

# 基于深度学习的手写文字识别

万茹月, 海玲, 谷铮, 刘文

(新疆工程学院 控制工程学院, 新疆 乌鲁木齐 830023)

**摘要:** 当前越来越多的场景需要将手写体的文字转换为电子格式, 手写体识别成为人机交互最便捷的手段之一, 拥有广泛的应用前景。文章提出了一种基于 TensorFlow 框架的深度学习手写识别方法, 包含手写数字识别和手写汉字识别。以 TensorFlow 为框架, 采用 CNN 神经网络模型建立训练集以降低识别错误率。实验结果最终表明, 对手写数字的识别率达到 95%, 对手写汉字的识别率达到 90%。

**关键词:** TensorFlow; 手写字体识别; 深度学习; 人工智能

**中图分类号:** TP18; TP391.4

**文献标识码:** A

**文章编号:** 2096-4706 (2021) 19-0089-04

## Handwritten Word Recognition Based on Deep Learning

WAN Ruyue, HAI Ling, GU Zheng, LIU Wen

(School of Control Engineering, Xinjiang Institute of Engineering, Urumqi 830023, China)

**Abstract:** At present, more and more scenes need to convert handwritten words into electronic format. Handwritten recognition has become one of the most convenient means of human-computer interaction and has a wide application prospect. This paper proposes a deep learning handwriting recognition method based on TensorFlow framework, including handwritten numeral recognition and handwritten Chinese character recognition. Taking TensorFlow as the framework, the CNN neural network model is used to establish a training set to reduce the recognition error rate. The experimental results finally show that the recognition rate of handwritten numeral reaches 95%, and the recognition rate of handwritten Chinese characters reaches 90%.

**Keywords:** TensorFlow; handwritten word recognition; deep learning; artificial intelligence

### 0 引言

随着计算机网络飞速的发展, 人机交互与手写汉字的结合既可为人们节省时间, 还能为他们的生活提供方便。虽然人类可以很容易地辨别出自己与他人的手写汉字书写的, 但由于许多汉字具有相似性、书写过程中的盲目性与不规范性, 会导致识别效果不理想等现象, 所以手写汉字识别仍然是人工智能研究中一个很大的难题<sup>[1]</sup>。和印刷体汉字相比, 脱机手写汉字对书写者没有书写规范的要求, 书写风格因人而异, 书写字体差别很大。和联机手写汉字相比, 脱机手写汉字没有笔画间的顺序信息, 每个汉字由图像中像素点组成, 所能提供的信息量较少<sup>[2]</sup>。

本文以 TensorFlow 作为框架, 使用 CASIA-HWDB1 数据集和 MNIST 数据集, 在实验中通过 CNN 神经网络模型进行训练, 对汉字的识别率最高可达 92%, 对数字的识别率最高可达 98%。

### 1 相关工作

目前针对手写文字研究的方法很多, 主要分为传统识别方法与深度学习识别方法。相对于英文识别而言, 汉字识别的结构复杂、种类繁多, 识别难度更大, 具有广泛的应用前景<sup>[3]</sup>。在传统文字识别研究中, 20 世纪 60 年代, 美国

IBM 公司用模板匹配法对印刷体汉字的模式识别研究工作成功的识别出了 1 000 个印刷体汉字; 在 80 年代我国也开始了对于手写汉字的研究。随着这些研究的需要, 相关数据集也随之诞生。如最大脱机手写汉字库 HCL2000, 中科院的 CASIA 中文数据集与英文数据集, 还有基于 Wacom 笔的 Text 数据集。在手写数字识别领域火热得到 MNIST 数据集, 美国邮政服务数据集 USPS, 用于手写数字识别的 UCT。

最近几年来, 深度学习理论不断颠覆了传统的识别算法, 已经成为模式识别领域研究热点之一<sup>[4]</sup>, 也在影响着手写体识别的研究。随着深度学习的不断深入研究出现了很多优秀的深度学习框架, 这些框架作为搭建深度学习模型的工具, 里面包含很多算法, 开发者利用这些算法可以轻松搭建符合自己的深度学习模型。如在 2012 年应用两种方法 (AlexNet, Dropout) 处理 ImageNet, 在 2013 年对 Atari 的应用, 2014 年的 Adam 优化器, 2015 年的残差网络, 2017 年的 Transformer 模型, 2017 年的阿尔法狗, 2018 年在深度学习的自动驾驶系统, 2020 年的自监督学习等, 深度学习的产物已经在我们的生活中扮演重要得多角色。自 2015 年 TensorFlow 面世, 因其稳定得到 API 和良好的兼容性受到众多开发者的青睐, 逐渐成为最常用的深度学习框架。TensorFlow 作为端到端的开放源码机器学习平台, 它而灵活的生态系统, 包括丰富的图书馆和社区资源<sup>[5]</sup>。

针对手写汉字结构复杂, 识别难度大的特点, 本文以 TensorFlow 作为框架, 使用中科院自动化所发布的中文手写

数据 CASIA 中文手写数据集 HWDB1。在实验中通过简单的神经网络模型训练,训练部分汉字的识别率能高达 90% 以上。以 MNIST 作为数据集实现对手写数字的识别,经实现,在简单的设计中,对数字的识别率能达到很高的准确率。主要贡献有:

(1) 基于 TensorFlow 的手写数字的识别:手写数字识别中用 TensorFlow 作为框架训练一个简单的卷积神经网络结构,用 MNIST 库作为数据集来源,使用 PyGame 做一个画板在电脑上绘制数字,然后完成对于手写数字的识别。

(2) 基于 TensorFlow 的手写汉字的识别:利用 TensorFlow 作为框架训练卷积网络,用 HWDB 作为训练集,使用 Flask 搭建一个微服务器上传图片,完成对手写汉字的识别。

## 2 手写字识别系统的实现过程

手写汉字识别系统基于 TensorFlow 和 CNN 神经网络模型完成。用 TensorFlow 作为框架,通过卷积网络,训练 HWDB1 库,使用 Flask 搭建了一个微服务器,能上传图片,实现图片识别。手写汉字识别系统流程图如图 1 所示。

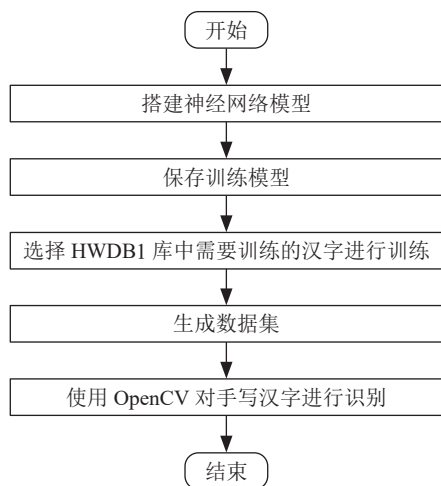


图 1 手写汉字流程图

### 2.1 深度学习

目前在人工智能应用最为广泛的便是深度学习,深度学习是让机器能够像人一样去执行某些任务,例如文字识别、动物图像识别以及其他模式识别等。深度学习不是执行已有的预定义好的模式,而是通过深度学习设置相关数据的参数,并通过使用多层处理识别模式来让计算机自己学习<sup>[6]</sup>。

### 2.2 手写字识别的 CNN 模型

手写数字的识别,不像对印刷体数字的识别,印刷体数

字具有工整规范的格式,手写体的数字是不同的人写的,具有不同的写字习惯与风格,所写出来的图片处理要求也不一样,对神经网络结构要求较大。因此在手写数字的识别中构建具有表征学学习的 CNN 神经网络模型,用于训练一个用于识别数字的模型。卷积神经网络能够处理多维的数据,对于训练一个简单的数字识别模型能够很轻松地实现。

本文采用 CNN 网络处理文字图像,网络通过降采样操作来提取图片的特征信息,再经过上采样处理提高图像的分辨率。为了确保特征信息的完整性,同时降低计算量、提高网络训练速度,网络采用局部连接。CNN 网络在处理手写数字图像时,利用 ReLu 函数将输入图像经过  $3 \times 3$  的卷积层处理,后经过池化处理得到特征图,最终将输出图像经过上采样处理得到原始图像的大小。

### 2.3 基于 TensorFlow 框架的手写字识别

TensorFlow 是 Google 开源的人工智能机器学习平台,是将复杂的数据结构传输至神经网络中,并进行分析和处理的系统。它并不仅仅是一个深度学习库,只要能够把将计算过程表示成一个数据流图的过程,就可以使用该框架来进行计算<sup>[7]</sup>。TensorFlow 的核心组件包括:分发中心、执行器、内核应用和最底端的设备层/网络层。执行 TensorFlow 基本步骤为:(1) 使用 Tensor 表示数据;(2) 通过变量 (Variable) 输入训练数据,维护状态;(3) 用计算图来表示计算任务;(4) 执行计算图。

Python 语言和 TensorFlow 框架的结合来开发手写数字的识别,很大程度降低了工作量,TensorFlow 框架中有很多模块能够很轻松的构建机器学习模型。

### 2.4 手写字数据集

手写汉字识别系统,一直是一个难度数倍高于数字识别,在汉字识别研究中,数据集大,中国汉字的博大精深,据统计的汉字具有八万多个,我们常用的汉字高达 3 000 多个,看似不多的 3 000 多个汉字如果能让计算机能够识别,需要的是数倍的训练集,如此大的训练集需要的一个结构十分复杂精巧的神经网络结构,才能训练出高识率的数据集,手写汉字的研究也不比印刷体汉字的研究,手写汉字的结构多变,写法多变,有的人写得比较正规,有的人写得比较潦草,要想多对这些手写字都能够进行识别,需要足够大的数据集来训练可靠的神经网络模型。在当今社会虽然语言分很多种但是阿拉伯数字确是通用的,人工智能的发展推动着社会的进步,手写数字的识别不在仅限于印刷体的识别,手写体数字的识别也需要技术性的突破。本文选用中科院自动化研究院的 HWDB1.0 中文手写数据集,其中有 3 866 个中文汉字训练集图片,如图 2 所示。

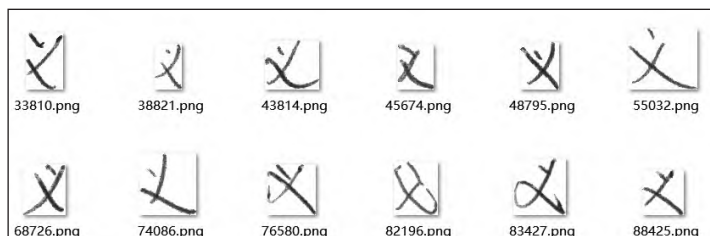


图 2 数据集中训练集图片

## 2.5 算法过程

下文具体论述算法的应用过程:

(1) 线性回归。满足公式  $\text{evidence}_i = \sum_j W_{i,j} x_j + b_i$ , 其中  $w$  表示权重,  $i, j$  的取值范围为  $[1: 784]$ , 亦可表示为:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \text{softmax} \begin{pmatrix} W_{1,1}x_1 + W_{1,2}x_2 + W_{1,3}x_3 + b_1 \\ W_{2,1}x_1 + W_{2,2}x_2 + W_{2,3}x_3 + b_2 \\ W_{3,1}x_1 + W_{3,2}x_2 + W_{3,3}x_3 + b_3 \end{pmatrix}$$

(2) 执行 softmax 函数。softmax 函数表示为  $S_i = \frac{e^{V_i}}{\sum_j e^{V_j}}$ , 线性回归的结果是离散值。举例说明: 通过加权求和得到  $(8, 2, 1, 2, 1, 3, 2, 0, 2, 2)$ , 通过 softmax 函数计算得到  $(8/23, 2/23, 1/23, 2/23, 1/23, 3/23, 2/23, 0/23, 2/23, 2/23)$  概率结果。

(3) 编程实现 MNIST。为了高效编程, 在 python 环境安装 numpy:

```
conda install -n python35tf numpy
import tensorflow as tf
```

```
x = tf.placeholder("float", [None, 784])
```

$x$  是一个占位符 (placeholder), 该张量的形状是  $[None, 784]$ ,  $None$  表示此张量的第一个维度可以是任何长度的, 可以输入任意数量的 MNIST 图像。

```
W = tf.Variable(tf.zeros([784,10]))
```

```
b = tf.Variable(tf.zeros([10]))
```

将  $tf.Variable$  赋值为零的张量来初始化  $W$  和  $b$ , 其初值可以随意设置, 其中  $W$  的维度是  $[784, 10]$ ,  $b$  的形状是  $[10]$ 。

```
y = tf.nn.softmax(tf.matmul(x,W) + b)
```

(4) 训练模型。将以上三个步骤, 关键算法用 python 实现, 并且训练 HWDB 数据集, 系统运行过程、损失率及识别结果如图 3、图 4 所示。

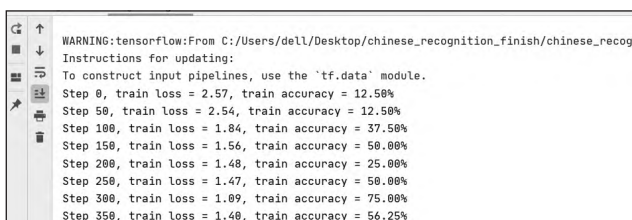


图 3 训练过程及损失率



图 4 训练结果

## 3 手写字识别系统的结果及分析

### 3.1 系统运行环境

系统操作系统为 Windows 10, 内存: 4 GB, 处理器: intel i5 7200, GPU: GoForce 940MX, 软件: Python3.8,

Opencv4.5.1, TensorFlow 等。

### 3.2 数据集

数据集分别采用 MNIST 数据集和 CASIA-HWDB1.0 数据集。将训练集和测试集的数据输入到 CNN 网络当中, 并经过多次实验得到最终的训练模型。

### 3.3 实验结果分析

通过试验一到实验三, 完成了对手写数字的识别, 并成功识别了数字 19, 237, 6 540, 如表 1 所示, 并且可以看出识别率较高, 通过一个简单的神经网络模型训练了 MNIST 库, 再用 Pygame 制作了一个画板, 在画板上我么你可以任意书写数字一到九, 系统会对手写的数字进行识别。三个实验分别识别了数字 0 ~ 9, 且都完成了识别的任务。

表 1 数字识别

名称	识别数字	准确率
实验一	19	100%
实验二	237	100%
实验三	6 540	100%

手写汉字识别系统中选取了中科院的中文数据集 HWDB1, 通过搭建简单的卷积神经网络对其中的 100 个汉字进行了训练, 使用 flask 框架搭建一个微服务器, 在对数据集中的汉字进行训练的时候我们可以看到随着训练步数的增多, 损失函数在不断地降低, 识别率也在不断地提高。并逐渐趋于一个稳定的识别概率。通过三次实验对三个不通的汉字进行了识别, 如表 2 所示, 识别率分别达到了 97%, 92%, 99%, 在三次的实验中很好地完成了对于已经训练的汉字识别, 且识别率都达到了 92% 以上。

表 2 汉字识别

名称	识别汉字	准确率
实验一	万	97%
实验二	丈	92%
实验三	与	99%

## 4 结 论

手写识别作为最自然的人机交互方式, 应用前景广泛, 本文对基于深度学习的手写文字识别进行研究, 经过多次实验训练, 得到识别准确率较高的训练模型。为了确保图片特征信息的完整性, 同时降低网络的计算量, 在 CNN 网络当中使用了局部连接。实验结果表明, 该方法对手写数字的识别准确率均值可达 95%, 对手写汉字的识别准确度均值可达 90%, 能较准确的识别手写字体。

### 参考文献:

- [1] 杨洁. 基于深度学习的手写汉字识别技术研究 [D]. 沈阳: 沈阳师范大学, 2019.
- [2] 张茹玉. 基于卷积神经网络和度量学习的脱机手写汉字识别 [D]. 上海: 华东师范大学, 2018.
- [3] 张达峰. 基于深度卷积神经网络的文字识别算法研究 [D]. 贵阳: 贵州大学, 2019.
- [4] 孙巍巍. 基于深度学习的手写汉字识别技术研究 [D]. 哈尔滨: 哈尔滨理工大学, 2017.



表 2 系统的数据设计示意

名称	系统描述	性能指标要求	系统完成功能
低慢小目标防御系统	执行低慢小目标防御系统的总体	探测指标和拦截指标满足需求方具体要求	全部功能
网弹拦截系统	实施网弹拦截的系统	拦截距离 20 m ~ 150 m, 拦截概率不小于 95%	拦截参数设置, 拦截控制, 拦截设备发射装订
激光拦截指控系统	激光拦截系统的指控单元	拦截概率不小于 95%	拦截参数设置, 拦截控制, 拦截设备发射装订
无人机拦截指控系统	无人机拦截系统的指控单元	拦截概率不小于 95%	光电探测, 雷达探测, 拦截设备信息共享, 空海节点通信, 情报信息预处理, 态势显示, 信息融合
探测系统	实施探测的系统	标准气象条件下, 探测概率不小于 95%	光电探测, 雷达探测, 信息融合
低慢小目标防御指控系统	低慢小目标防御实施指挥控制的单元, 包括人和信息系统	信息融合和指挥命令下达的时间不大于 1 分钟	拦截区域管理, 拦截设备后勤保障, 态势生成, 信息融合, 拦截方案生成, 威胁判断
.....	.....	.....	.....

5 结 论

本文针对低慢小目标防御系统设计中缺乏统一规范的问题, 研究了低慢小目标防御系统的体系结构框架, 该框架基于 DoDAF, 由全视图、作战视图、系统视图、技术视图组成。结合低慢小目标防御系统体系结构核心要素分析, 本文提出了低慢小目标防御系统的作战视图设计过程和系统视图设计过程, 并设计了作战视图和系统视图中的几类关键视图产品, 包括高级概念图 (OV-1)、作战活动分解树 (OV-5a)、作战活动模型 (OV-5b)、作战节点连接关系图 (OV-2)、作战规则模型 (OV-6a)、系统功能描述 (SV-4)、系统组成描述 (SV-1a)。通过对低慢小目标防御系统体系结构的规范化设计, 可为低慢小目标防御系统研制和集成应用提供指导和技术支持。下一步研究将关注于低慢小目标防御系统体系结构设计产品的评估与优化等方面。

参考文献:

[1] ZHOU C, LIU Y, SONG Y Y. Detection and tracking of a

UAV via hough transform [C]//2016 CIE International Conference on Radar (RADAR).Guangzhou: IEEE, 2017: 1-4.

[2] WU Y F, SUN H J, LIU P X. A novel fast detection method of infrared LSS-Target in complex urban background [J].International Journal of Wavelets, Multiresolution and Information Processing, 2018, 16 (1) : 1-17.

[3] 吕敬. “低慢小” 无人机威胁与防范 [J]. 现代世界警察, 2017 (7) : 92-93.

[4] 韩晓飞, 蒙文, 李云霞, 等. 激光防御低慢小目标的关键技术分析 [J]. 激光与红外, 2013, 43 (8) : 867-871.

[5] 何加浪, 傅军, 王辉. 基于元模型的体系结构设计方法概念辨析 [J]. 指挥信息系统与技术, 2016, 7 (1) : 1-6.

[6] 张志鹏, 苏中. 拦截低慢小目标的指控系统建模与仿真 [J]. 系统仿真学报, 2018, 30 (11) : 4340-4347+4358.

作者简介: 王飞 (1984—), 男, 汉族, 山东寿光人, 高工, 博士研究生, 主要研究方向: “低慢小” 目标防控。

(上接 91 页)

[5] 苏日娅. 基于深度学习的手写汉字识别的研究 [D]. 呼和浩特: 内蒙古大学, 2019.

[6] 齐照辉. 基于 TensorFlow 的卷积神经网络应用 [D]. 武汉: 武汉大学, 2018.

[7] 周甲甲. 基于深度学习的汉字识别方法研究 [D]. 武汉: 华

中师范大学, 2019.

作者简介: 万茹月 (1993.06—), 女, 汉族, 新疆乌鲁木齐人, 讲师, 硕士, 研究方向: 嵌入式系统; 通讯作者: 刘文 (1982.02—), 男, 汉族, 新疆乌鲁木齐人, 教师, 博士, 研究方向: 人工智能、大数据。