

基于图像深度特征的稻种识别方法研究

黄崇渝, 华云松

上海理工大学光电信息与计算机工程学院, 上海

收稿日期: 2022年11月8日; 录用日期: 2022年12月1日; 发布日期: 2022年12月16日

摘要

随着水稻品种不断增加, 如何识别稻种成为业内的一个难点, 机器视觉技术较人工识别稻种识别准确度和时间更具优势, 因而成为一种趋势。针对传统图像特征识别方法能力有限、识别精度低、速度慢的问题, 本文提出使用L2正则化和dropout的卷积神经网络(convolutional neural networks, CNN)将原来得到的线性不可分图像深度特征数据变为线性可分的数据, 通过支持向量机(support vector machine, SVM)代替传统CNN的softmax分类器对得到的线性数据进行分类, 最后利用10个常用CNN模型 + SVM, 通过准确性、敏感性、特异性、假阳性率、F1 Score和测试时间等方面对模型效果进行评估后, 得出水稻识别结果。实验表明, 该方法在稻种识别方面准确率达到99.8%, 每张图片识别速度低至0.07 s, 可用于混杂稻种的识别和分类。

关键词

稻种识别, 深度学习, 卷积神经网络, 支持向量机

Research on Rice Seed Recognition Method Based on Image Depth Feature

Chongyu Huang, Yunsong Hua

School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Nov. 8th, 2022; accepted: Dec. 1st, 2022; published: Dec. 16th, 2022

Abstract

With the increasing number of rice varieties, how to identify rice seeds has become a difficult point in the industry. Machine vision technology has become a trend as it has more advantages over manual identification of rice seeds in terms of accuracy and time. To address the problems of limited capability, low recognition accuracy and slow speed of traditional image feature recogni-

tion methods, in this paper, we propose a convolutional neural network (CNN) using L2 regularization and dropout to change the originally obtained linear non-separable image depth feature data into linear separable data by support vector machine (SVM) instead of traditional. Finally, the rice recognition results were obtained by evaluating the model effects in terms of accuracy, sensitivity, specificity, false positive rate, F1 Score and testing time using 10 commonly used CNN models + SVM. Experiments show that the method achieves 99.8% accuracy in rice seed recognition, and the recognition speed is as low as 0.07 s per image, which can be used for the recognition and classification of mixed rice seeds.

Keywords

Rice Seed Identification, Deep Learning, Convolutional Neural Networks, Support Vector Machines

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在谷类中, 稻米是世界大部分人口的主要食物, 是一种富含碳水化合物和淀粉的谷物。根据联合国农业组织提供的数据, 水稻每年的产量位于主要谷类前列[1]。稻种质量的好坏直接影响人的身体健康, 确保同一批次的稻种纯度是有必要的。由于市场上出现了品种越来越多的种子, 受过训练的专业人员也很难识别; 另外, 人工识别数量较多的稻种时, 也会花费较长时间并且容易出错。机器视觉、近红外光谱、高光谱技术凭借快速、准确、无损检测等优点目前已广泛应用于农作物识别。运用机器视觉和图像处理技术可以对稻种外形、颜色、纹理等物理特征进行提取实现分类。

国内外在稻种识别方面的研究现状: Zhengjun Qiu [2]利用两种不同光谱范围的高光谱成像技术对水稻种子品种进行鉴定, 分类准确率达到 87%; Fabiyi [3]选取 90 种稻种利用高分辨率图像的空间特征, 将其与高光谱数据的光谱特征融合, 分类准确率达到 78%; Castillo [4]基于树莓派开发了一个便携式稻种识别系统, 采用基于监督学习的卷积神经网络来实时识别稻种, 准确率达到 90%; Jin [5]把近红外光谱与深度学习结合, 使用 ResNet 对十种稻种进行分类, 准确率达到 86.08%; Huang [6]根据三种稻种的形态特征, 建立反向传播神经网络(BPNN), 分类准确率达到 95.56%; Kuo [7]利用光学显微镜对 30 种稻种, 采用图像处理和基于稀疏表示的分类(SRC)进行识别, 准确率达到 89.1%; Sethy & Chatterjee [8]依据水稻几何纹理特征, 采用多类支持向量机进行分类, 准确率达到 92%; Kantip [9]分别采用支持向量机(SVM)和 InceptionResNetV2 的深度学习模型, 对 14 种 5000 个样本的水稻进行分类, 准确率达到 95.14%。

本文提出一种使用 CNN 提取图像深度特征并将原来线性不可分的数据变得线性可分后, 再使用对线性数据分类效果优秀的 SVM 进行分类的方法, 最后与传统的图像识别方法灰度共生矩阵(Gray Level Concurrence Matrix, GLCM)、局部二值模式(Local Binary Patterns, LBP)、方向梯度直方图(Histogram of Oriented Gradients, HOG)进行对比, 结果表明 EfficientNet_b0 + SVM 模型对稻种的分类在精度和速度上都有比较优秀的效果。

2. 理论方法

2.1. 卷积神经网络

CNN 凭借权值共享大大降低了 CNN 系统的复杂性, 减少了需要计算的参数量。卷积是对图像进行

过滤直到遍历整个图像的过程[10], 其卷积公式可以描述为:

$$y_j^l = f \left(\sum_{i \in m} x_i^{l-1} * k_{ij}^l + b_j^l \right) \quad (1)$$

$$f(x) = \max(0, x) \quad (2)$$

x_i^{l-1} 是 $l-1$ 层第 i 个输入特征, y_j^l 是 l 层第 j 个输出特征, m 是 $l-1$ 层的特征集合, k_{ij}^l 是介于在输入特征 i 第 $l-1$ 层和输出特征 j 第 l 层之间的卷积核, b_j^l 是输出特征 j 在 l 层的偏差。 $f(x)$ 为线性激活函数。

CNN 主要包含卷积层、池化层和全连接层[11]。在卷积层中输入图像有高度 H 、深度 D 、宽度 W , 在经过卷积层后宽度为:

$$W_1 = \frac{(W - F + 2p)}{S + 1} \quad (3)$$

其中 F 为卷积核的尺寸大小, p 为需要补零的数目, S 为步长。在经过卷积层后高度为:

$$H_1 = \frac{(H - F + 2p)}{S + 1} \quad (4)$$

D_1 为扩张的通道数。新的图像输出大小为 $W_1 \times H_1 \times D_1$ 。通常会在连续的卷积层中加入池化层, 其中池化层的作用是缩小前一层图像的尺寸, 从而减少计算量、内存使用量、参数个数。池化层有最大池化和平均池化两种方式, 本研究采用了平均池化的方式。在池化层中输入图像高度 H 、深度 D 、宽度 W , 在经过池化层后宽度为:

$$W_2 = \frac{W - F}{S + 1} \quad (5)$$

高度为:

$$H_2 = \frac{H - F}{S + 1} \quad (6)$$

深度 D 保持不变。最后经过全连接层整合图像的特征, 采用 softmax 进行分类。

在训练一个大型的卷积神经网络时, 训练数据不够多时易发生过拟合现象。所以本研究在代价函数方面上引入 L2 正则化防止模型过拟合, 降低泛化误差, 提高泛化能力。在梯度下降法中, 用于迭代计算参数 θ 的迭代式为:

$$\theta_j := \theta_j - \alpha \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \quad (7)$$

其中 α 是学习率, 上式是没有添加 L2 正则化项的迭代公式。在原代价函数之后添加 L2 正则化, 则迭代公式为:

$$\theta_j := \theta_j \left(1 - \alpha \frac{\lambda}{m} \right) - \alpha \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \quad (8)$$

其中 λ 是正则化参数。从公式得知每次迭代 θ_j 都会乘上一个小于 1 的因子, 使得 θ_j 不断减小。较大的 λ 将约束模型的复杂度, 通过减小权重消除部分隐藏单元的影响, 降低发生过拟合的可能性。

从修改神经网络结构上引入 dropout 提升网络的泛化能力。训练过程中设置 dropout_rate = 0.2, 即每一个输出节点以 20% 概率不工作。因为 dropout 的随机性, 每次 dropout 后, 网络模型都可以看成是一个不同结构的神经网络, 所以将多个不同神经网络的预测结果取平均得到输出结果。由于随机地让一些节点不工作, 因此可以避免某些特征只在固定组合下生效, 可以有意识地让网络去学习一些普遍的共性从

而提高模型泛化能力。

实验采用 VGG、GoogLeNet、ResNet、EfficientNet、DenseNet、MobileNet、ShuffleNet 网络进行测试。VGG 网络通过堆叠多个 3×3 的卷积层替代大尺寸的卷积层, 在拥有相同感受野的情况下能够减少所需参数[12]。GoogLeNet 引入 Inception 结构来融合不同尺寸的特征信息[13]。ResNet 该网络提出超深的网络概念, 引入残差模块, 使用 Batch Normalization (BN)加速训练[14]。EfficientNet_b0 网络提出同时增加网络的宽度、深度以及输入网络的分辨率来提升网络的性能[15]。DenseNet121 设计的是一种稠密卷积神经网络, 让每一层都接受它前面所有层的输出。具有缓解梯度消失问题, 加强特征传播, 鼓励特征复用, 减少计算量的优点[16]。

由于传统的神经网络内存需求大、运算量大导致无法在移动设备以及嵌入式设备上运行。轻量级 CNN 例如 MobileNet_v2 和 ShuffleNet_v2 被应用于移动端。MobileNet_v2 提出 Depthwise Convolution 大大减少运算量和参数数量[17]。ShuffleNet_v2 提出影响模型性能的不仅是 FLOPS, 内存的使用量不能忽略。提出同等通道大小可以最小化内存访问量, 过量使用组卷积会增加内存消耗, 网络碎片化会降低并行度, 不能忽略元素级操作。实验发现如果将 ResNet 中残差单元中的 ReLU 和 shortcut 移除的话, 速度有 20% 的提升[18]。

2.2. 支持向量机

支持向量机是在分类与回归分析中分析数据的监督式学习模型。给定一组训练实例, 每个训练实例被标记为属于两个类别中的一个或另一个, SVM 训练算法创建一个将新的实例分配给两个类别之一的模型, 使其成为非概率二元线性分类器。线性分类可以理解为在 2 维空间中, 可以通过一条直线来分类。在 p 维空间中, 可以通过一个 $p-1$ 维的超平面来分类。SVM 线性可分的示意图如图 1 所示。

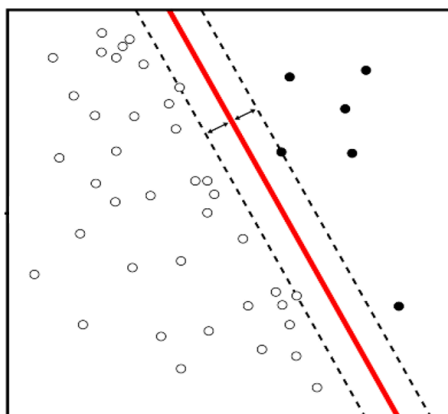


Figure 1. Linear separable diagram of SVM

图 1. SVM 线性可分示意图

SVM 模型是将实例表示为空间中的点, 这样映射就使得单独类别的实例被尽可能宽的明显的间隔分开。然后, 将新的实例映射到同一空间, 并基于它们落在间隔的哪一侧来预测所属类别。

2.3. SVM 代替 Softmax 分类器

通常 softmax 用作神经网络的最后一个激活函数, 以将网络的输出标准化为预测输出类的概率分布。Softmax 是 CNN 方法中用于分类的逻辑回归层, 它是基于回归的原则, LR 的损失函数也是得自于一个考虑了全局数据的自概率模型。逻辑回归的损失函数为:

$$a_0 = -\frac{1}{m} \sum_{i=1}^m [x_i \log(z_i) + (1-x_i) \log(1-z_i)] \tag{9}$$

m 为批次大小, x_i 为在批次中第 i 次的预测标签, z_i 为在批次中第 i 次的实际标签。

SVM 的分类精度都主要依赖于数据中位于分割平面的一部分支持向量。支持向量机的损失函数为:

$$b_0 = \min \left[\frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \max(0, 1 - y_i'(w^T x_i + b))^2 \right] \tag{10}$$

$\|w\|^2$ 为欧几里得范数, C 是惩罚参数, y_i' 是实际标签, $w^T x_i + b$ 是决策边界。

从损失函数得知 SVM 与 LR 很大的一个区别是: SVM 性能主要受支持向量的影响, 而 LR 考虑全局数据的同时分类超平面也会受到每一个样本的干扰。所以可以认为当数据的特征处于一个线性不可分状态的时候, SVM 与 LR 是从两个不同的角度去解决了同一个分类问题。

CNN 所连接的 LR 与单纯的 LR 的区别在于 CNN 在训练的同时就会趋于提取样本中 feature 的 signature 特征, 也就是说 CNN 的卷积过程在训练的过程中就是对一个线性不可分的数据增加其线性可分程度的过程[19]。当 CNN 的 feature 提取生效之后, 原本线性不可分的分类样本趋于线性可分的时候, SVM 只利用部分支持向量样本的分类原则的优势就能够得到进一步体现。将 CNN 的 softmax 分类器换成 SVM 分类器的示意图如图 2 所示:

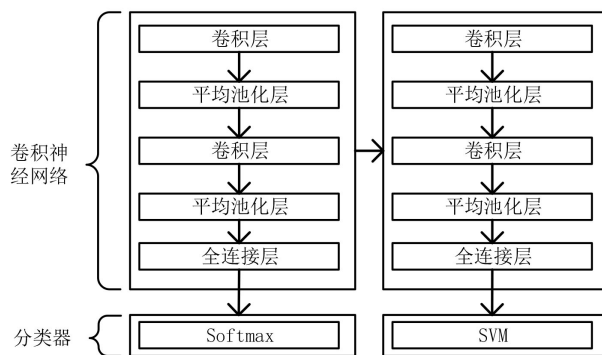


Figure 2. Schematic diagram of replacing the softmax classifier of CNN with SVM classifier

图 2. CNN 的 softmax 分类器换成 SVM 分类器的示意图

3. 数据集和实验流程

3.1. 数据集

样本数据集有 6 个类别: 郑早 10 号、丹早稻 53、绿早 1 号、绿银占、强两优 698、扬粳优 918。6 个类别中包含了粳型常规稻、籼型常规稻和籼型杂交稻等主要栽培类型。

采用一种改进的 Canny 边缘检测算法[20], 使用自适应平滑滤波代替高斯滤波, 在平滑图像的同时锐化了边缘。使用水平垂直、45°和 135°四个方向梯度模板计算图像梯度, 改善了传统 Canny 算法在计算梯度时对噪声的敏感性。对稻种数据集进行边缘检测, 图像分割, 使单个稻种成为一个样本。经过图像分割后得到 6307 个稻种样本。

根据数据增强的原理, 对训练集样本进行旋转、翻转来扩大训练集的样本量。通过数据增加, 数据集样本增加了一倍。对于每次训练和验证选择的样本都保证是随机的。样本数据集随机 80% 作为训练集和验证集, 20% 为测试集。

表 1 列举了实验稻种的名称与数量。

Table 1. Name and quantity of experimental rice varieties

表 1. 实验稻种的名称与数量

稻种类型	原始图像个数	图像增强个数	训练和验证的图像个数	测试的图像个数
丹早稻 53	1035	1035	1656	379
绿早 1 号	1079	1079	1727	387
强两优 698	1093	1093	1677	378
扬粳优 918	1056	1056	1675	375
绿银占	1020	1020	1632	368
郑早 10 号	1024	1024	1639	366
总和	6307	6307	10006	2253

3.2. 稻种识别流程

选择 SVM 对稻种进行分类, SVM 目的是最小化经验风险和置信风险, 以获得良好的样本统计规则, 提高机器学习的泛化能力。激活 CNN 的 GPU 批量大小为 32, 且 GPU 有足够内存空间来容纳图像数据集。SVM 使用 one-vs-all 编码设计。从 CNN 全连接层或平均池化层提取深度特征输入到 SVM 进行训练, 流程如图 3 所示。

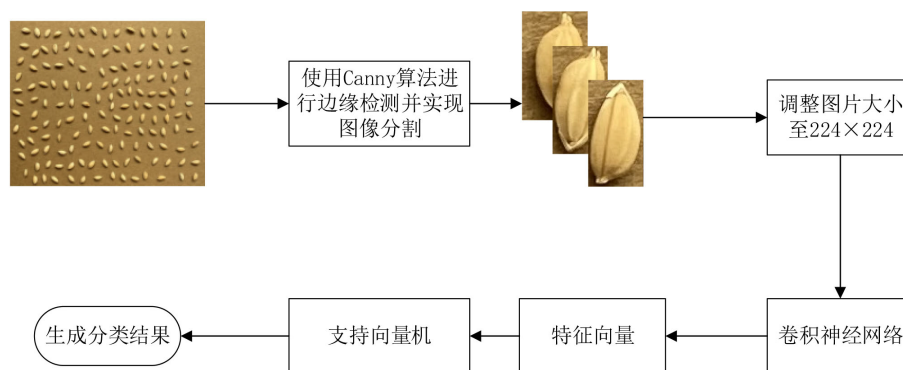


Figure 3. Flow chart of rice seed identification

图 3. 稻种识别流程图

4. 实验结果

4.1. 实验环境与参数配置

研究的实验环境为 Window10 64 位系统, 采用 Pytorch 深度学习开源框架, 选用 Python 作为编程语言。CPU 为 Intel@Core-i7-9750H, 并采用英伟达 RTX2060 显卡加速图像处理。

采用 10 种卷积神经网络对稻种特征进行提取。根据网络将图像尺寸调整为 224×224 像素, 网络优化器采用随机梯度下降 SGD, 动量设置为 0.9, 学习率为 0.001, 权值衰减设为 0.0001, 训练 50 个 epoch。对每个 epoch 中每个图像随机进行裁剪、缩放、旋转、归一化操作, 这样使神经网络可以更准确提取图像特征[21]从准确性、敏感性、特异性、假阳性率(FPR)、F1 Score 和训练时间等方面衡量各分类器的性能。

本实验采用 One-Vs-all 方法和线性核函数作为 SVM 分类器参数, 用网格搜索和交叉验证的方法确定参数 C。混淆矩阵参数表示为:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (11)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (12)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (13)$$

$$\text{FPR} = \frac{FP}{TN + FP} \quad (14)$$

$$F_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (15)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (16)$$

TP 为真阳性, FP 为假阳性, TN 为真阴性, FN 为假阴性, $(TP + FP + TN + FN)$ 为测试集所有样本数量。

4.2. 使用预训练的 CNN 模型对稻种进行识别

从头开始训练深度学习网络需要大量的数据集并且也是一个耗时的过程。在本次研究中使用了迁移学习技术, 初始化的权重来自预训练的模型, 该模型已经在著名的 ImageNet 数据集上训练过。用自己的目标数据集重新训练预训练的权重, 根据要分类的类型更新权重, 同时调整学习参数以适应研究目的并减少训练时间。

使用 Vgg16、Vgg19、GoogLeNet、Resnet18、Resnet50、Resnet101、Densenet121、EfficientNet_b0、Mobilenet_v2、Shufflenet_v2 10 种常用的卷积模型。

只使用预训练的 CNN 对稻种进行识别, 识别准确率如表 2 所示。

Table 2. Identifying rice seeds by the pre-trained CNN model
表 2. 使用预训练 CNN 模型对稻种进行识别

分类模型	Accuracy
VGG16	0.900143
VGG19	0.915241
GoogLeNet	0.954113
ResNet18	0.956231
ResNet50	0.961324
ResNet101	0.963241
DenseNet121	0.964231
EfficientNet_b0	0.969121
MobileNet_v2	0.848231
ShuffleNet_v2	0.917468

4.3. 使用预训练的 CNN 模型对稻种进行识别

CNN + SVM 在训练集上的性能如表 3 所示。通过对比表 2 和表 3 的结果, 可以得知利用 CNN 提取特征后用 SVM 进行分类可以提升对稻种识别的准确率。

表 3 列举了各个 CNN + SVM 模型的测试性能。从表 3 中得知 EfficientNet_b0 和 Densenet121 准确率最高, 性能最好。Resnet 和 GoogLeNet 网络性能较好, Vgg 网络和轻量级网络 Mobilenet_v2 和 Shufflenet_v2 性能比较一般。EfficientNet_b0 凭借比 Densenet121 有着更短的测试时间, 作为本研究最终使用模型, 该模型在测试集上的识别结果如表 4 所示。

Table 3. Depth features of various CNN models were used for system analysis based on SVM

表 3. 利用各种 CNN 模型的深度特征基于 SVM 进行系统分析

分类模型	Accuracy	Recall	Specificity	FPR	F1 Score	Testing Time
VGG16	0.921851	0.921852	0.961459	0.014216	0.921852	482.572134
VGG19	0.934197	0.934196	0.972811	0.013236	0.934196	579.018826
GoogLeNet	0.984818	0.984818	0.996965	0.003034	0.984796	116.845401
ResNet18	0.987614	0.987615	0.997521	0.002478	0.987614	89.087674
ResNet50	0.988413	0.988414	0.997682	0.002317	0.988418	252.264768
ResNet101	0.990411	0.990412	0.998085	0.001914	0.990410	406.727547
DenseNet121	0.998401	0.998402	0.999680	0.000319	0.998401	239.412328
EfficientNet_b0	0.998401	0.998402	0.999681	0.000318	0.998402	154.782743
MobileNet_v2	0.920894	0.920895	0.984175	0.015824	0.920979	116.050966
ShuffleNet_v2	0.935277	0.935278	0.987059	0.012940	0.935206	60.194926

Table 4. Identification results based on EfficientNet_b0+SVM model on the test set

表 4. 基于 EfficientNet_b0+SVM 模型在测试集上的识别结果

稻种类型	丹早稻 53	绿早 1 号	强两优 698	扬粳优 918	绿银占	郑早 10 号
Accuracy	0.984168	0.992248	0.989417	0.992000	1.000000	1.000000
Recall	0.984169	0.992248	0.989418	0.992000	1.000000	1.000000
Specificity	0.998399	0.998392	1.000000	1.000000	0.997347	0.997880
FPR	0.001600	0.001607	0.000000	0.000000	0.002652	0.002119
F1Score	0.988079	0.990968	0.994681	0.995984	0.993252	0.994565

在传统的图像处理和机器学习的方法中, 采用 GLCM + SVM、LBP + SVM、HOG + SVM 进行图像分类。这些方法的 F1 Score 如表 5 所示。

Table 5. F1 Score for traditional image recognition method

表 5. 传统图像识别方法的 F1 Score

方法	GLCM + SVM	LBP + SVM	HOG + SVM
F1 Score	0.425894	0.875781	0.671626

从图 4 EfficientNet_b0 + SVM 在测试集上的混淆矩阵中可以看出, 模型对丹早稻 53 和强两优 698 识别效果较为一般, 绿早 1 号和扬粳优 918 识别效果良好, 绿银占和郑早 10 号识别效果很好达到 100%。

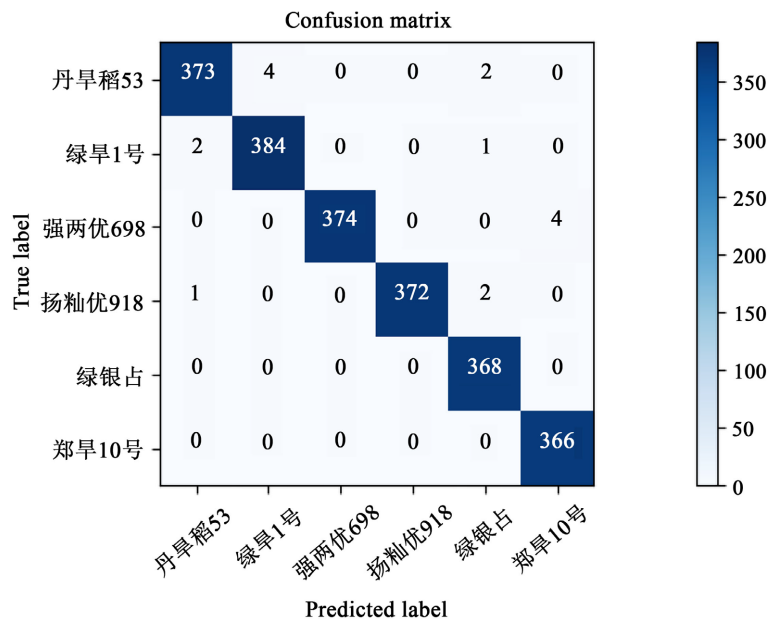


Figure 4. EfficientNet_B0 + SVM Confusion matrix on test set
 图 4. EfficientNet_b0 + SVM 在测试集上的混淆矩阵

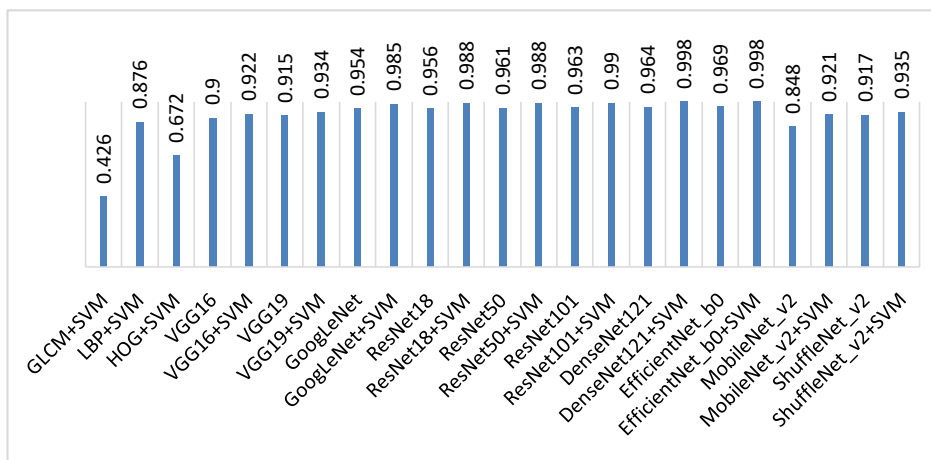


Figure 5. F1 score for all test models
 图 5. 所有测试模型的 F1 score

5. 结论

本文提出了一种使用卷积神经网络提取图像深度特征并将原来线性不可分的数据变得线性可分后,再使用对线性数据分类效果优秀的支持向量机进行分类的方法。通过实验数据得知 EfficientNet_b0 + SVM 对稻种识别效果最好,识别准确率达到 99.8%, 每张图片识别速度为 0.07 s, 可在保持高识别准确率的条件下保持较快的识别速度。根据图 5 可以发现, 在 CNN 的基础上加上 SVM 可以不同程度地提高模型识别性能, CNN+SVM 相比传统图像识别准确率有着明显的提高。

参考文献

[1] FAO of the United Nations (2022) Faostat Database. <http://www.fao.org/faostat>

- [2] Qiu, Z.J., Chen, J., Zhao, Y.Y., *et al.* (2018) Variety Identification of Single Rice Seed Using Hyperspectral Imaging Combined with Convolutional Neural Network. *Applied Sciences—Basel*, **8**, 212. <https://doi.org/10.3390/app8020212>
- [3] Fabiyi, S.D., Vu, H., Tachtatzis, C., *et al.* (2020) Varietal Classification of Rice Seeds Using RGB and Hyperspectral Images. *IEEE Access*, **8**, 22493-22505. <https://doi.org/10.1109/ACCESS.2020.2969847>
- [4] Castillo, L.J.L., Galindo, J.A.M. and Rosal, J.E.C. (2020) A Supervised Learning Approach on Rice Variety Classification Using Convolutional Neural Networks. Association for Computing Machinery, Seoul, 18-23.
- [5] Jin, B.C., Zhang, C., Jia, L.Q., *et al.* (2022) Identification of Rice Seed Varieties Based on Near-Infrared Hyperspectral Imaging Technology Combined with Deep Learning. *ACS Omega*, **7**, 4735-4749. <https://doi.org/10.1021/acsomega.1c04102>
- [6] Huang, K.Y. and Chien, M.C. (2017) A Novel Method of Identifying Paddy Seed Varieties. *Sensors*, **17**, 809. <https://doi.org/10.3390/s17040809>
- [7] Kuo, T.Y., Chung, C.L., Chen, S.Y., *et al.* (2016) Identifying Rice Grains Using Image Analysis and Sparse Representation Based Classification. *Computers and Electronics in Agriculture*, **127**, 716-725. <https://doi.org/10.1016/j.compag.2016.07.020>
- [8] Sethy, P.K. and Chatterjee, A. (2018) Rice Variety Identification of Western Odisha Based on Geometrical and Texture Feature. *International Journal of Applied Engineering Research*, **13**, 35-39.
- [9] Kiratiratanapruk, K., Temniranrat, P., Sinthupinyo, W., *et al.* (2020) Development of Paddy Rice Seed Classification Process Using Machine Learning Techniques for Automatic Grading Machine. *Journal of Sensors*, **2020**, Article ID: 7041310. <https://doi.org/10.1155/2020/7041310>
- [10] Guo, S., Chen, S., Li, Y., *et al.* (2016) Face Recognition Based on Convolutional Neural Network and Support Vector Machine. *IEEE International Conference on Information and Automation (ICIA)*, Ningbo, 1-3 August 2016, 1787-1792. <https://doi.org/10.1109/ICInfA.2016.7832107>
- [11] Wiatowski, T. and Bolcskei, H. (2018) A Mathematical Theory of Deep Convolutional Neural Networks for Feature Extraction. *IEEE Transactions on Information Theory*, **64**, 1845-1866. <https://doi.org/10.1109/TIT.2017.2776228>
- [12] Simonyan, K. and Zisserman, A. (2015) Very Deep Convolutional Networks for Large-Scale Image Recognition. *International Conference on Learning Representations (ICLR)*, San Diego, 7-9 May 2015, 1-14.
- [13] Szegedy, C., Liu, W., Jia, Y., *et al.* (2015) Going Deeper with Convolutions. *The Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, 7-12 June 2015, 1-9. <https://doi.org/10.1109/CVPR.2015.7298594>
- [14] He, K., Zhang, X., Ren, S., *et al.* (2016) Deep Residual Learning for Image Recognition. *The Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, 27-30 June 2016, 770-778. <https://doi.org/10.1109/CVPR.2016.90>
- [15] Tan, M. and Le, Q. (2019) EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *Proceedings of the International Conference on Machine Learning, PMLR*, Volume 97, 6105-6144.
- [16] Huang, G., Liu, Z., Maaten, L.V.D., *et al.* (2017) Densely Connected Convolutional Networks. *The Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, 21-26 July 2017, 4700-4708. <https://doi.org/10.1109/CVPR.2017.243>
- [17] Sandler, M., Howard, A., Zhu, M., *et al.* (2018) Mobilenetv2: Inverted Residuals and Linear Bottlenecks. *The Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-22 June 2018, 4510-4520. <https://doi.org/10.1109/CVPR.2018.00474>
- [18] Ma, N., Zhang, X., Zheng, H.T., *et al.* (2018) Shufflenet v2: Practical Guidelines for Efficient CNN Architecture Design. *The Proceedings of the European Conference on Computer Vision (ECCV)*, Munich, 8-14 September 2018, 116-131. https://doi.org/10.1007/978-3-030-01264-9_8
- [19] Chen, Y.S., Jiang, H.L., Li, C.Y., *et al.* (2016) Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks. *IEEE Transactions on Geoscience and Remote Sensing*, **54**, 6232-6251. <https://doi.org/10.1109/TGRS.2016.2584107>
- [20] 许宏科, 秦严严, 陈会茹. 一种基于改进 Canny 的边缘检测算法[J]. *红外技术*, 2014, 36(3): 210-214.
- [21] Cubuk, E.D., Zoph, B., Mane, D., *et al.* (2019) Autoaugment: Learning Augmentation Strategies from Data. *Proceedings of the 32nd IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 16-21 June 2019, 113-123. <https://doi.org/10.1109/CVPR.2019.00020>