

# MNIST 实验指导

## 实验概览

### 实验目的

1. 了解 **深度学习** 的基本概念，掌握神经网络的基本原理。
2. 掌握 **TensorFlow.js** 的基础知识，学会使用 **JavaScript** 进行机器学习模型的搭建、训练和推理。
3. 提高代码实现能力，掌握模型训练、优化和评估的流程。

### 实验要求

1. 了解 **深度学习** 的基本概念。
2. 搭建基于 **TensorFlow.js** 的机器学习环境。
3. 训练并测试 **LeNet手写数字识别模型**，分析其性能。
4. 观察实验中的**异常情况**，思考其原因，找到接解决办法。
5. 能够对**LeNet模型**进行优化，提高识别精度。

### 提交内容

实验文件夹里除了 `node_modules/` 之外的所有内容，打成压缩包。

### 提交地址

<https://icloud.qd.sdu.edu.cn:7777/link/5E280616573EE369D997AD5132CABFDE>

### 截止日期

2025-04-13

## 1. 深度学习

深度学习（Deep Learning）是一种基于人工神经网络的机器学习方法，旨在模拟人脑的学习方式，从大量数据中自动提取特征并进行预测或决策。它是**机器学习**和**人工智能**的一个重要分支，在计算机视觉、自然语言处理、语音识别等领域取得了巨大成功。深度学习利用大数据和强大的计算能力，实现了许多过去难以完成的任务，并在多个领域取得了突破性进展。

### 核心概念

#### 1. 神经网络（Neural Networks）

- 由**输入层**、**隐藏层**和**输出层**组成，每一层包含多个**神经元（Neuron）**。
- 通过**权重（Weights）**和**激活函数（Activation Function）**来调整和传播信息。

#### 2. 深度神经网络（DNN）

- 多层神经网络的总称，通常包含多个隐藏层，能够学习复杂的特征表示。

#### 3. 反向传播（Backpropagation）

- 通过**梯度下降算法（Gradient Descent）**优化网络参数，使误差不断减小。

#### 4. 常见深度学习模型

- 卷积神经网络（CNN）**：擅长处理图像数据，如目标检测、人脸识别等。
- 循环神经网络（RNN）**：适用于处理序列数据，如语音识别、文本生成等。
- Transformer**：基于自注意力机制，在自然语言处理（如ChatGPT、BERT）等领域表现优异。

### 深度学习的应用

- 计算机视觉**：图像分类、人脸识别、自动驾驶等。
- 自然语言处理（NLP）**：机器翻译、聊天机器人、文本摘要等。
- 语音识别**：智能助手（如Siri、Alexa）。
- 医学诊断**：疾病预测、医学影像分析。
- 游戏 AI**：AlphaGo 等。

## 2. 卷积神经网络

卷积神经网络（**CNN**，Convolutional Neural Network）是一类特殊的深度神经网络，广泛应用于图像处理、计算机视觉等任务。CNN通过模拟人类视觉系统的处理方式，能够自动从图像中提取特征，并进行分类、识别等任务。与传统神经网络相比，CNN在处理图像等高维数据时具有更高的效率和准确性。

### CNN的核心结构

#### 1. 卷积层（Convolutional Layer）

- 卷积层是CNN的核心，通过**卷积操作**对输入图像进行处理。
- 通过使用多个**卷积核（滤波器）**，提取不同的特征（如边缘、纹理等）。
- 卷积操作是局部的，意味着每个卷积核只关注输入图像的局部区域，有效减少计算量。

#### 2. 激活函数（Activation Function）

- 通常使用**ReLU（Rectified Linear Unit）**作为激活函数，引入非线性，使网络能够学习复杂的特征。

#### 3. 池化层（Pooling Layer）

- 池化层用于降低图像的空间尺寸，从而减少计算量和过拟合的风险。
- 最大池化（Max Pooling）**和**平均池化（Average Pooling）**是常见的池化方法。最大池化选取局部区域中的最大值，减少图像的尺寸。

#### 4. 全连接层（Fully Connected Layer）

- 在CNN的最后阶段，通常会使用全连接层将卷积层和池化层提取的特征映射到输出空间，如分类标签。
- 该层将所有神经元连接到每个输出神经元，用于最终的分类或回归任务。

#### 5. 输出层（Output Layer）

- 输出层通常采用Softmax激活函数，用于多类分类任务，输出每一类的概率。

### CNN的优点

- 参数共享**：卷积操作中，同一个卷积核会应用到输入图像的不同区域，这大大减少了网络的参数数量，提高了训练效率。
- 局部连接**：卷积核只关注局部区域，有助于捕捉局部特征，如边缘、纹理等。
- 自动特征学习**：CNN能自动从数据中学习特征，不需要手工设计特征。

## CNN的应用

- 图像分类**：如识别图像中的物体、动物等。
- 目标检测**：如在图像中检测出人脸、车辆等目标。
- 语音识别**：通过音频信号的图像化处理，进行语音识别。
- 医学影像分析**：用于疾病诊断，如癌症的检测。

### 3. 二维卷积层

卷积神经网络（Convolutional Neural Network, CNN）是一类包含卷积层（Convolutional Layer）的神经网络结构，广泛应用于计算机视觉任务。与传统的全连接神经网络不同，卷积神经网络能够有效地捕捉数据的局部特征，并通过共享权重减少参数数量，提高计算效率。

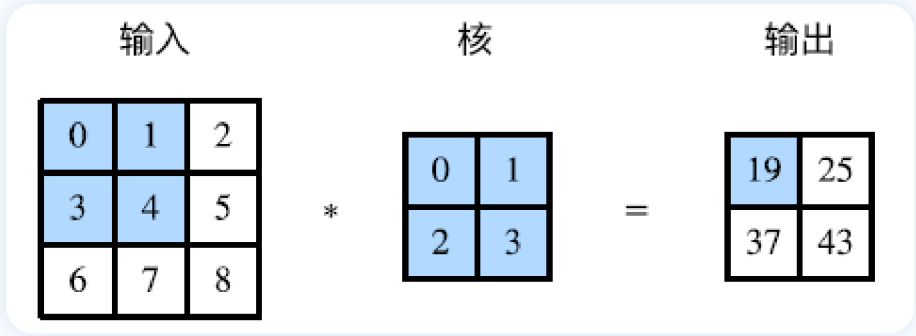
这里主要讨论 **二维卷积层**，它是最常见的卷积层类型，具有 **高度** 和 **宽度** 两个空间维度，主要用于处理图像数据。

#### 互相关运算与卷积层

虽然卷积层的名称来源于数学中的 **卷积（Convolution）** 运算，但在深度学习中，通常使用的是更加直观的 **互相关（Cross-correlation）** 运算。在 **二维卷积层** 中，输入数据以 **二维数组** 形式存在，而 **卷积核（Kernel）**（也称 **过滤器（Filter）**）也是一个二维数组。二者通过 **互相关运算** 生成 **输出特征图**，从而提取局部特征信息。

#### 卷积窗口的形状

卷积窗口（又称 **卷积窗口大小**）的形状取决于 **卷积核的高度和宽度**。常见的卷积核尺寸包括 **2x2**、**3x3**、**5x5**、**7x7** 等，尺寸的选择影响特征提取的精度和计算复杂度。较小的卷积核适用于捕捉局部细节，而较大的卷积核则有助于提取更高级的全局特征。



## 二维互相关运算的过程

通过一个具体的例子来理解 二维互相关运算 的工作原理。

### 1. 卷积窗口滑动：

- 计算时，卷积核（Kernel）作为滑动窗口，从输入数组的左上角开始，按照从左到右、从上到下的顺序依次移动。
- 这种滑动方式允许卷积核覆盖输入数组的不同区域，提取不同局部特征。

### 2. 计算输出值：

- 当卷积窗口滑动到某一位置时，窗口内的输入子数组与卷积核进行按元素相乘，并将所得结果求和，得到输出特征图中对应位置的元素值。
- 这个过程会在输入数据的各个区域重复进行，最终生成完整的输出特征图。

$$\begin{aligned} 0 \times 0 + 1 \times 1 + 3 \times 2 + 4 \times 3 &= 19, \\ 1 \times 0 + 2 \times 1 + 4 \times 2 + 5 \times 3 &= 25, \\ 3 \times 0 + 4 \times 1 + 6 \times 2 + 7 \times 3 &= 37, \\ 4 \times 0 + 5 \times 1 + 7 \times 2 + 8 \times 3 &= 43. \end{aligned}$$

二维卷积层通过将输入数据与卷积核（Kernel）进行互相关运算，并加上一个标量偏差（Bias）来生成输出特征图。在神经网络中，卷积层的核心参数包括卷积核权重和偏差项，这些参数决定了网络的特征提取能力。

在模型训练过程中，通常先对卷积核进行随机初始化，以打破对称性，防止所有神经元学习相同的特征。随后，卷积核和偏差通过梯度下降和反向传播不断优化，使网络逐步提升特征提取能力，从而在图像分类、目标检测、语音识别等任务中发挥关键作用。

## 4. LeNet

**LeNet**是由Yann LeCun等人于1998年提出的卷积神经网络（CNN）架构，是深度学习和计算机视觉领域的经典模型之一。LeNet的设计对于后来的CNN架构和现代深度学习的发展具有重要影响，它主要用于手写数字识别，尤其是在**MNIST数据集**上取得了显著成功。

### LeNet的特点

- **结构简单**：相比于现代复杂的深度网络，LeNet的结构较为简单，包含了卷积层、池化层、全连接层，适合处理小规模的数据（如手写数字识别）。
- **卷积操作**：LeNet使用了卷积操作来提取图像特征，减少了传统全连接神经网络中的参数数量。
- **池化层**：通过池化操作减少特征图的尺寸，降低了计算复杂度。

### LeNet的贡献

1. **深度学习的早期应用**：LeNet是卷积神经网络在图像识别任务中的早期成功应用之一，奠定了深度学习在计算机视觉领域的基础。
2. **权重共享和局部连接**：LeNet提出的卷积层和池化层的设计思想（如权重共享和局部连接）是现代CNN架构的基石。
3. **手写数字识别**：LeNet在MNIST数据集上取得了优异的成绩，推动了手写数字识别技术的发展。

尽管LeNet的结构在今天看来相对简单，但它为后来的更复杂的神经网络模型（如AlexNet、VGG、ResNet等）提供了宝贵的设计思想。

## 5. MNIST 数字识别

本次实验要求使用 **TensorFlow.js** 构建一个简单的卷积神经网络LeNet，进行 **MNIST** 识别。

具体实验流程参考压缩包里的 `readme.md`。