

DOI: 10.13718/j.cnki.xdzk.2021.08.018

基于 CapsNet 神经网络的树叶图像分类模型

张冬妍, 韩睿, 张瑞, 曹军

东北林业大学 机电工程学院, 哈尔滨 150040

摘要: 对树木研究的基础是对其进行分类处理. 本文结合 CapsNet 神经网络模型, 以提高树叶分类的准确率为目的, 使用实验室拍摄的 10 种树叶图片建立树叶分类模型. 考虑到模型效率和图像大小, 在原有 CapsNet 上与传统卷积神经网络相结合, 通过优化动态路由算法对 CapsNet 进行改进, 得到了 E-CapsNet 网络模型, 同时与经典的神经网络模型 AlexNet 和 Inception V3 模型进行对比. 经过 50 次 epoch 的训练, 模型训练准确率最高达到 99.15%, 验证集的准确率为 98.51%, 测试集准确率为 98.63%, 对比原 CapsNet 网络, 测试集准确率提高了 2.51%. 实验结果表明, 改进后的 E-CapsNet 模型实现了更高的精度.

关键词: 胶囊网络; 神经网络; 图像分类; 树叶识别; 动态路由

中图分类号: TP391.4; S718.3

文献标志码: A

文章编号: 1673-9868(2021)08-0143-09

目前, 世界上大约有 6 万多种树木, 对树木研究的基础是对其进行分类处理, 所以对其分类的方法进行研究具有很重要的意义. 相对于树干、树形和树木等生物信息, 树叶具有特定的外形轮廓和纹理等固有特征, 易于区分和鉴定, 常作为区分树木种类的重要标志^[1]. 树叶的分类与识别对于区分树的种类、对树木物种种群的跟踪和保护、建立树木种类数据库、植物基础医学研究、树木分布等具有十分重要的意义^[2].

传统的树叶分类识别主要基于图像处理技术, 对树叶图像进行认为选定的特征来进行分类, 识别效率较低且泛化能力弱. 近年人工智能快速发展, 基于深度学习的算法在图像分类识别方面表现出巨大的优势, 可以快速而有效地提取和分析图像中复杂深层次的特征, 对于图像分类识别具有很好的效果^[3].

文献[4-6]研究了卷积神经网络在图像和语音中的应用, Krizhevsky^[7] 2012 年在 ILSVRC 图像识别比赛中使用 AlexNet 模型获得冠军, 首次实现了 15.40% 的错误率, 引起了深度学习和卷积神经网络研究的兴起和爆炸性增长. 在后续的 ZFnet, VGG, GoogleNet, Resnet 等深度学习网络模型中^[8-11], 图像识别的错误率逐渐下降到 2.25%, 展现出深度学习在图像分类识别中的优势. Grinblat 等^[12] 利用静脉形态学的模式识别方法对植物品种进行识别, 能够准确识别 3 种豆科植物. 2016 年, Sladojevic 等^[13] 利用深层神经网络识别了 13 种植物病虫害, 平均准确率为 96.30%. 同年, 张帅等^[14] 设计了一个 8 层卷积神经网络学习系统, 并用 softmax 和 SVM 两种分类器进行树叶分类, 对单一背景下的树叶分类准确率高达 91.11% 和 90.90%. 2017 年, Cruz 等^[15] 研究了一种基于迁移学习和深度学习的可视化植物病害检测系统, 采用卷积神经网络结合迁移学习的方式, 对少量图像样本进行训练和验证, 可以自动、定量地检测早期植物病害, 降低了诊断时间和成本. 2019 年, Barbedo 等^[16] 采用 CNN 和 GoogleNet 结构进行深度学习, 建立了有效的病虫害检测与识别模型, 并与原始图像进行了病虫害和斑点的特征提取, 而原始图像在不使用手工去除背景的情况下, 整体准确率提高到 94%.

收稿日期: 2020-10-20

基金项目: 黑龙江省自然科学基金项目(C2017005).

作者简介: 张冬妍, 副教授, 博士, 主要从事机器学习, 图像处理研究.

为了提高树叶分类的准确性,本文使用 CapsNet 神经网络模型对树叶进行分类识别.首先,建立了树叶图像的数据集;其次,对数据集中的树叶图像进行预处理和数据增强;最后,对 CapsNet 网络进行优化改进,对处理后的图像进行训练和预测.实验结果表明,改进后的 CapsNet 在树叶分类上的效果明显优于原 CapsNet,有助于实现树叶种类的快速识别,为建立树种数据库、树木种群的跟踪和保护提供了重要的技术储备.

1 CapsNet 网络模型

1.1 CapsNet 神经网络

在深度学习中,卷积神经网络(CNN)在图像分类识别方面效果非常好,但 CNN 的内部数据表示没有考虑特征之间的重要空间层级关系.虽然 CNN 可以通过最大池化或者添加后续卷积层来减少通过网络的数据空间的大小,从而增加上层神经元的“视野”,即获得输入图像较大区域的高阶特征,在一定程度上提高了模型对图像的识别能力,但是,最大池化会损失有价值的特征信息,使用传统卷积网络也无法解决低级特征和高级特征空间位置关系的根本问题.Sabour 等^[17]在 2017 年提出了基于胶囊系统和动态路由机制的 CapsNet 神经网络,在 MNIST 数据集上实现了极高的分类准确率.胶囊用于对图像中特征的性质和空间关系进行封装编码,例如位置、大小、方向、变形、速度、反射率、颜色和纹理等^[18].在多层胶囊系统中,低层胶囊将其输出发送给对该输出结果表示“认同”的高一层胶囊,这一步是通过囊间动态路由算法实现的^[19].

1.2 胶囊之间的传播过程

如图 1 所示,CapsNet 网络模型的输入 u_i 是一个矢量,它是一个胶囊单元,它类似卷积神经网络中的神经元,但输入输出不是标量而是矢量.将 u_i 分别乘以不同的权重 W_{ij} ,然后得到底层胶囊的输出,即预测矢量 $\hat{u}_{j|i}$,然后 $\hat{u}_{j|i}$ 再乘以相应的耦合系数 c_{ij} ,即进行加权求和得到 s_j :

$$\hat{u}_{j|i} = W_{ij} u_i \quad (1)$$

$$s_j = \sum_i c_{ij} \hat{u}_{j|i} \quad (2)$$

式中, c_{ij} 是由迭代动态路由过程确定的耦合系数,而且胶囊网络模型中不存在偏置项.

CapsNet 中使用非线性激活函数 Squashing 将矢量的大小缩放到 0~1 之间.得到 s_j 后,再通过这个 Squashing 函数将矢量 s_j 转化成矢量 v_j 如式(3)所示:

$$v_j = \frac{\|s_j\|^2}{1 + \|s_j\|^2} \frac{s_j}{\|s_j\|} \quad (3)$$

由于这个函数是非线性的,输出的 v_j 保持了 s_j 的维度,不改变矢量的方向,只改变了矢量的大小,其输出矢量的长度代表着胶囊检测的给定特征的概率,因此,每个胶囊的概率在 0~1 之间.

1.3 囊间动态路由算法

使 CapsNet 网络模型中的耦合系数 c_{ij} 的更新迭代是通过动态路由算法实现的.动态路由算法的输入是预测矢量 $\hat{u}_{j|i}$ 和路由迭代次数 r . b_{ij} 是一个临时变量,它的值会在迭代过程中进行更新,当整个动态路由算法运行完毕后,它的值将被保存到 c_{ij} .在训练开始时, b_{ij} 被初始化为零.

以下过程会被迭代 r 次:首先使用 softmax 计算所有权重 c_{ij} 的值,如式(4)所示:

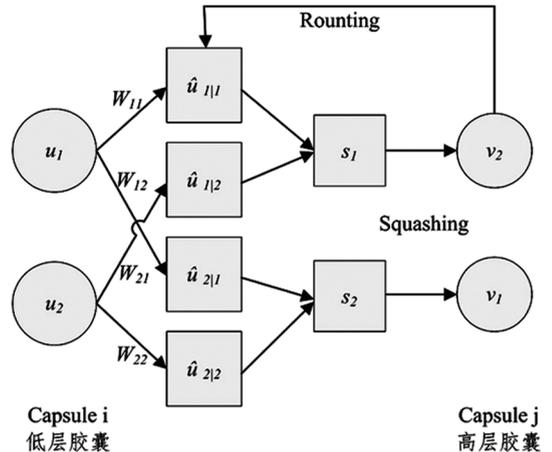


图 1 胶囊之间的传播过程

$$c_{ij} = \frac{\exp b_{ij}}{\sum_k \exp b_{ik}} \quad (4)$$

softmax 可以确保所有权重 c_{ij} 均为非负数, 且使其总和等于 1.

然后将 $\hat{\mathbf{u}}_{j|i}$ 加权求和得到 \mathbf{s}_j , 再使用 squashing 函数计算得到 \mathbf{v}_j 的操作. 接下来进行迭代的最后一步, 检查每个高层胶囊输入并更新相应的权重 b_{ij} :

$$b_{ij} \leftarrow b_{ij} + \hat{\mathbf{u}}_{j|i} \cdot \mathbf{v}_j \quad (5)$$

在更新权重之后, 算法将返回并重新开始计算 c_{ij} 的值, 重新开始这一过程, 并重复 r 次. 动态路由算法很容易收敛, 但也存在着过拟合问题, 虽然增加迭代次数可以提高准确率, 但是会增加泛化误差, 所以不宜过多迭代.

1.4 损失函数

对于图像中出现的每个类别 k , 胶囊使用单独的利润损失 L_k 如式(6)所示:

$$L_k = T_k \max(0, m^+ - \|\mathbf{v}_k\|)^2 + \lambda(1 - T_k) \max(0, \|\mathbf{v}_k\| - m^-)^2 \quad (6)$$

如果存在 k 类图片, 则设置 $T_k=1$, 并且 $m^+=0.9$, $m^-=0.1$, 将 λ 取值为 0.5 来降低某些类别没有出现时的损失, 这样可以保证表示该类别的数字胶囊的向量模值尽可能大, 其他类别的向量模值尽可能小, 以实现准确分类. 总损失就是所有数字胶囊损失的总和.

2 实验设计

2.1 数据准备和预处理

使用的树叶是实验室自行拍摄采集即建立的深度学习数据集, 数据集中采集了桉树、龙眼、木荷、珊瑚树、水翁树、香樟、柚树、桃花心木、荷花木兰、深山含笑共 10 种树木的树叶(图 2), 共计 3 193 张图片, 每一种树叶图片约 300 张左右.



图 2 树叶图片

在使用含 3 193 张图像的数据集训练网络模型之前, 本文将图像像素大小从原始的 $1\ 536 \times 1\ 536$ 调整为 147×147 . 由于深度学习对数据集的数据量有一定的要求, 研究中每种树叶的图片数量在 300 张左右, 这无法满足算法在实践过程中的泛化能力, 会导致模型过拟合. 虽然可以通过现有的其他数据集和网络途径获取更多的图片数据, 但获取的图片质量参差不齐, 无法使数据集数量有效增多. 为了减少树叶图像在训练时的过拟合, 增强模型的泛化能力, 本文尝试用数据增强来解决, 数据增强主要通过通过对树叶图像进行图像变换处理, 例如旋转、缩放、平移、尺度变换等, 本文使用的图像增强方法如下:

1) 将原始图像向随机方向平移 10 到 20 像素;

- 2) 随机选择角度对图像进行旋转;
- 3) 对图像进行轻微模糊处理;
- 4) 向图像添加随机噪声.

经过图像增强处理后, 10 种树叶图像增强至 15 965 张, 分别是桉树 1 485 张, 深山含笑 2 215 张, 龙眼 1 530 张, 木荷 1 310 张, 荷花木兰 1 090 张, 珊瑚树 2 810 张, 水翁树 1 335 张, 桃花心木 1 535 张, 香樟 1 295 张, 柚树 1 360 张, 用于训练的数据集数量如表 1 所示. 其中, 桉树叶的增强示例图像如图 3 所示.

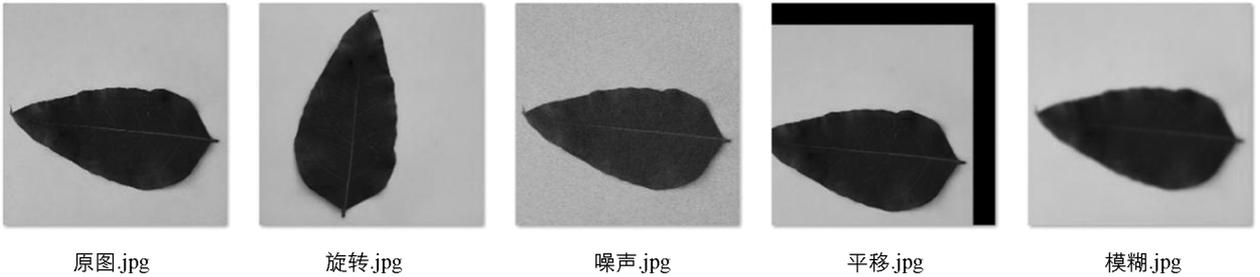


图 3 数据增强示例

表 1 数据集数量

	原图像数量/张	数据增强后图像数量/张		原图像数量/张	数据增强后图像数量/张
桉树	297	1 485	水翁树	267	1 335
深山含笑	443	2 215	桃花心木	307	1 535
龙眼	306	1 530	香樟	259	1 295
木荷	262	1 310	柚树	272	1 360
荷花木兰	218	1 090			
珊瑚树	562	2 810	总计	3 193	15 965

在扩充后的数据集中随机抽取 12 772 张图片(占数据集的 80%)作为训练集, 1 597 张(约占数据集的 10%)作为验证集, 剩余 1 596 张(约占数据集的 10%)作为测试集. 训练集用于训练 CapsNet 的模型参数, 验证集在每一个 epoch 完成后测试当前模型的准确率, 根据准确率对模型的超参数(学习率, 迭代次数等参数)进行调整, 从而得到一个更好的模型, 测试集不参与模型的训练过程, 用于测试已训练好的模型的性能.

2.2 动态路由算法的改进

采用两种方法对动态路由环节进行优化改进. 一种是引入 Adam 函数, 另一种是改进压缩函数.

陈立潮等^[20]在动态路由环节加入了 Adam 算法, 可以使模型收敛更加平稳. 算法对胶囊间参数的误差进行移动指数平均, 动态调整路由算法的 Learning Rate, 对权重 b_{ij} 进行调整.

本文在其基础上优化了 Adam 算法的部分环节, 对动态路由环节进行改进, 迭代时用平移后的 Sigmoid 乘以 $s_j / \|s_j\|$ 代替 Squashing 函数, 在最后输出的时候使用 Squashing 函数, 改进后的算法如下:

算法 1: 改进动态路由

初始化: $b_{ij} \leftarrow 0$, 一阶矩 $m_t = 0$, 二阶矩 $n_t = 0$, 一阶矩指数衰减率 $\alpha = 0.9$, 二阶矩指数衰减率 $\beta = 0.999$, 稳定数值常数 $\epsilon = 10^{-8}$, 时间步 $t = 0$.

1: procedure I-ROUTING($\hat{u}_{j|i}$, r , l)

2: for all capsule i in layer l and capsule j in layer $(l+1)$: $b_{ij} \leftarrow 0$, $m_t \leftarrow 0$, $n_t \leftarrow 0$, $\alpha \leftarrow 0.9$, $\beta \leftarrow 0.999$, $\epsilon \leftarrow 10^{-8}$, $t = 0$, $\eta \leftarrow 0.001$.

3: for r iterations do

4: $c_i \leftarrow \text{softmax}(b_t)$

- 5: $\mathbf{s}_j \leftarrow \mathbf{s}_j = \sum_i c_{ij} \hat{\mathbf{u}}_{j|i}$
- 6: $\mathbf{v}_j \leftarrow \text{Sigmoid}(\|\mathbf{s}_j\| - 5) \cdot \mathbf{s}_j / \|\mathbf{s}_j\|$
- 7: $m_t \leftarrow \alpha m_{t-1} + (1 - \alpha) \mathbf{u}_{j|i} \mathbf{v}_j$
- 8: $n_t \leftarrow \beta n_{t-1} + (1 - \beta) (\mathbf{u}_{j|i} \mathbf{v}_j)^2$
- 9: $\hat{m}_t \leftarrow m_t / (1 - \alpha^t)$
- 10: $\hat{n}_t \leftarrow n_t / (1 - \beta^t)$
- 11: $b_{ij} \leftarrow b_{ij} + \eta \cdot \hat{m}_t / (\sqrt{\hat{n}_t} + \epsilon)$
- 12: return squash(\mathbf{v}_j)

其中 softmax 函数的定义由公式(4)给出, Squashing 函数定义由公式(3)给出, Sigmoid 函数如下:

$$\text{Sigmoid}(\|\mathbf{s}_j\|) = 1 / (1 + e^{-\|\mathbf{s}_j\|}) \tag{7}$$

由于 CapsNet 网络最后是用向量的模的大小衡量某个结果的概率, 模的大小和概率大小直接相关, 模值越大, 相应概率也越大. Sigmoid 函数比 Squashing 函数收敛速度更快, 因此采用平移后的 Sigmoid 函数乘以 $\mathbf{s}_j / \|\mathbf{s}_j\|$ 在迭代环节替代 Squashing 函数能在结果接近 0 或 1 时起到放大作用, 从而提高对某些特定种类图片的识别准确率, 同时可以提高模型运行效率.

2.3 使用的 CapsNet 网络结构

由于原 CapsNet 仅支持 28×28 大小的图片输入, 为了适应数据集的图片尺寸, 提升模型性能, 将 CNN 网络与改进的 CapsNet 相结合, 得到一种新的强化胶囊网络(Enhanced-CapsNet)模型, 如图 4 所示.

模型开始部分有一层卷积核大小为 7×7 , 步长为 2 的卷积层, 具有 ReLU 非线性激活函数. 第二层也是卷积层, 卷积核大小为 6×6 , 步长为 1. 之后连接一个池化层, 池化核大小为 2×2 , 步长为 1, 采用最大池化. 池化层后再连接一个 6×6 步长为 1 的卷积层, 将数据重塑之后再输入 PrimaryCaps 层. 这样可以尽可能地获得图片的主要特征信息, 由于数据量较大, 所以引入了池化层进行数据降维, 提高模型性能和运行效率.

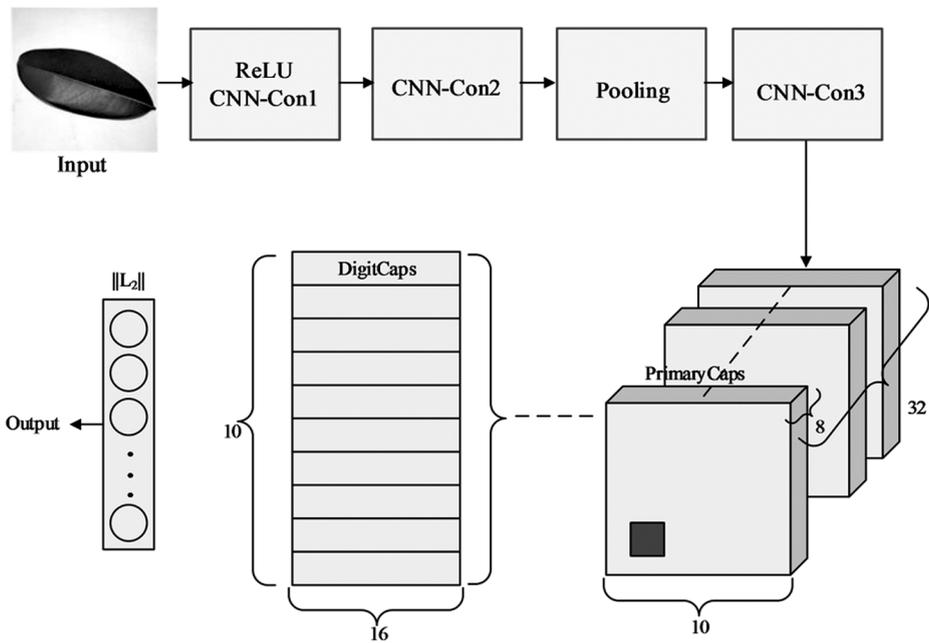


图 4 胶囊网络模型

PrimaryCaps 层: PrimaryCaps 中使用 9×9 大小的卷积核在 28×28 的空间中移动, 步长为 2, 通道数为 32, 设置为无填充. 这一层的作用类似常规的卷积层, 但最重要的作用是将每 8 个 $10 \times 10 \times 1 \times 32$ 的输出封装在一起, 形成一个长度为 8 的矢量, 为改进的 Routing 算法做准备. 在 PrimaryCaps 层最后得到的

输出形状为 $10 \times 10 \times 8 \times 32$, 这一层共输出 $3\,200(10 \times 10 \times 32)$ 个矢量, 也就是 $3\,200$ 个胶囊单元。

DigitCaps 层: 在 DigitCaps 中将长度为 8 的胶囊转换为长度 16 的胶囊. 由于使用的树叶数据集包含有 10 个类别, 因此 DigitCaps 中包含有 10 个长度为 16 的矢量, 这 10 个矢量分别代表不同的分类. 本层中使用了 Routing 算法并经过计算获得最终输出值向量 \mathbf{v}_j , 根据 \mathbf{v}_j 获得图像被识别为特定类别的概率.

整个模型参数设置如表 2 所示.

表 2 E-CapsNet 参数设置

Name	Input shape	Output shape
Input(输入层)	(None, 147, 147, 1)	(None, 147, 147, 1)
Con-1	(None, 147, 147, 1)	(None, 71, 71, 8)
Con-2	(None, 71, 71, 8)	(None, 66, 66, 16)
Maxpool	(None, 66, 66, 16)	(None, 33, 33, 16)
Con-3	(None, 33, 33, 16)	(None, 28, 28, 256)
Pcaps(Con2d)	(None, 28, 28, 256)	(None, 10, 10, 256)
Pcaps(Reshape)	(None, 10, 10, 256)	(None, 3200, 8)
Pcaps(Squash)	(None, 3200, 8)	(None, 3200, 8)
DigitCaps	(None, 3200, 8)	(None, 10, 16)
OutCaps	(None, 10, 16)	(None, 10)

2.4 实验参数设定

使实验设定批尺寸为 100, epoch 设为 50, 同时使用原始的 CapsNet 网络和经典网络 AlexNet 以及 Inception V3 深度学习模型进行对比.

2.5 实验环境

实验使用的计算机内存为 16GB, 搭载 Inter(R) Core(TM) i7-7700HQ CPU 处理器, 使用 NVIDIA Geforce GTX 1070 8G 显卡, 软件环境为 Ubuntu18.04 操作系统, 终端软件为 Pycharm, Python 版本为 3.6, 使用 Tensorflow+Keras 框, Cuda 版本为 10.0.130.

2.6 模型评价指标

由于实验是一个多分类任务(共 10 种类别), 所以采用准确率(Accuracy)评价 10 类图像的分类效果, 准确率用 η_{Accuracy} 表示, 计算公式如下:

$$\eta_{\text{Accuracy}} = (TP + TN) / (P + N) \quad (8)$$

其中 TP 为正类预测且结果为正类的个数, TN 为负类预测且结果为负类的个数, P 为所有正类预测个数, N 为所有负类预测个数.

3 结果与分析

使用 E-CapsNet 网络训练和分类树叶数据集的 15 965 张图像, 我们最终获得的结果如表 3 所示: 经过 50 次 epoch 的训练, 模型训练集准确率最高达到 99.15%, 损失率最低为 0.015 89, 验证集的准确率为 98.51%, 损失率为 0.020 49, 测试集准确率为 98.63%, 损失率和准确率分别由公式(6)和公式(8)得出.

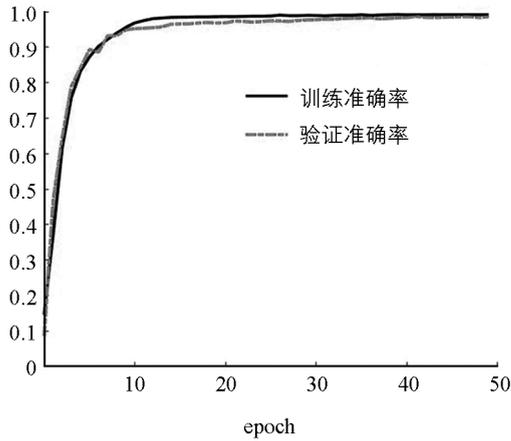
图 5 为模型训练图, 模型从 epoch12 左右开始趋近收敛, 收敛速度较快, 训练准确率和验证准确率曲线接近, 训练损失率和验证损失率都较为平滑, 模型拟合效果较好.

表 3 E-CapsNet 模型的分训练结果

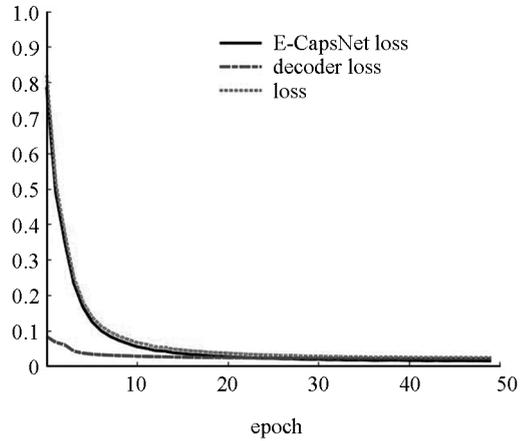
方 法	模型训练准确率/%	模型训练损失率/%	验证集准确率/%	验证集损失率/%	测试集准确率/%
E-CapsNet	99.15	0.01589	98.51	0.02049	98.63

本文在 AlexNet 和 Inception V3 模型上使用相同的树叶数据集进行测试, 采用与原 CapsNet 同样的学习率和迭代次数, E-CapsNet 模型的分精度比 AlexNet 要高 11.63%, 比 Inception V3 模型高 7.63%,

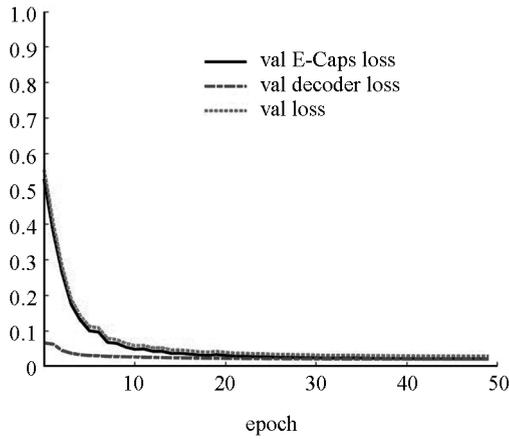
比原 CapsNet 网络高 2.51%。对比实验结果如表 4 所示。



(a) 训练准确率和验证准确率曲线



(b) 训练损失率



(c) 验证损失率

图 5 模型训练曲线

表 4 测试集准确率对比

方式	测试集准确率	方式	测试集准确率
AlexNet	87.0	CapsNet	96.12
Inception V3	91.0	E-CapsNet	98.63

为了测试动态路由算法的改进对模型精度的影响, 本文分别在保证其他参数相同的情况下, 使用 E-CapsNet 模型, 将原 Rounting 算法、使用 Adam 算法但不改动压缩函数和本文方法进行对比. 测试结果如表 5 所示.

表 5 在动态路由环节使用不同优化方法的准确率对比

方法	训练集准确率	测试集准确率
原动态路由算法	96.32%	95.86%
使用 Adam 优化	98.65%	97.47%
本文方法	99.15%	98.63%

由表 5 可见, 原动态路由算法训练集准确率为 96.32%, 测试集准确率为 95.86%. 仅使用 Adam 算法进行优化时, 训练集准确率为 98.65%, 测试集准确率为 97.47%. 使用 Adam 进行优化, 并且改进压缩函数部分后, 训练集准确率为 99.15%, 测试集准确率为 98.63%. 由图 6 可见, 使用了 Adam 优化算法后, 训练准确率曲线收敛相对平滑, 拟合度更好, 收敛速度也略微增加; 同时, 在优化了动态路由环节的压缩

函数后,训练准确率和验证准确率更高.可以看出,本文的优化可以取得更好的分类效果.

4 结 论

在本文中,我们采用改进的 CapsNet 神经网络对树叶进行分类实验,将 CNN 结构与 CapsNet 网络结合,更好地适应了数据集图像尺寸,同时提高了模型性能,克服了传统人工卷积神经网络分类器的某些局限性,并使用数据增强来减少模型的过拟合,提高了模型的泛化能力.在优化了 Rounting 算法和 Squashing 函数环节后,模型收敛更加平滑,收敛速度更快,分类识别准确率也有提升,实现了比原结构更好的性能.同时,我们还设计了对比实验,实验结果表明,本文训练出的用于树叶分类的 E-CapsNet 网络模型的测试准确率达到 98.63%,明显高于 AlexNet 和 InceptionV3 模型和原 CapsNet.

CapsNet 目前仍处在起步阶段,在以后的研究工作中,把重点放在优化内部结构或者与已有的成熟网络模型相结合上,可能会在优化模型性能上更有发展.

参考文献:

- [1] 刘春爽. 基于植物叶形状和叶脉的植物叶自动分类研究 [D]. 杭州: 浙江理工大学, 2016.
- [2] 恩 德, 忽胜强. 基于集成神经网络的植物叶片识别方法 [J]. 浙江农业学报, 2015, 27(12): 2225-2233.
- [3] 于业达, 顾偌铤, 唐运林, 等. 基于深度学习的草地贪夜蛾自动识别 [J]. 西南大学学报(自然科学版), 2019, 41(9): 24-31.
- [4] LECUN Y, BENGIO Y. Convolutional Networks for Images, Speech, and Time Series [J]. Handbook of Brain Theory & Neural Networks, 1995, 3361(10): 1995.
- [5] 杨 珍, 王 俊, 辛春花. 利用 DCE-MRI 结合改进卷积神经网络的 MR 图像自动分割与分类方法 [J]. 重庆理工大学学报(自然科学), 2020, 34(2): 147-157.
- [6] 崔少国, 熊舒羽, 刘 畅, 等. 基于深度哈希卷积神经网络的医学图像检索 [J]. 重庆理工大学学报(自然科学), 2020, 34(8): 134-142.
- [7] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet Classification with Deep Convolutional Neural Networks [J]. Communications of the ACM, 2017, 60(6): 84-90.
- [8] ZEILER M D, FERGUS R. Visualizing and Understanding Convolutional Networks [C] //Computer Vision-ECCV 2014, 2014: 818-833. DOI:10. 1007/978-3-319-10590-1_53.
- [9] SIMONYAN K, ZISSERMAN A. Very Deep Convolutional Networks for Large-Scale Image Recognition [EB/OL]. (2014-09-01) [2020-09-15]. https://www.researchgate.net/publication/319770291_Very_Deep_Convolutional_Networks_for_Large-Scale_Image_Recognition.
- [10] SZEGEDY C, LIU W, JIA Y Q, et al. Going Deeper with Convolutions [C] //2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 7-12, 2015, Boston, MA, USA. IEEE, 2015: 1-9.
- [11] HE K, ZHANG X Y, REN S Q, et al. Deep Residual Learning for image Recognition [C] //2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 770-778.
- [12] GRINBLAT G L, UZAL L C, LARESE M G, et al. Deep Learning for Plant Identification Using Vein Morphological Patterns [J]. Computers and Electronics in Agriculture, 2016, 127: 418-424.
- [13] SLADOJEVIC S, ARSENOVIC M, ANDERLA A, et al. Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification [J]. Computational Intelligence and Neuroscience, 2016, 2016: 3289801.
- [14] 张 帅, 淮永建. 基于分层卷积深度学习系统的植物叶片识别研究 [J]. 北京林业大学学报, 2016, 38(9): 108-115.
- [15] CRUZ A C, LUVISI A, DE BELLIS L, et al. Vision-Based Plant Disease Detection System Using Transfer and Deep

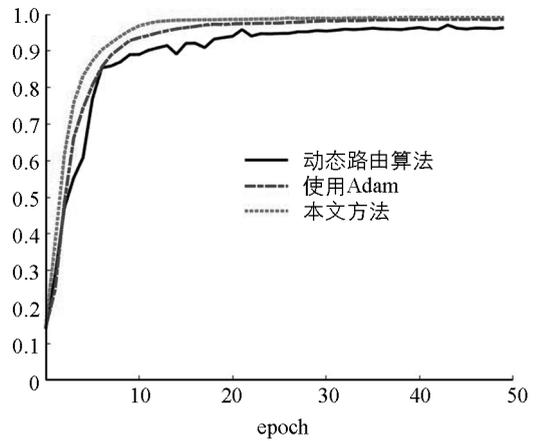


图 6 不同优化方法的训练准确率曲线

- Learning [C] //2017 Spokane, Washington July 16-July 19, 2017. St. Joseph, MI: American Society of Agricultural and Biological Engineers, 2017: 1.
- [16] BARBEDO J G A. Plant Disease Identification from Individual Lesions and Spots Using Deep Learning [J]. *Biosystems Engineering*, 2019, 180: 96-107.
- [17] SABOUR S, FROSST N, HINTON G E. Dynamic Routing Between Capsules [EB/OL]. (2020-09-22) [2020-10-15]. <https://www.cnblogs.com/MTandHJ/p/13659001.html>.
- [18] ZHANG X Q, ZHAO S G. Fluorescence Microscopy Image Classification of 2D HeLa Cells Based on the CapsNet Neural Network [J]. *Medical & Biological Engineering & Computing*, 2019, 57(6): 1187-1198.
- [19] MOBINY A, VAN NGUYEN H. Fast CapsNet for Lung Cancer Screening [M] //Medical Image Computing and Computer Assisted Intervention-MICCAI 2018. Cham: Springer International Publishing, 2018: 741-749.
- [20] 陈立潮, 郑佳敏, 曹建芳, 等. 基于胶囊网络的智能交通标志识别方法 [J]. *计算机应用*, 2020, 40(4): 1045-1049.

A Neural Network-Based Leaf Image Classification Model

ZHANG Dong-yan, HAN Rui, ZHANG Rui, CAO Jun

College of Mechanical and Electrical Engineering, Northeast Forestry University, Harbin 150040, China

Abstract: The basis of research of trees is to classify them. Combined with the CapsNet neural network model, 10 kinds of leaf pictures taken in the laboratory are used to establish a leaf classification model so as to improve the accuracy of leaf classification. Considering the efficiency and image size of the model, an E-CapsNet network model is obtained by combining the original CapsNet with the traditional convolutional neural network and optimizing the dynamic routing algorithm to improve the CapsNet. At the same time, the E-CapsNet network model is compared with the classical neural network model AlexNet and InceptionV3 model. After 50-epoch training, the highest accuracy of model training is 99.15%, the accuracy of verification set is 98.51%, and the accuracy of test set is 98.63%. Compared with that of the original CapsNet network, the accuracy of the test set is improved by 2.51%. The experimental results show that the improved E-CapsNet model achieves higher accuracy.

Key words: capsule network; neural network; image classification; leaf recognition; dynamic routing

责任编辑 汤振金