大量的时间和算力。而使用迁移学习, 模型的权重参数无需从头开始训练, 可以显著降低对训练数据和训练时间的需求。在实际的图像分类任务中, 可以利用预先在大批量、高质量的数据集

上训练完成的模型, 迁移到不同的图像分类任务上, 只要进行少量的微调训练, 就可以快速地得到一个较高的识别准确率<sup>[16]</sup>。良好预训练的模型可以发现数据的高维度特征, 后续层次能够从这些高维度特征中抽取目标任务需要的特征。迁移学习在源域学习到的知识可以迁移到目标域中, 这些知识可以帮助目标域的任务具有更好的模型性能<sup>[17-18]</sup>。

本文使用的迁移学习训练是把在 ImageNet 大型数据集中训练出的带有特征参数的 InceptionV3 模型, 先进行拆分。以 Inception 块为单位, 将原 InceptionV3 模型拆分成 Inception 块之前的 0 号块、1~10 号 Inception 块、及全部 Inception 块输出 11 号块。在拆分出的每个 Inception 块的矩阵拼接 (Concatenate) 运算之后添加一个 CBAM 模块, 然后再把这些模块拼接起来。通过一个全局平均池化 (GlobalAverage Pooling) 层, 连接 101 个输出的全连接层, 再添加第 1 个 CBAM 模块, 共计添加 11 个 CBAM 模块。最后使用 softmax 函数对结果进行分类输出, 构成一个嵌入了 11 个 CBAM 模块的改进版 CBAM-InceptionV3 模型。用此模型进行迁移学习的结构示意图如图 5 所示。

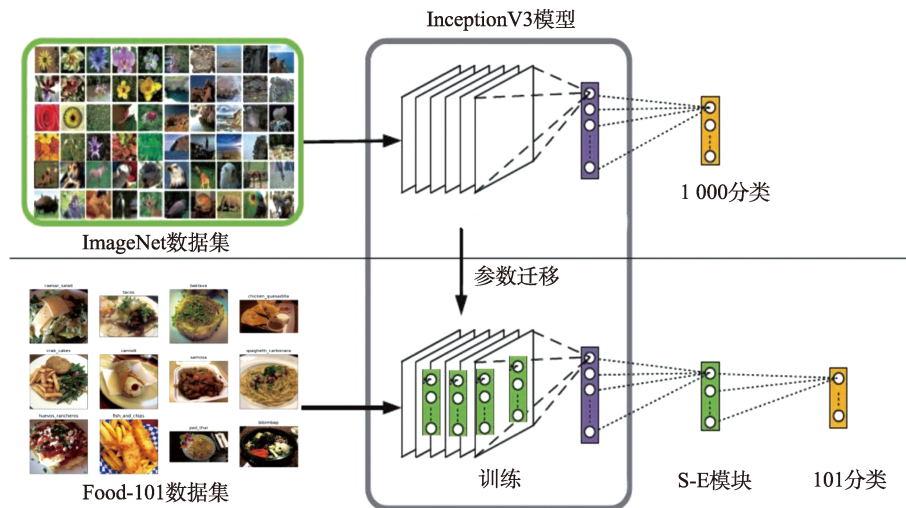


图 5 CBAM-InceptionV3 模型迁移学习示意图

Fig.5 Schematic diagram of transfer learning on CBAM-Inception V3 model

## 3 实验结果

### 3.1 实验环境

本文所使用计算机硬件: CPU 为 AMD Ryzen5 5500 处理器, 32G 内存, GPU 为 NVIDIA 3080 10GB 显存; 操作系统为 Manjaro Linux, CUDA

版本为 12.1, Tensorflow 2.12 深度学习框架。

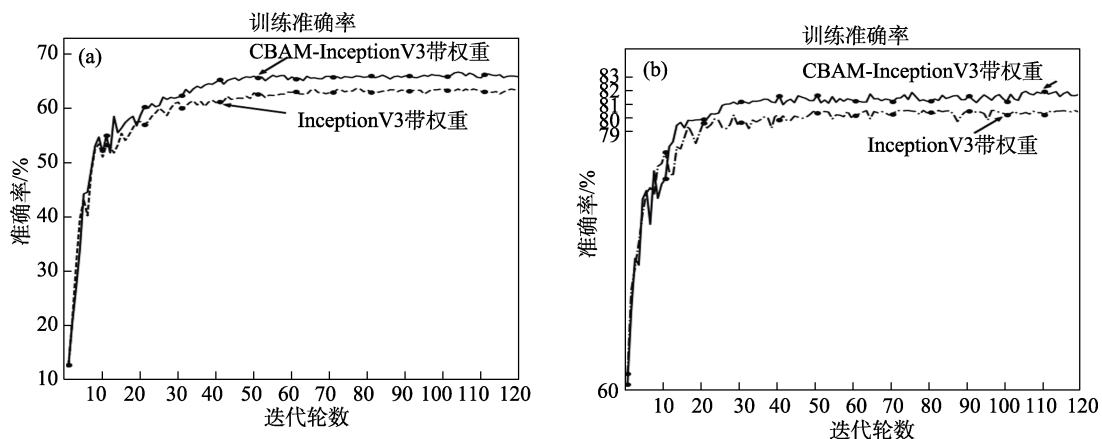
### 3.2 训练与结果分析

为了验证本文提出模型的效果, 对 Food-101 数据集在相同参数的设置下做对比实验。分别比较: InceptionV3 不带 ImageNet 权重参数从头开

始训练、InceptionV3 带 ImageNet 权重参数后迁移学习训练、嵌入 CBAM 模块的 InceptionV3 不带 ImageNet 权重参数从头开始训练，和嵌入 CBAM 模块的 InceptionV3 带 ImageNet 权重学习训练的准确率。训练参数为：训练迭代 epoch 设置为 120，优化器选用学习率 0.002、动量为 0.9 的随机梯度下降法（Stochastic gradient descent, SGD）。在传统梯度下降算法的基础上添加动量项，可以有效避免局部最优解附近的振荡，加速逼近全局最优解<sup>[19]</sup>。批大小 batch\_size 设置为 16，其主要受限制于 GPU 的显存容量，显存容量大则可以将批大小适当调高。对数据集中的 101 000 张食品图像，采用数据集默认的训练集和验证集划分，75 750(101×1 000×0.75)张作为训练集，25 250(101×1 000×0.25)张作为验证集。为了比较准确地对比模型之间的性能差异，避免由图像增强过程中某些随机操作可能带来的偏差，选择不

使用图像增强操作。

训练过程的准确率曲线如图 6 所示。从图中可以看出，对于原始的 InceptionV3 模型，迁移学习训练比从头开始训练的损失值更小、准确率更高。在 ImageNet 数据集上训练得到的权重参数从第一轮次的训练就显示出较高的准确率。在训练结束时，基于迁移学习训练的方法具有更高的精度。以整个训练过程中出现的最优准确率数据比较，带有 ImageNet 权重参数的原始 InceptionV3 模型比不带 ImageNet 权重参数的原始 InceptionV3 模型，准确率高出 17.09%。带有 ImageNet 权重参数的 CBAM InceptionV3 模型比不带 ImageNet 权重参数的 CBAM InceptionV3 模型，准确率高出 15.44%。由此分析，迁移学习对于缺少足够大规模数据集和大规模训练过程的深度学习分类任务来说，可以显著提高分类结果的准确率<sup>[20]</sup>。



注：(a) 不带 ImageNet 权重 (b) 带 ImageNet 权重

Note: (a) without ImageNet weights (b) with ImageNet weights

图 6 训练过程的准确率曲线

Fig.6 The accuracy curve of the training process

另一方面，嵌入了 CBAM 模块的 InceptionV3 模型，在带 ImageNet 权重参数的情况下，准确率比原始 InceptionV3 模型的准确率提升了 1.40%。而在不带 ImageNet 权重的情况下，嵌入了 CBAM 模块的 InceptionV3 模型对比原始 InceptionV3 模型的准确率提升则很明显，提升了 3.05%。可以分析得出，CBAM 模块在未经过大规模数据集充分训练的模型上，对分类准确性的提升更加明显；在已经经过大规模数据集和大规模训练过程的模

型上，相较于其模块增加的少量计算量而言，其分类准确率的提升表明，将通道注意力机制和空间注意力机制嵌入已有的成熟模型对于性能提升仍然是有意义和值得的<sup>[21-22]</sup>。

### 3.3 模型对比

分别使用 MobileNetV2<sup>[23]</sup>、ResNet50，与 InceptionV3、CBAM InceptionV3 模型在 Food-101 数据集上进行对比实验。训练参数为：训练迭代 epoch 设置为 120，优化器选用学习率 0.002、动

量为 0.9 的带动量随机梯度下降法 SGD, 批大小 batch\_size 设置为 16。不同模型训练准确率对比如表 2 所示。实验数据显示: 嵌入 CBAM 模块的 InceptionV3 模型的准确率明显提高、对食品进行分类识别效果最佳。同时, 最优迭代轮次也因为计算量的增加而更晚出现。

表 2 不同模型平均分类准确率对比  
 Table 2 Comparison of average classification accuracy of models

模型	ImageNet 权重	最优迭代轮次	准确率/%
MobileNetV2	带	86	77.83
ResNet50	带	10	74.17
InceptionV3	不带	78	63.52
CBAM-InceptionV3	不带	109	66.57
InceptionV3	带	93	80.61
CBAM-InceptionV3	带	104	82.01

Food-256 数据集是比 Food-101 更大的食品图像数据集。为了测试本文提出模型的性能, 在 Food-256 数据集中选取了与 Food-101 类别相同的图像 1 000 张, 作为测试集, 对模型进行测试。结果显示, 基于 CBAM-InceptionV3 的模型总体预测准确率达到 77.36%, 略低于验证集准确率, 并没有出现明显的过拟合现象。

使用带有在 ImageNet 数据集上预先训练好的权重参数的模型, 在新数据集上进行迁移学习训练, 可以显著缩短训练时间, 获得比从头开始训练更高的精度, 降低模型训练的成本<sup>[24]</sup>。ImageNet 数据集训练产生的权重参数在其它数据集上的泛化能力强, 在 ImageNet 数据集上的精确率的小小的提升就能够导致在其它数据集上性能的提升<sup>[25]</sup>。


#### 4 结语

为提高食品图像自动识别分类的准确率, 本文提出一种嵌入多个 CBAM 模块的改进 InceptionV3 模型进行迁移学习训练的食品图像分类方法。对 InceptionV3 模型进行拆分, 在每个 Inception 块和最后的输出分类操作之后嵌入通道注意力机制 CBAM 模块, 再拼接为完整的 CBAM-InceptionV3 模型, 输入食品图像数据集, 进行迁移学习训练, 最后进行食品图像分类。

实验结果表明: 加入 CBAM 通道注意力机制有利于增强网络识别精度, 对比分析显示, 本文提出的模型具有良好的食品识别分类能力, 对于餐饮行业人工智能应用具有良好的支撑能力。

#### 参考文献:

- [1] WAZIRY R, RYAN C, CORCORAN D. Effect of long-term caloric restriction on DNA methylation measures of biological aging in healthy adults from the calerie trial[J]. *Nature Aging*, 2023, 3(3): 248-257.
- [2] ANDERSON R, WEINDRUCH R. The caloric restriction paradigm: implications for healthy human aging[J]. *American Journal of Human Biology: The Official Journal of the Human Biology Council*, 2012, 24(2): 101-106.
- [3] 彭耿, 刘宁钟. 基于注意力机制的食物识别与定位算法[J]. *计算机技术与发展*, 2022, 32(11): 121-126.
- [4] PENG G, LIU N Z. Food recognition and localization algorithm based on attention mechanism[J]. *Computer Technology and Development*, 2022, 32(11): 121-126.
- [5] ZHENG W, FU X, YING Y. Spectroscopy-based food classification with extreme learning machine[J]. *Chemometrics and Intelligent Laboratory Systems*, 2014, 139: 42-47.
- [6] BETTADAPURA V, THOMAZ E, PARNAMI A, et al. Leveraging context to support automated food recognition in restaurants[C]//2015 IEEE Winter Conference on Applications of Computer Vision. Waikoloa, HI, USA: IEEE, 2015: 580-587.
- [7] 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.
- [8] CHEN J, NGO C. Deep-based ingredient recognition for cooking recipe retrieval[C]//Proceedings of the 24th ACM international conference on Multimedia. Amsterdam The Netherlands: ACM, 2016: 32-41.
- [9] MEZGEC S, SELJAK B. NutriNet: a deep learning food and drink image recognition system for dietary assessment[J]. *Nutrients*, 2017, 9(7): 657.
- [10] BOSSARD L, GUILLAUMIN M, VAN GOOL L. Food-101-mining discriminative components with random forests[C]// European Conference on Computer Vision, 2014: 446-461.
- [11] TEAM K. Keras documentation: InceptionV3[EB/OL]. [2023-04-24]. <https://keras.io/api/applications/inceptionv3/>.
- [12] KRIZHEVSKY A, SUTSKEVER I, HINTON G. ImageNet classification with deep convolutional neural networks[J]. *Communications of the ACM*, 2017, 60(6): 84-90.
- [13] SAINATH T, KINGSBURY B, SAON G, et al. Deep convolutional neural networks for large-scale speech tasks[J]. *Neural Networks*, 2015, 64(4): 39-48.

- [13] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV: IEEE Computer Society, 2016: 2818-2826.
- [14] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift[C]//International Conference on International Conference on Machine Learning. Cambridge MA: JMLR, 2015: 448-456.
- [15] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[J]. In Computer Vision-ECCV, 2018: 3-19.
- [16] TAN C, SUN F, KONG T, et al. A survey on deep transfer learning[C]//Artificial Neural Networks and Machine Learning-ICANN 2018. Cham: Springer International Publishing, 2018: 270-279.
- [17] PAN S, YANG Q. A Survey on transfer learning[J]. IEEE Transactions on Knowledge and Data Engineering, 2010, 22(10): 1345-1359.
- [18] 崔庭源, 杨其长, 张义, 等. 基于迁移学习和卷积神经网络的生菜鲜重估测[J]. 中国农业大学学报, 2022, 27(11): 197-206.  
CUI T Y, YANG Q C, ZHANG Y, et al. Fresh reestimation of lettuce based on transfer learning and convolutional neural networks[J]. Journal of China Agricultural University, 2022, 27(11): 197-206.
- [19] 史加荣, 王丹, 尚凡华, 等. 随机梯度下降算法研究进展[J]. 自动化学报, 2021, 47(9): 2103-2119.  
SHI J R, WANG D, SHANG F H, et al. Progress in the stochastic gradient descent algorithm[J]. Journal of Automation, 2021, 47(9): 2103-2119.
- [20] 石祥滨, 房雪键, 张德园, 等. 基于深度学习混合模型迁移学习的图像分类[J]. 系统仿真学报, 2016, 28(1): 167-173+182.  
SHI X B, FANG X J, ZHANG D Y, et al. Image classification based on deep learning hybrid model transfer learning[J]. Journal of System Simulation, 2016, 28(1): 167-173+182.
- [21] 徐沁, 梁玉莲, 王冬越, 等. 基于 SE-Res2Net 与多尺度空谱融合注意力机制的高光谱图像分类[J]. 计算机辅助设计与图形学学报, 2021, 33(11): 1726-1734.  
XU Q, LIANG Y L, WANG D Y, et al. Hyperspectral image classification based on the attention-attention mechanism of SE-Res2Net and multi-scale empty spectral fusion[J]. Journal of Computer Aided Design and Graphics, 2021, 33(11): 1726-1734.
- [22] 张晓倩, 罗建, 杨梅, 等. 基于改进 SE-Net 网络与多注意力的脑肿瘤分类方法[J/OL]. 西华师范大学学报(自然科学版): 1-12[2023-04-25]. <http://kns.cnki.net/kcms/detail/51.1699.N.20230317.1725.002.html>  
ZHANG X Q, LUO J, YANG M, et al. Brain tumor classification method based on an improved SE-Net network with multiple attention[J/OL]. Journal of China West Normal University (Natural Science Edition): 1-12 [2023-04-25]. <http://kns.cnki.net/kcms/detail/51.1699.N.20230317.1725.002.html>
- [23] MobileNetV2: inverted residuals and linear bottlenecks | IEEE Conference Publication | IEEE Xplore[EB/OL]. [2023-04-24]. <https://ieeexplore.ieee.org/abstract/document/8578572/>.
- [24] 张钢, 张石清. 基于 DCNN 和迁移学习的食物图像识别[J]. 实验室研究与探索, 2019, 38(6): 111-114.  
ZHANG G, ZHANG S Q. Food image recognition based on DCNN and transfer learning[J]. Laboratory Research and Exploration, 2019, 38(6): 111-114.
- [25] KORNBLITH S, SHLENS J, LE Q. Do better ImageNet models transfer better?[M/OL]. arXiv, 2019[2023-04-24]. <http://arxiv.org/abs/1805.08974>. 
- 备注: 本文的彩色图表可从本刊官网 (<http://lyspkj.ijournal.cn>)、中国知网、万方、维普、超星等数据库下载获取。