

2 0 X X

基于卷积神经网络 的手写数字识别

01

前言

02

内容

03

解决方案

04

实现步骤

05

代码实现

06

结束语

基于卷积神经网络的手写数字识别

前言

手写数字识别，作为机器视觉入门项目，无论是基于传统的OpenCV方法还是基于目前火热的深度学习、神经网络的方法都有这不错的训练效果。当然，这个项目也常常被作为大学/研究生阶段的课程实验。可惜的是，目前网络上关于手写数字识别的项目代码很多，但是普遍不完整，对于初学者提出了不小的挑战

基于卷积神经网络的手写数字识别

内容

我们的要做的是，训练出一个人工神经网络，使它能够识别手写数字实现输图片对所检测到的数字进行检测并识别输出

基于卷积神经网络的手写数字识别

解决方案

图1 检测流程

灰度化是将彩色图像转换为灰度图像的过程。在灰度图像中，每个像素的值表示了该点的亮度，通常用0(黑色)到255(白色)之间的整数表示

基于卷积神经网络的手写数字识别

二值化是将灰度图像转换为黑白图像的过程，即将图像中的像素值转换为只有两个值，通常是0和255(或1和0)，分别代表黑色和白色。二值化可以帮助凸显图像中的目标物体的轮廓和形状，简化了图像处理和分析的复杂度

剔除小连通区域是图像处理中常用的操作，用于去除图像中的噪声或不重要的细小区域，剔除小连通区域可以帮助提高图像处理和分析的准确性，保留了主要的目标区域

ROI (Region of Interest)指的是图像中感兴趣的区域，通常是指包含了我们关注的目标或信息的部分。在图像处理中，提取ROI区域可以帮助我们集中精力分析和处理感兴趣的部分，而忽略其他部分。ROI区域提取在图像处理和计算机视觉中应用广泛，例如在目标检测、人脸识别、医学图像分析等领域中都有重要的作用。提取ROI区域可以帮助集中注意力和资源在感兴趣的部分，提高处理效率和准确性

基于卷积神经网络的手写数字识别

常见的应用包括将图像缩小以适应屏幕显示、将图像放大以进行细节分析、将图像缩放到特定的输入尺寸以适应深度学习模型等



在图像处理中，缩放是指改变图像的尺寸大小，通常是为了适应特定的显示或处理需求









缩放可以帮助适应不同的需求和平台，是图像处理中常用的操作之一

基于卷积神经网络的手写数字识别

实现步骤

4.1 数据集选择

此次使用手写数字识别经典数据集：本文数据集选择的 FashionMint 数据集 中的 t10k，共含有一万张 28*28 的手写图片（二值图片）

 t10k-images-idx3-ubyte.gz	2024/1/11 22:54	Bandizip.gz	1,611 KB
 t10k-labels-idx1-ubyte.gz	2024/1/11 22:54	Bandizip.gz	5 KB
 train-images-idx3-ubyte.gz	2024/1/11 22:54	Bandizip.gz	9,681 KB
 train-labels-idx1-ubyte.gz	2024/1/11 22:54	Bandizip.gz	29 KB
 t10k-labels.idx1-ubyte	2024/1/11 22:54	IDX1-UBYTE 文件	10 KB
 train-labels.idx1-ubyte	2024/1/11 22:54	IDX1-UBYTE 文件	59 KB
 t10k-images.idx3-ubyte	2024/1/11 22:54	IDX3-UBYTE 文件	7,657 KB
 train-images.idx3-ubyte	2024/1/11 22:54	IDX3-UBYTE 文件	45,938 KB

基于卷积神经网络的手写数字识别

→ 4.2 构建网络

- 残差网络 (Residual Network, 通常缩写为ResNet) 是卷积神经网络 (CNN) 的一种分支。它在2015年由微软亚洲研究院的研究员提出, 并在ImageNet图像识别挑战赛中取得了非常显著的成绩
- 残差网络的核心思想是引入了残差学习, 通过引入跨层的连接, 使得网络可以学习残差映射, 而不是直接学习原始映射。这种残差学习的方式可以帮助解决深度神经网络训练过程中的梯度消失和梯度爆炸问题, 使得网络可以更深更容易地训练
- 在残差网络中, 每个基本的残差单元包含了两个分支, 一个是恒等映射 (即输入直接连接到输出), 另一个是残差映射 (对输入进行变换后与恒等映射相加)。这种结构使得网络可以学习到残差, 从而更好地适应于复杂的特征学习任务

基于卷积神经网络的手写数字识别

残差网络的出现极大地推动了深度学习在计算机视觉领域的发展，成为了图像分类、目标检测、语义分割等任务中的重要模型架构

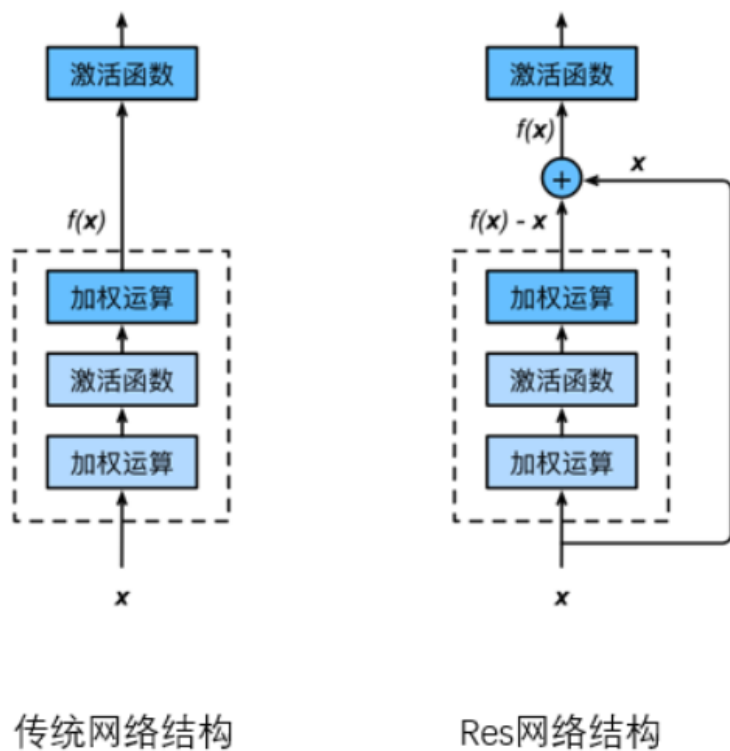
采用Resnet (残差网络)，残差网络的优势在于

- 1 更易捕捉模型细微波动
- 2 更快的收敛速度

基于卷积神经网络的手写数字识别

本次设计的网络结构如下图所示

图4 残差神经网络图



4.3 训练网络

本次设计设置训练次数为100个循环，网络的训练过程是这样的

- 1 给网络：模型“喂”数据(图像+标签)
- 2 网络根：据“喂”来的数据不断自我修正权重
- 3 本文一：共“喂”100次1万张图像

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：
<https://d.book118.com/406004231153010130>