

基于卷积神经网络的手写数字识别研究与设计

朱姣蓉

(上海大学文化遗产与信息管理学院, 上海 201900)

摘要: 在信息化时代, 手写数字识别在计算机视觉和模式识别中具有广泛应用。为了提升手写数字识别的精度和效率, 设计并测试了四种基于卷积神经网络(CNN)的算法模型。通过对 MNIST 数据集的实证研究, 比较不同模型的训练效果。实验结果表明, 多层卷积神经网络模型表现最优, 准确率达到 98.9%, 且每轮训练时间仅需约 20 s。研究表明, 增加卷积层数和选择高阶 API(application programming interface)有助于提升识别精度, 进一步推动 CNN 在手写数字识别中的应用性能。这一结果为高效、准确的手写数字识别模型设计提供了新的思路和方法。

关键词: 卷积神经网络; 手写数字识别; MNIST 数据集; 深度学习

中图分类号: TP391.41; TP183 **文献标志码:** A **文章编号:** 1671-1807(2025)10-0069-08

随着现代科技的不断发展, 人工智能成为一个备受关注的领域。在人工智能的诸多应用领域中, 图像识别一直是备受关注的研究方向之一。手写识别作为一种常见的静态图像识别, 广泛应用于数据统计、财务金融及成绩判定等多个领域^[1]。在这些领域中, 使用最频繁且最便利人们生活的便是手写数字识别(handwritten digit recognition), 它属于光学字符识别技术的一个分支, 是模式识别学科的传统研究领域。

深度学习作为人工智能实现的重要技术, 它的发展是时代的大势所趋。其中, 卷积神经网络(convolutional neural network, CNN)是一类包含卷积计算且具有深度解构的前馈神经网络, 是深度学习的代表算法之一^[2], 具有自动学习特征的能力, 其应用涉及制造业、农业等多个领域, 用于完成图像分类、目标检测等任务。

近年来, 卷积神经网络技术在图像识别领域广泛应用, 各种算法层出不穷, 以 LeNet-5 为起点, 学者们对模型不断优化, 衍生出 AlexNet、VGG、GoogLeNet 等模型^[3]。随着卷积神经网络技术的发展, 利用卷积神经网络的各种框架对图像进行分割与识别, 可以更好地对模型训练, 实现对手写静态图像的识别, 尤其是在手写数字识别中表现出良好的性能。

本文旨在研究和设计几种基于卷积神经网络的用以实现手写数字识别的不同模型, 并对其进行

实证研究。具体而言, 使用 Python 语言, 通过对手写数字数据集 MNIST 进行预处理和特征提取, 完成测试和训练, 再利用卷积神经网络的各种不同算法构建模型框架、进行特征学习, 形成可视化的结果图用以比较分析算法优劣。最终实现准确性高、损失率低的手写数字识别。

因此, 本文的实现从研究意义上说, 许多机器学习和模式识别领域的新理论和算法都是先用手写数字识别验证有效性的, 便于研究与验证。从社会意义上来说, 本文的完成可以帮助人们实现对手写数字的自动识别, 便利人们使用电子设备的输入; 从经济意义上来说, 本文的完成可以减少手工输入的工作量, 提高工作效率, 节约人工成本。

1 文献综述

卷积神经网络(CNN)在手写数字识别领域的应用始于 1989 年, 当时 LeCun 等^[4]首次提出反向传播算法在手写邮政编码识别中的应用, 为卷积神经网络的发展奠定基础。此后, LeCun 及其团队在 1998 年推出具有里程碑意义的 LeNet-5 模型^[5], 为 CNN 的研究开辟了新篇章。随着深度学习的兴起, Hinton 和 Salakhutdinov^[6]于 2006 年在 *Science* 上发表了关于深层神经网络的开创性论文, 进一步推动了卷积神经网络模型的优化与发展。这一研究不仅推动了 CNN 模型的多层次结构设计, 也开启了深度学习在农业^[7]、工业^[8]、地质^[9]等图像识别任务中的广泛应用, 尤其是在手写数字识别领域, 卷

收稿日期: 2024-11-21

作者简介: 朱姣蓉(2000—), 女, 浙江湖州人, 硕士研究生, 研究方向为深度学习。

积神经网络的优越性能得到了充分验证^[10-11]。

然而,在中国卷积神经网络的研究热潮直到 2004 年雷国伟等^[12]提出基于 CNN 的图像特征提取方法时才真正起步。此后,随着深度学习理论的逐步普及,中国学者也开始专注于卷积神经网络模型的创新与优化,尤其是在手写数字识别中的应用。例如,樊继慧等^[13]通过改进 Sigmoid 激活函数提升了卷积神经网络在手写数字数据集上的分类精度,谢东阳等^[14]基于 AlexNet 模型引入 Inception-ResNet 模块,显著提高了数字特征的提取能力。此外,陈群贤^[15]在 TensorFlow 框架下应用 Keras 接口,构建了一种具有更高运行效率和鲁棒性的手写数字识别系统。这些研究成果不仅为 CNN 模型的发展提供了理论基础,也为实践中各类数字识别任务提供技术支持。

国外的研究也呈现类似的趋势,Sulaiman 等^[16]提出一种高效的离线字符识别 OCR 系统,并在多种手写数字识别场景中取得了显著成果。同时,Zhou 和 Yu^[17]对 LeNet-5 进行改进,加入局部二值模式(LBP)图像信息,进一步提高模型的准确性,达到 99.8% 的识别率。全球科技公司如 IBM、HP、东芝等也在这一领域投入大量的资源,不仅在手写数字识别的基础研究中取得突破,也推动了相关技术在智能医疗、自动驾驶和人脸识别等领域的应用。

尽管已有诸多研究取得了重要进展,但仍存在一些局限。首先,目前的卷积神经网络模型在手写数字识别中的准确性仍未达到 100%,因此模型的优化和迭代仍是一个重要的研究方向。其次,尽管新兴模块和网络架构的引入确实提升了模型的性能,但这一过程往往忽视了模型本身结构优化的潜力。在快速发展的背景下,学者们更多地专注引入新的模块和技术,却较少关注如何在现有结构上进行有效的优化,以提升模型的泛化能力和稳定性。这使得卷积神经网络在处理复杂、噪声较多的手写数字图像时仍面临一定的挑战。

本文正是在这一背景下展开的,通过对现有卷积神经网络结构的分析与优化,尤其是应用程序编程接口(application programming interface, API)选择和卷积层数的调节,旨在深入探讨不同设计对模型性能的影响,并在此基础上提出新的改进方向,从而提升模型在手写数字识别任务中的应用效果。

2 方法

2.1 数据采集与预处理

本文以在手写体数字识别系统性能测试方面

具有较高认可度的 MNIST 数据库为实验样本集,其中包含 7 万张图片,6 万张为训练样本,1 万个为测试样本,用于对模型的训练及验证。在 NIST 中,手写数字 MNIST 是目前被最多人使用且数据最多的数据库,经常被用在深度学习中^[18]。采集步骤如下。

(1)原始图像收集:使用手写数字的原始图像,这些图像可能来自多种来源。

(2)图像标注:手动标注每个图像的正确数字标签,以便用于监督学习。

(3)数据集划分:将整个数据集分为训练集、验证集和测试集,确保模型能够在不同的数据集上进行评估。

数据预处理是确保数据适合模型训练的关键步骤。本文采取以下预处理步骤。

(1)图像归一化:将像素值缩放到特定的范围,通常是 $[0, 1]$ 。这有助于提高模型训练的稳定性。

(2)图像标准化:对图像进行标准化处理,使其具有零均值和单位方差。这有助于加速模型的收敛。

(3)图像大小调整:将图像调整为模型输入的固定大小,通常是 28×28 像素。这有助于确保所有输入数据具有相同的格式。

(4)标签编码:将数字标签进行独热编码(one-hot encoding),将其转换为模型能够理解的格式。

2.2 卷积神经网络设计

2.2.1 网络架构选择

卷积神经网络是感知器(MLP)的一种扩展和扩展,是一种基于监督学习的多层神经网络模型^[19]。它包括输入层、隐藏层(卷积层、池化层)和输出层,其中,隐藏层可以有多个,每一层都包含一个或更多的二维平面,每一层又包含若干个单独的神经元。

卷积神经网络中最关键的两个步骤是卷积和池化运算,前一步的目的是从这一层的图像样本中抽取特征,并去噪;后一步的目的是对前一层的特征矢量进行降维(降低图像分辨率,以获取有效的信息)。由此,可以得出结论,卷积神经网络是一种由两个不同的层间交替构造而成的模型。卷积神经网络的导出层通常选用线性全联结,当下最为常见的应当是 Softmax 划分类别的方式。

2.2.2 卷积层、池化层配置

卷积神经网络是以多个卷积核为基础,对不同的特征进行卷积运算,并对其进行训练,并由此获

得对应的特征表达式。通过将上一层的多个输出图像(上一层是输入图像,下一层是特征图)通过学习得到的卷积图,在特征图上“滑动”,每滑动一次,即可得到下一层的特征图,并与上一层的特征图形成内在联系。每个卷积核可以抽取一种类型的特征,通过 n 个卷积核可以获得 n 个特征图,也就是表示已经抽取 n 种类型的特征,之后再加上偏移 b ,就可以得到卷积的结果^[16]。卷积层的计算公式为

$$a_j^{(l)} = \sum_{i \in M_j} x_j^{(l-1)} * k_{ij}^{(l)} + b_j^{(l)} \quad (1)$$

$$x_j^{(l)} = f(a_j^{(l)}) \quad (2)$$

式中: $*$ 为卷积云运算符号; $a_j^{(l)}$ 为第 l 层第 j 个神经元节点的输入; $x_j^{(l)}$ 为第 l 层卷积后第 j 个神经元的输出; $x_j^{(l-1)}$ 为第 $l-1$ 层的第 j 个神经元的输出; $k_j^{(l)}$ 为第 l 层卷积核; $b_j^{(l)}$ 为第 l 层第 j 个神经元的偏置; $f(\cdot)$ 为非线性函数; M_j 为第 j 个神经元的局部感受野。

池化的目的是使图像具有某种空间不变性,以降低图像分辨率。在此基础上,提出一种新的图像处理算法,该算法将特征图与之相应卷积层的特征图相关联,并对特征图进行采集。可以用以下公式来计算池化层:

$$x_j^{(l)} = f[\beta_j^{(l)} \text{down}(x_j^{(l-1)}) + b_j^{(l)}] \quad (3)$$

式中: β 为不同的 Feature Map 的系数; $\text{down}(\cdot)$ 为一个下采样函数,如果下采样因子是 n ,输出的图像分辨率就变为原来的 $1/n$ 。

2.2.3 激活函数

在神经网络中,激活函数(activation function)是一类重要的映射函数。人工神经网络的计算模式是每一层神经网络的特征值和权值相乘,从本质上来讲,就是一种矩阵的乘法运算。这便意味着即使加深了神经网络的层数,仍然是多次的矩阵相乘,层与层之间也是线性相关的^[20]。因此,该神经网络从本质上来讲,仍然是一个原始的多层感知器模型,对某些非线性函数具有很强的拟合能力。为了确保卷积神经网络可以对非凸函数进行拟合,需要采用非线性激活函数,在每一层的操作完成之后,经过激活函数的映射后,输出计算结果。因此,要对卷积结果使用激活函数,来获得新的特征图。

Sigmoid(S型非线性函数)作为最早的激活函数,它的具体实现是将输出实数规整到 $0 \sim 1$ 使输出幅度有界(取值范围为 $[0, 1]$),此函数使网络具备对非线性信息处理的能力,一直被广泛地使用,

其数学表达式为

$$f(x) = \frac{1}{1 + e^{-x}} \quad (4)$$

为解决这些传统函数存在的梯度弥散问题,提出使用非饱和修正线性 ReLU 作为激活函数。ReLU 在近些年的研究中表现出色。ReLU 是一个分段函数,它可以使数据特征最大限度得以保留,具体方法是:将为正的所有数据全部保留,其他部分清零,其实数学表达式为

$$f(x) = \begin{cases} \max(0, x), & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (5)$$

2.3 模型训练

2.3.1 优化

为防止卷积神经网络在学习过程中出现局部极值,需对其学习算法进行优化。该方法将学习过程划分为两个阶段。一是信息从底层传递到高层的正向传播;二是在当前向传播得出的结果与预期不一致时,从高层面向低层面传播训练的反向传播^[21]。训练过程如下。

(1)对加权网络进行初始化。

(2)通过卷积层、池化层和全连接层,将输入的数据向前传递,获得输出数值。

(3)对上述网络目标值与所述输出值之间的误差进行计算。

(4)返回基于以上得到的结果并进行分析,如果错误误差与我们的预期一致,则停止训练。

(5)根据求得误差进行权值更新,然后再进入到第二步。

尽管随机梯度下降方法对传统的梯度降阶方法进行一定的改进,但是该方法仍存在着易陷入局部极值和易出现振荡等问题。为了对此进行改善,提出动量法,这是一种在物理学中模仿动量的概念,用积累以前的动量来代替真正的梯度^[22],参数更新公式如下:

$$v_t = w_t + \eta \Delta_{\theta} J(\theta) \quad (6)$$

$$\theta = \theta - v_t \quad (7)$$

式中: v_t 为表示时刻的动量值; Δ_{θ} 为学习率; η 为动量因子。

通过对基于动量的随机梯度下降法的计算公式进行分析,发现当梯度的方向和动量相同时,调整参数的振幅将增大,从而加快收敛速度;反之,调整参数的振幅会降低,从而降低振荡的发生。若在局部极值情形中,则可能出现由动能累积而跳出极值情形。

2.3.2 学习率

模型训练的另一个困难的问题是如何设定学习率^[23]。学习率的设定对整体模型的效果和训练效率有直接影响。较大的学习率,会造成参数更新的幅度较大、训练的时间较短,更容易陷入局部最优,或在最优范围内震荡;较低的学习率,意味着时间越长,所需的计算资源越多。所以,如何选择合适的学习量,是一个相当困难的问题。由于损失通常对参数空间中的某个维度或者其他维度非常敏感,因此,理论上可以在敏感度更高的情形下选择更低的学习率,或者在训练速度更快的情形下选择更高的学习率。

3 实证研究

本文使用深度学习框架 TensorFlow,它提供了很多 API,可以被简单地分为高低两个级别,因此本文用于对比的四个模型的区别之一就在于高阶和低阶 API 的选择;理论上来说,多层卷积可以更有效地更新权重来拟合训练数据,使得准确率更高,因此本文四个模型的区别之二在于针对卷积层的数量设计了单层、双层、三层卷积进行识别性能对比。

3.1 单层卷积神经网络低阶 API 的手写数字识别实现

第一个模型利用单层卷积神经网络低阶 API 实现手写数字识别。低阶 API 主要包括各种张量操作、计算图和自动微分,如在该模型中用到的 `tf.ragged` 张量操作、`tf.random` 随机数生成等。具体的实现步骤如下。

(1)导入库及获取数据集。首先需要导入 TensorFlow 和 Keras 库。数据集采用的是 Keras 中常用的数据集 Datasets 中的 MNIST 手写字符数据集。

(2)整理数据集。将数据集拉伸为二维,并划分为训练集和验证集,并将其归一化处理。

(3)定义参数创建模型。对数据处理完成后,需要定义模型的参数,并使用激活函数将线性模型转为非线性,在该模型中使用的是 Softmax 函数。

(4)神经网络训练。在训练模型之前需要先编译模型,定义训练的次、速率等,使用梯度下降速度快且容易在最优值附近震荡的 Adam 优化器。训练模型时要指定训练/测试样本和训练/测试标签以及损失率和准确率,并打印进度和日志信息。最后,将准确率和损失率的变化绘制成可视化的二维图展示,图 1 和图 2 所示。

(5)模型及图像预测。在测试集上评估模型准确率。再从索引的第 0 个开始及其后 10 个对图像识别进行预测,把图像送入模型进行预测,返回结

果按照实际数字、预测数字及数字图像可视化展示,识别结果如图 3 所示。

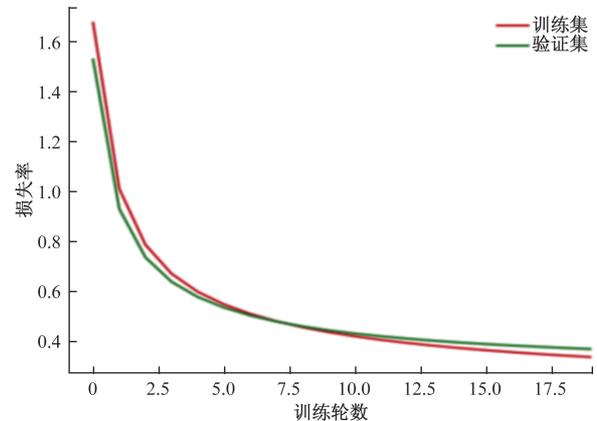


图 1 损失率变化

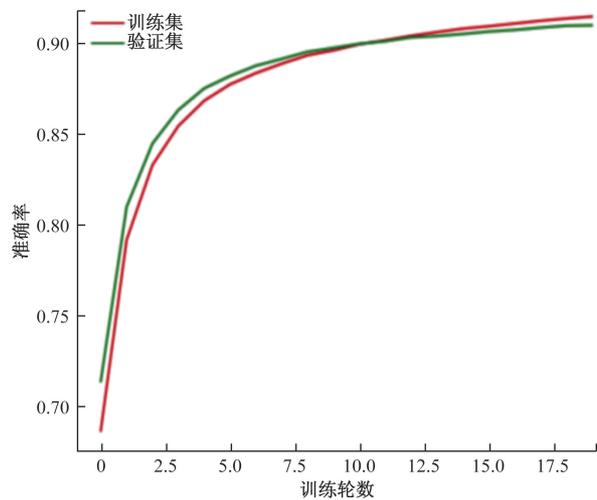


图 2 准确率变化

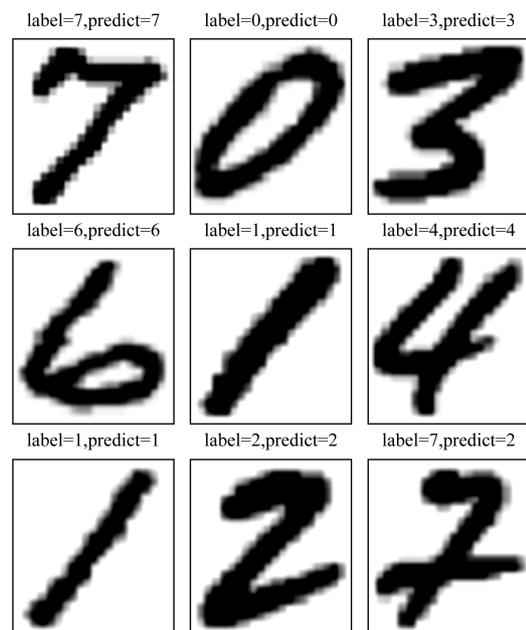


图 3 部分预测结果

3.2 单层卷积神经网络高阶 API 的手写数字识别实现

从低阶 API 的训练过程可以看出,整个训练过程都被暴露出来。第二个模型使用了高阶 API,一般为按照 OOP 方式封装的高级 API,在该模型中便增加高阶 API 提供的函数,如激励函数 ReLU 及评估函数 metrics。具体实现同第一个模型,但与第一个模型不同的是,在该模型中定义第一层隐藏层和输出层的权重和偏执项变量,并增加激励函数 ReLU 用以解决 gradient vanishing 问题及加快收敛速度。损失率和准确率的变化结果分别如图 4 和图 5 所示。

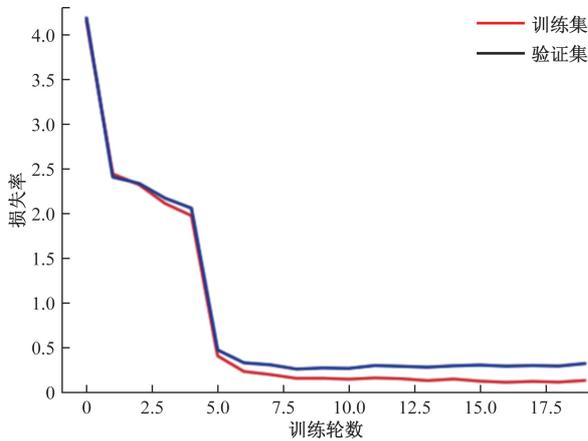


图 4 损失率变化

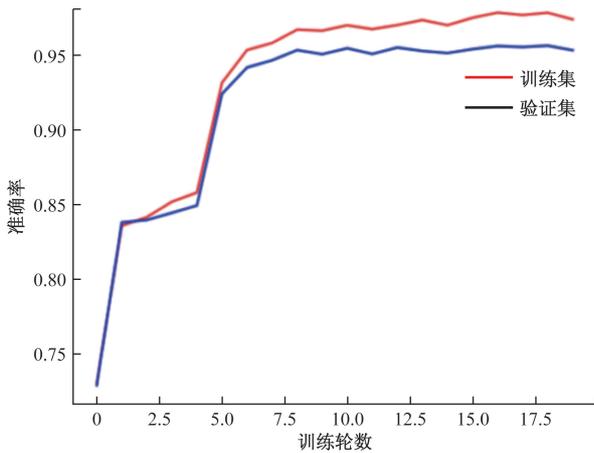


图 5 准确率变化

3.3 双层卷积神经网络的手写数字识别实现

不同于只有一层卷积的卷积神经网络,双层卷积神经网络有两层卷积层。多层神经网络可以从各个角度分析复杂数据,而单层只能做简单的分类和分析。在该模型中,最重要的就是有两层卷积,并且卷积模式相同,在每次卷积完成后都需要进入池化层对

特征图进行压缩,再一起进入全连接层进行后续输出。与前两个模型同样使用训练轮数(epoch)将训练集中每个样本都参与训练一次,打印出模型信息、损失率及准确率变化图,如图 6~图 8 所示。

```

Model:"sequential"
-----
Layer(type)
-----
conv2d(Conv2D) (None,28,28,32) 32θ
max_pooling2d(MaxPooling2D) (None,14,14,32) θ
conv2d_1(Conv2D) (None,14,14,64) 18 496
max_pooling2d_1(MaxPooling2D) (None,7,7,64) θ
dropout(Dropout) (None,7,7,64) θ
flatten(Flatten) (None,3 136) θ
dense(Dense) (None,512) 6 144
dense_1(Dense) (None,1θ) 513θ
-----

```

图 6 模型信息

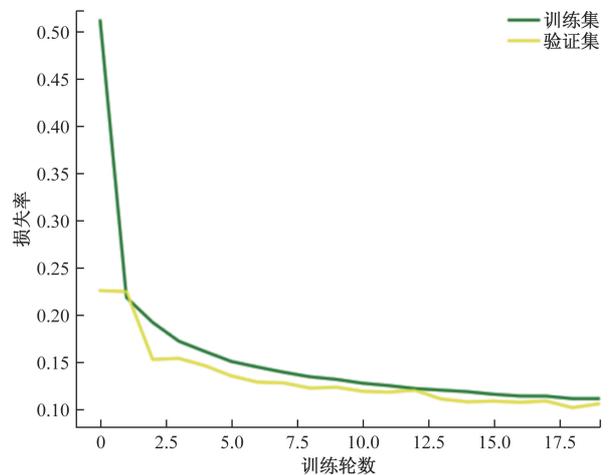


图 7 损失率变化

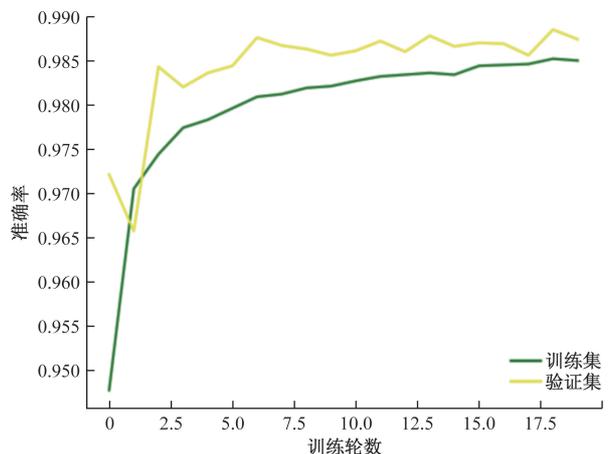


图 8 准确率变化

3.4 三层卷积神经网络的手写数字识别实现

第四个模型在双层卷积的基础上又增加了一个卷积层,由此可以纵向比较单层、双层、三层卷积

层模型的识别效果,并且利用该模型识别自己手写的数字。需要特别注意的是,与双层卷积模型不同的是,该模型有三层卷积并且卷积模式相同,在每次卷积完成后分别进入池化层,再一起进入全连接层进行后续输出,此处使用的激活函数是 ReLU 以及使用 Softmax 分为 10 类。最后展示该模型的信息及损失率、准确率变化,如图 9~图 11 所示。

```

Model:"sequential_1"
-----
Layer(type)          Output Shape          Param #
-----
conv2d(Conv2D)       (None,28,28,32)      320
max_pooling2d_2(MaxPooling2D)  (None,14,14,32)      0
conv2d_1(Conv2D)     (None,14,14,64)      18 496
max_pooling2d_3(MaxPooling2D)  (None,7,7,64)        0
conv2d_5(Conv2D)     (None,7,7,64)        36 928
max_pooling2d_4(MaxPooling2D)  (None,3,3,64)        0
dropout_1(Dropout)   (None,3,3,64)        0
flatten_1(Flatten)   (None,576)            0
dense_2(Dense)       (None,512)            295 424
dense_3(Dense)       (None,10)             510
-----
Total params:356 298
Trainable params:356 298
Non-trainable params:0
    
```

图 9 模型信息

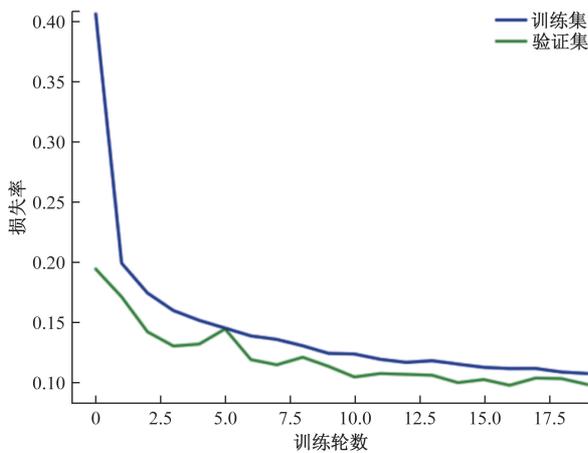


图 10 损失率变化

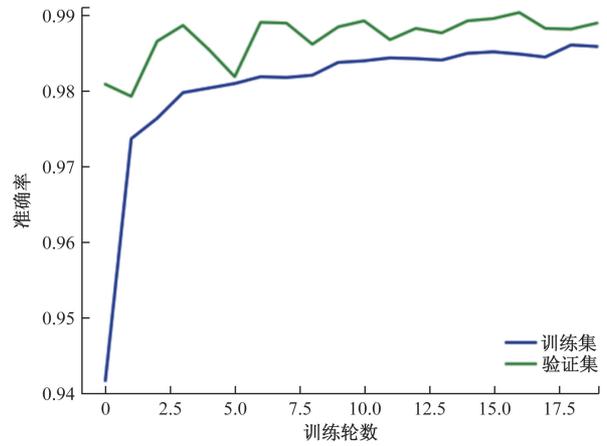


图 11 准确率变化



图 12 自己的手写数字



图 13 处理后的手写数字

3.5 识别自己的手写数字

模型训练完成后保存可用于预测,如果预测样本没有发生比较大的变化,那么训练好的模型理论上可以一直使用。根据以上训练,模型识别准确率很高,可以使用画图板手写数字,此处手写了如图 12 所示的数字“3”进行识别,手写数字图片保存为 jpg. 格式后,对该图片进行预处理、读取(图 13)、预测(图 14)。

3.6 讨论

本文尝试了四种不同的模型进行训练,对训

```

1/1 [=====] 13 ms/step
预测数字: 3
    
```

图 14 预测数字

练结果进行对比,模型区别的依据其一是根据 API 的不同对比低阶和高阶 API 的模型性能,其二是根据卷积层数的不同对比单层、双层和三层的模型性能。表 1 展示了同一数据集(MNIST)下,按 1:4 划分训练集和测试集,采用不同模型的卷积神经网络训练 20 次后的训练损失率和验证准确率。

表 1 各模型在 MNIST 数据集上的性能

模型	训练 损失率	测试 损失率	训练 准确率	测试 准确率
单层卷积神经网络低阶 API	0.335	0.367	0.914	0.910
单层卷积神经网络高阶 API	0.114	0.303	0.974	0.953
双层卷积神经网络高阶 API	0.112	0.106	0.985	0.987
三层卷积神经网络高阶 API	0.107	0.098	0.986	0.989

训练损失率是用于评估模型如何拟合训练数据的指标,也就是说,它评估模型在训练集上的误差,损失越大误差越大;测试准确率用于评估模型在测试集上的准确性,准确率越高越准确。

实验结果表明,在同样是单个卷积层的卷积神经网络下,低阶 API 训练损失率高达 0.335 而高阶 API 仅 0.114。高阶 API 不仅通过封装复杂计算步骤,简化了模型构建流程,还通过内置优化算法提高模型的计算效率和泛化能力,因此显著降低了识别误差。测试准确率也从 0.910 提升至 0.953,表明高阶 API 能够更高效地捕捉特征并提升模型性能。

在增加了卷积层之后,多层卷积模型的训练损失率进一步降低至 0.112(双层)和 0.107(三层),而测试准确率分别提升至 0.987 和 0.989。这是因为增加卷积层数可以提取更丰富、更深层次的特征,从而增强了模型的表达能力和对复杂数据的适应性。此外,尽管卷积层的增加延长了训练时间,但由于单次训练仅需 20 s 左右,因此计算成本仍在可接受范围内。

进一步分析表明,多层模型的性能提升还得益于卷积层对噪声的抑制和对局部特征的强化,使得特征提取更加精准。然而,卷积层数的增加也可能带来过拟合风险,因此在实际应用中需根据数据集的复杂性和任务需求,合理选择卷积层数和模型参数,以实现性能与效率的平衡。

4 总结与展望

在人工智能技术高速发展的信息化时代,基于人们对全球信息化和自动化程度的要求不断提升,以及在诸多领域中仍存在耗费大量人物财力实现人工识别手写数字的情况,一种利用深度学习技术高效准确的手写数字识别方式是极具研究价值、社会价值和经济价值的。本文理清了卷积神经网络及手写数字识别的发展趋势;实现使用基于卷积神经网络的手写数字识别的四种模型。最后的结果表明,在四个模型中高阶 API 三个卷积层的模型识别效果最优,在未来的研究中以及在各个行业领域中具有很大的应用潜力。

除此之外,本文还存在一定的问题有待进一步改进与研究:①在多层卷积神经网络构建时只测试了两层和三层卷积的模型,随着卷积层的增加识别准确率在何处会达到峰值,何时会过拟合有待进一步研究;②在构建模型时训练次数、验证集的比例、卷积核的大小等参数都选择了目前研究中多数人选定的数值,并未对参数进行修改、测试、分析,何时最优有待进一步研究;③MNIST 数据集中的每一个手写体数字都是完整的,然而现实中的手写体数值由于种种原因可能会不完整,将不完整的手写体数字添加到该数据集中,可以提高其鉴别的准确度。

参考文献

- [1] 唐鉴波,李维军,赵波,等.基于卷积神经网络的手写数字识别方法研究[J].电子设计工程,2022,30(21):189-193.
- [2] GOODFELLOW I, BENGIO Y, COURVILLE A. Deep learning(Vol. 1)[M]. Cambridge: MIT press, 2016.
- [3] 杨培伟,周余红,邢岗,等.卷积神经网络在生物医学图像上的应用进展[J].计算机工程与应用,2021,57(7):44-58.
- [4] LECUN Y, BOSER B, DENKER J S, et al. Backpropagation applied to handwritten zip code recognition[J]. Neural Computation, 1989(4): 541-551.
- [5] LECUN Y, BENGIO Y. Convolutional networks for images, speech, and time series[J]. The Handbook of Brain Theory and Neural Networks, 1995(10): 96-104.
- [6] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(3): 504-507.
- [7] 崔兆韵,崔焕森,宋德娟,等.改进卷积神经网络的冬小麦提取方法[J].科技和产业,2024,24(1):281-286.
- [8] 胡敏,陈楠,毕进娜.基于 SegNet 网络和概率图模型的工区障碍物提取[J].科技和产业,2023,23(17):266-272.
- [9] 刘善伟,马志伟,魏世清,等.基于轻量级卷积神经网络的岩石图像岩性识别方法[J/OL].地质科技通报,1-13 [2024-12-13]. <https://doi.org/10.19509/j.cnki.dzqk.tb20240348>.
- [10] 余国庆,杨燕婷,宗兆星,等.基于卷积神经网络的手写数字识别技术研究[J].安徽电子信息职业技术学院学报,2024,23(3):1-5.
- [11] 谭琬滢,左珊珊,邱佩琳,等.基于深度卷积神经网络的手写数字识别研究[J].智能计算机与应用,2024,14(8):138-142.
- [12] 雷国伟,吕迎阳,纪安妮,等.图像特征的 CNN 提取方法及其应用[J].计算机工程与应用,2004(14):204-206.
- [13] 樊继慧,滕少华,金弘林.基于改进 Sigmoid 卷积神经

- 网络的手写体数字识别[J]. 计算机科学, 2022, 49(12): 244-249.
- [14] 谢东阳, 李丽宏, 苗长胜. 基于改进 AlexNet 卷积神经网络的手写体数字识别[J]. 河北工程大学学报(自然科学版), 2021, 38(4): 102-106.
- [15] 陈群贤. TensorFlow 下基于 CNN 卷积神经网络的手写数字识别研究[J]. 信息记录材料, 2022, 23(9): 159-161.
- [16] SULAIMAN K, HAFEEZ A, HAZRAT A, et al. Pioneer dataset and recognition of handwritten pashto characters using convolution neural networks[J]. Measurement and Control, 2020, 53(9/10): 2041-2054.
- [17] ZHOU L J, YU W H. Improved convolutional neural image recognition algorithm based on LeNet-5[J]. Journal of Computer Networks and Communications, 2022, 35(5): 1-5.
- [18] 余国庆, 杨燕婷, 宗兆星, 等. 基于卷积神经网络的手写数字识别技术研究[J]. 安徽电子信息职业技术学院学报, 2024, 23(3): 1-5.
- [19] 纪绪. 基于卷积神经网络的手写数字识别研究与实现[J]. 信息与电脑(理论版), 2023, 35(12): 169-171.
- [20] 张趁香, 陈黄宇. 基于 LeNet-5 的手写数字识别的改进方法[J]. 电脑知识与技术, 2024, 20(12): 27-30.
- [21] 张焕, 张庆, 于纪言. 卷积神经网络中激活函数的性质分析与改进[J]. 计算机仿真, 2022, 39(4): 328-334.
- [22] 闫建红, 段运会. 动量余弦相似度梯度优化图卷积神经网络[J]. 计算机工程与应用, 2024(14): 133-143.
- [23] 蔡靖, 谷承睿, 刘光达, 等. 基于改进 AlexNet 卷积神经网络人脸识别的研究[J]. 电子技术应用, 2024(11): 42-46.

Research and Design of Handwritten Digit Recognition Based on Convolutional Neural Network

ZHU Jiaorong

(School of Cultural Heritage and Information Management, Shanghai University, Shanghai 201900, China)

Abstract: In the era of informatization, handwritten digit recognition has wide applications in computer vision and pattern recognition. In order to improve the accuracy and efficiency of handwritten digit recognition, four convolutional neural network (CNN)-based algorithm models were designed and tested. An empirical study was conducted on the MNIST dataset to compare the training performance of different models. The experimental results show that the multilayer CNN model performs best, with an accuracy rate of 98.9%, and each training cycle takes about 20 seconds. This indicates that increasing the number of convolutional layers and using high-level APIs help improve recognition accuracy, further enhancing the application performance of CNN in handwritten digit recognition. New insights and methods are provided for designing efficient and accurate handwritten digit recognition models.

Keywords: convolutional neural network; handwritten digit recognition; MNIST dataset; deep learning