**研究生课程考核试卷**

**（适用于课程论文、提交报告）**

**科 目： 人工神经网络及模式识别 教 师： 罗小刚**

**姓 名： 梁宵 学 号： 20241901014**

**专 业： 生物医学工程 类 别： 博士研究生**

**上课时间： 2024年 9 月至 2024年 11 月**

**考 生 成 绩：**

|  |  |  |
| --- | --- | --- |
| **卷面成绩** | **平时成绩** | **课程综合成绩** |
|  |  |  |

**阅卷评语：**

**阅卷教师 (签名)**

重庆大学研究生院制

**重庆大学研究生《人工神经网络及模式识别》课程考核要求**

注：1、本试卷格式用于考核方式为“提交报告”、“课程论文”、“考查”等各类别研究生课程的考核。

2、要有明确的课程考核要求：如课程论文（报告）题目（范围）、篇幅（字数）、必须的参考资料、提交时间等。并提前将课程考核试卷发给学生。

**作业要求：基于minist数据集，设计一个手写数字识别网络，要求该网络具有五层以上网络，中间层分别采用 sigmoid、relu作为激活函数(函数类型可选)，方差误差函数、交叉熵函数(误差函数类型可选) 在学习过程中采用两种以上不同的分布学习方法(方法可选，不做硬性要求)，分别探讨不同学习方法、学习率、激活函数、误差函数以及batchsize对学习精度、收敛性以及误差的影响。**

**(1)提交设计文档，文档中以流程图、拓扑图、公式、文字、伪码等方式结合，详细说明设计原理、过程、结果、结果讨论、课程学习体会等。**

**(2)提交源程序、训练数据集、测试数据集、程序使用说明(如有必要)并打包，要能实现解程序 后能直接运行测试。**

3、提交课程论文撰写格式参考《重庆大学博士、硕士学位论文撰写格式标准》。

## 一、引言

近年来，随着计算能力的提升和大规模数据集的普及，深度学习在人工智能领域取得了显著进展，尤其是卷积神经网络（Convolutional Neural Networks, CNN）在图像处理领域表现尤为突出。深度学习的核心在于通过多层非线性变换对数据的高阶特征进行自动提取和学习，摆脱了传统机器学习中对特征工程的依赖。针对不同任务，研究者们提出了多种深度神经网络架构，如用于图像分类的 AlexNet、ResNet，以及广泛应用于自然语言处理的 Transformer 等，这些模型推动了计算机视觉、语音识别和文本分析等领域的飞跃发展。然而，尽管深度学习展现出了强大的表现能力，其在模型优化、网络设计、泛化能力和计算资源需求等方面仍然面临挑战。因此，设计高效、鲁棒的网络结构，并深入探索参数对模型性能的影响，依然是深度学习领域的重要研究方向。

在本研究中，我设计并实现了一个包含多层卷积层和全连接层的深度学习模型，结合多种激活函数和优化策略，探索不同网络结构及参数对分类性能的影响。模型采用交叉熵损失函数和 Adam 优化器，以提升收敛速度和精度，同时通过 Dropout 技术有效缓解了过拟合问题。实验结果表明，该模型在 MNIST 数据集上的分类准确率达到 99.72%，展现出较高的泛化能力和鲁棒性。本研究不仅验证了深度学习方法在手写数字分类任务中的有效性，还为优化网络结构提供了实践依据。

## 二、神经网络设计原理

**2.1 网络结构设计**

这是一个包含至少五层的神经网络架构，其中包含多种类型的激活函数和误差函数选择。网络的基本结构如下：

网络分为以下层级：

输入层：接收 28x28 的灰度图像，并将其展平为一维向量（784维），故输入层设置784个节点。

中间层：

第一层卷积层 + 最大池化层：16 个神经元，激活函数使用 ReLU。

第二层卷积层 + 最大池化层：32 个神经元，激活函数使用 ReLU（Tanh）。

第三层卷积层：64 个神经元，激活函数使用 ReLU。

输出层：10 个神经元（分别对应数字 0-9），激活函数使用 Softmax。



图1 神经网络设计流程图

**2.1.1 卷积层**

卷积层是 CNN 的核心模块，用于提取局部特征。它通过卷积核（或称过滤器）对输入数据进行滑动窗口操作，捕获空间关系（如边缘、纹理等）。在该网络的设计中，第一层卷积层采用了 32 个 3×3的卷积核，输入为 MNIST 数据集的灰度手写数字图片（28×28）。这一层的作用是提取低级特征，如边缘和角点，得到的特征图大小为16×27×27。

第一层最大池化采用了 32 个 2×2的池化核，池化后尺寸减半，得到的特征图大小为16×13×13。

第二层和第三层的卷积核数量分别为 32 和 64，更多的卷积核允许网络捕获更加复杂的高级特征，如曲线、数字形状等。卷积操作后每个特征图通过 ReLU 激活函数，引入非线性以提升网络的表达能力。最终提取得到64×2×2的特征图。

**2.1.2 激活函数**

激活函数是网络的关键组成部分，用于为模型引入非线性特性。在卷积层之后，使用 ReLU（Rectified Linear Unit）作为激活函数，其定义为：

相比于传统的 Sigmoid 或 Tanh 函数，ReLU 的优势在于其计算效率高且能够减轻梯度消失问题。通过 ReLU 激活后的特征图保留了非线性信息，为网络提供了更强的拟合能力。

**2.1.3 池化层**

池化层是 CNN 中的重要组成部分，用于降低特征图的空间维度，同时保留主要特征信息。在本网络中，每次卷积操作后都紧跟一个2×2 的最大池化层。池化的作用主要体现在以下几个方面：  
 相比于传统的 Sigmoid 或 Tanh 函数，ReLU 的优势在于其计算效率高且能够减轻梯度消失问题。通过 ReLU 激活后的特征图保留了非线性信息，为网络提供了更强的拟合能力。

**2.1.4 全连接层**

在卷积和池化操作的基础上，将提取的多维特征通过 Flatten 层展平为一维向量，输入到全连接层进行分类。全连接层的主要目的是将高维特征映射到输出类别。

在本设计中，首先使用一个具有 64 个神经元的全连接层，通过 ReLU 激活函数进一步处理高阶特征，并在最后的输出层中使用 Softmax 激活函数。Softmax 函数的输出是归一化的概率分布，表示输入属于每个类别的可能性。

**2.1.5 数据预处理与归一化**

在输入网络之前，MNIST 数据集的像素值被归一化到 [0,1] 范围，以加速训练过程并提升模型性能。此外，对标签进行了独热编码，将原始标签（0-9）转换为 10 维的向量，以便于在最后的分类层中使用交叉熵损失函数。

**2.1.6 损失函数设计**

损失函数采用 交叉熵损失函数（Categorical Crossentropy），其公式为：

其中 是真实标签，是预测概率。交叉熵能够很好地衡量分类问题中模型输出概率与真实标签分布之间的差距。

优化器选择了 Adam，结合了动量和自适应学习率的优点，能快速收敛并适应不同梯度的变化，公式如下：

**2.2 参数设计**

**学习率（Learning Rate）**：1e-4。学习率默认值为1e-4，决定了模型在每次迭代时更新参数的幅度。较低的学习率通常能提升稳定性，但会减缓收敛速度。

**训练迭代次数（Training Iterations）**：37500。其中每个 epoch 的训练迭代次数是750，该参数指定了模型训练过程中参数更新的次数。

**Dropout 比例（Dropout）**：0.5。在训练时随机"杀死"50%的神经元以降低过拟合风险。

**批次大小（Batch Size）**：64。每次训练中使用的样本数量。模型会在这个批次中完成一次前向传播和反向传播，更新一次权重。

**2.3 分布学习方法**

随机梯度下降法 (SGD)：

更新规则：

Adam 优化器：

包含动量和二阶矩估计，更新规则：

**2.4 网络设计思想与优点**

本网络设计基于经典的 CNN 结构，逐层提取特征并逐步聚合。以下是其优点：

**分层特征提取**：从简单的边缘到复杂的全局形状逐层提取。

**参数共享**：卷积核共享权重，显著降低了参数规模。

**空间降维**：通过池化操作减少计算量，防止过拟合。

**非线性增强**：ReLU 减轻了梯度消失问题，提高了训练效率。

## 三、结果

基于 MNIST 数据集训练了一个卷积神经网络 (CNN)，网络结构包含三层卷积层和两层全连接层。经过 200 个 epoch 的训练，模型在测试集上的表现如下：

测试集准确率 (Test Accuracy)：99.72%

测试集损失值 (Test Loss)：0.0263

为了更直观地展示模型的训练过程，绘制了训练和验证的准确率随 epoch 变化的曲线，如图2所示。由图可知，训练准确率和验证准确率在逐渐收敛，且验证准确率略低于训练准确率，表明模型具有良好的泛化能力。

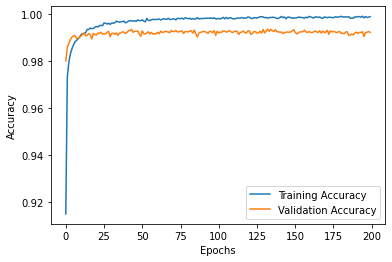


图2 模型准确率预测曲线

在训练过程中，模型的训练精度快速上升并趋于稳定，最终达到了较高的精度水平。这表明模型成功学习到了训练数据中的特征。验证精度的变化趋势相对平稳，与训练精度基本一致，说明模型具有一定的泛化能力，没有出现明显的过拟合现象。在验证集上的精度曲线表明，模型在新数据上的分类能力较为稳定，但仍存在少许波动。这可能与验证数据的复杂性或模型对特定模式的学习偏差有关。综合来看，该模型在当前任务上的性能表现良好，能够较好地完成分类任务。

为了直观显示模型在每个类别上的分类表现，反映分类错误的分布，通过混淆矩阵来验证模型实际的预测性能，如图3所示。行表示实际类别，列表示预测类别。对角线上的值越高，说明模型分类越准确。图中对角线上的值代表模型预测正确的样本数量，可以看到大部分样本集中在对角线上，说明模型的分类性能较好。非对角线上的值表示分类错误的样本数量，这些值相对较小，表明模型的误分类率较低。

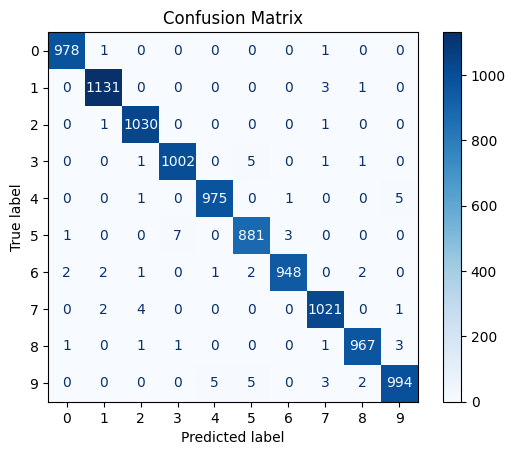


图3 混淆矩阵

为了进一步说明模型的稳定性， 采用交叉验证对模型泛化能力进行评估。如图4所示，这是一个 5 折交叉验证的准确率变化情况，每个折叠的准确率通过蓝色的线连接，每个数据点对应一个折叠的准确率值。第一个折叠的准确率大约为 0.9980，显示了较高的性能。第二折叠和第三折叠的准确率略有提升，但它们之间的差异非常小，维持在 0.9982 左右。第四折叠的准确率有显著提升，接近 0.9990，表明在这个折叠中模型表现得尤为好。第五折叠的准确率仍然很高，维持在 0.9990 左右。较高的准确率（接近 0.998-0.999）表明，该模型的整体表现是稳定的，能够很好地学习到手写数字识别任务中的特征。

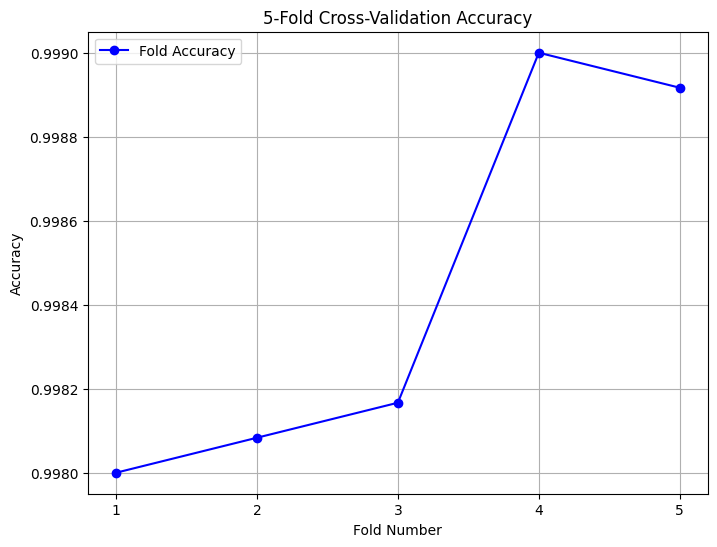


图4 五折交叉验证

## 四、讨论

本研究中构建的卷积神经网络 (CNN) 在 MNIST 数据集上的表现优异，验证集准确率达到了 99.72%，损失值仅为 0.0263。这表明模型能够精准地捕获手写数字的特征，并在分类任务中实现了高效的性能。从训练和验证准确率随 epoch 变化的趋势来看，模型在训练过程中快速收敛，训练与验证准确率曲线接近，且验证集上的精度波动较小，说明模型具备良好的泛化能力，并未出现明显的过拟合现象。

通过交叉验证进一步评估模型的稳定性，5 折验证结果显示每个折叠的准确率均维持在 99.8%-99.9% 的高水平，这表明模型在不同数据分布下均表现出一致性，能够有效学习到数据的内在特征。同时，混淆矩阵结果显示，大部分样本集中于对角线，说明模型对每个类别的分类性能较好，仅在少数类别间存在轻微混淆，误分类率较低。然而，模型的表现仍可进一步优化，例如在数据增强上引入更复杂的变换以增加数据多样性，或尝试更深层次的网络结构以进一步提升模型的特征提取能力。此外，针对误分类的类别，可以结合注意力机制或类别加权策略，进一步降低错误率。整体架构除了MLP外也可以采用KAN网络结构可能可以实现更精确的效果。综合来看，该模型在手写数字识别任务中展现了极高的准确性和稳定性，是一个有效的解决方案。

## 五、学习体会

作为生物医学工程专攻生物信息学的一名博士，学习神经网络课程让我深刻认识到这一技术在生物信息学领域的潜力和应用价值。在课程中，我不仅掌握了神经网络的基本原理、激活函数、优化算法等理论知识，还深入了解了深度学习如何解决复杂的分类和预测问题。通过课程作业，我尝试构建了手写数字识别神经网络，并得到了高准确率的预测结果。结合专业知识，我尝试将神经网络模型应用于实际生物数据的分析，如基因表达模式分类、蛋白质结构预测和活性肽的功能筛选等。

神经网络在解决高维数据的非线性关系方面展现出独特优势，与生物信息学中常见的挑战完美契合。同时，这门课程也让我意识到，设计高效的模型需要在网络结构、超参数优化和数据处理之间找到平衡，且模型性能的提升离不开对生物学问题的深入理解。通过这段学习，我不仅提升了算法设计能力，更增强了将计算工具应用于生物学研究的信心，这帮助我在未来开展生物医学领域的创新研究奠定了坚实的基础。