

武汉理工大学

本科生毕业设计（论文）开题报告

学 生 姓 名： 丁 聪

导师姓名、职称： 黄龙庭 讲师

所 属 学 院： 信息工程学院

专 业 班 级： 电信 1704

设计（论文）题目： 基于深度学习的手写汉字数字识别算法研究与实现

2021 年 3 月 22 日

开题报告填写要求

1. 开题报告应根据教师下发的毕业设计（论文）任务书，在教师的指导下由学生独立撰写。
2. 开题报告内容填写后，应及时打印提交指导教师审阅。
3. “设计的目的及意义”至少 800 汉字（外语至少 500 字），“基本内容和技术方案”至少 400 汉字（外语至少 200 字）。进度安排应尽可能详细。
4. 指导教师意见：学生的调研是否充分？基本内容和技术方案是否已明确？是否已经具备开始设计（论文）的条件？能否达到预期的目标？是否同意进入设计（论文）阶段？

撰写内容要求（可加页）：

1. 目的及意义（含国内外的研究现状分析）

随着 alphago 在围棋比赛中击败韩国围棋大师李世石，便有众多研究者加入神经网络的研究当中。深度学习的应用作为人工智能领域的一个重要研究分支，给予了计算机识别图像犀利的“眼睛”和复杂的大脑。神经网络模拟人的大脑活动的过程，尽可能地去模拟一个人的思考行为，赋予了计算机意想不到的功能。但是，因此 90 年代的计算能力的限制，神经网络曾经停滞不前，直到显卡的计算能力大幅上升，并将并行计算运用在训练神经网络上，深度的神经网络才得以实际应用，并发挥意想不到的作用。

神经网络的最初提出者 Lecun 在 1989 年就实现了第一个卷积神经网络，他在 06 年的《Notes on Convolutional Neural Networks》论文当中详细描述卷积神经网络结构^[1,2,3]，包括卷积层、池化层以及全连接层，每一层都有众多的权值、偏置参数；关于参数的自动更新，他引用了反向传播、自动梯度计算的概念来解释神经网络如何通过损失函数来自动更新每一个可变参数的值，以达到最终识别图像的目的。目前，国外众多研究者将神经网络应用于图像的分类问题当中，并且在改进自己的网络当中^[4,5,6]，引用了很多有名的图像分类数据集，如 ImageNet 、CIFAR-10 、MNIST 等等，其中 MNIST 一类是关于手写阿拉伯数字，如今的识别率已经达到 99%以上^[7]。

对于数字的识别研究如火如荼，但对于中文的数字以及某些特定文字数据集的研究近期缺比较少，并且对于中文数字识别的研究属于汉字识别的研究范畴，许多研究从上世纪 90 年代已经开始，但其研究方法都是基于传统的特征分解^[8,9]，如提取汉字的笔画密度特征、外围特征，亦或者使用算子进行方向分解^[10]，找到特定的比划特征如横、竖等，再将特征放入支持向量机进行分类^[11,12]。该类方法识别的效果不差，但是比起神经网络的准确度，其效果远远不及。

通过研究神经网络在图像分类的良好效果，将其运用到中文的数字识别，准确度一定能够更上一个层次，而计算力的几何倍数的提升也让数字识别的时间大大减小，时间上的劣势不再是问题。每年都有应用在图片分类的良好的神经网络诞生，倘若能够测试到比较好的网络，应用在中文数字识别上^[13]，那么文字的识别系统准确度将会随之得到极大的改良^[14,15]。

2. 研究（设计）的基本内容、目标、拟采用的技术方案及措施

2.1 研究（设计）的基本内容、目标

1. 学习卷积神经网络的基础知识，数据集训练方法以及训练模型存储。
2. 学习数字图像处理的理论，如图像的形态学运算、图像分割、二值化处理等，用于手写数字图像数据的预处理。

3. 学习 python 相关基础语法知识和 Tensorflow 神经网络框架的使用。
4. 开发一个轻量的 python GUI 用于测试、检验。

其总体识别流程如下所示：

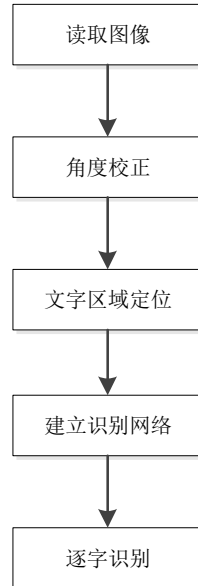


图 1 识别流程图

2.2 拟采用的技术方案及措施

1. 卷积神经网络（CNN）

（1）神经元

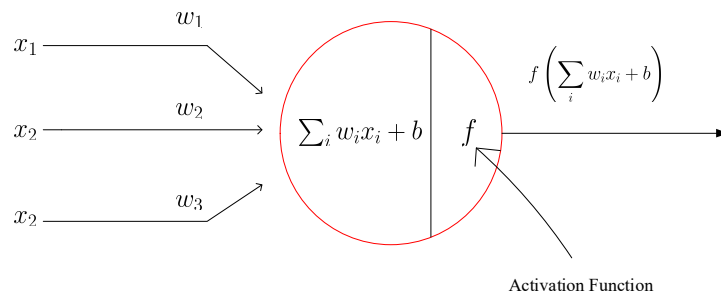


图 2 神经元

神经元是神经网络的基本单元，由一个或多个输入，图 1 当中拥有三个输入 x ，每一个输入都对应不同的权重 w ，把输入的三个信号累加之后加上预设偏置 b ，便可以得到神经元的值，这个值送入激活函数（如 ReLU），若达到某一个值，便可以触发神经元传递信号到下一个神经元。

（2）卷积神经网络（CNN）

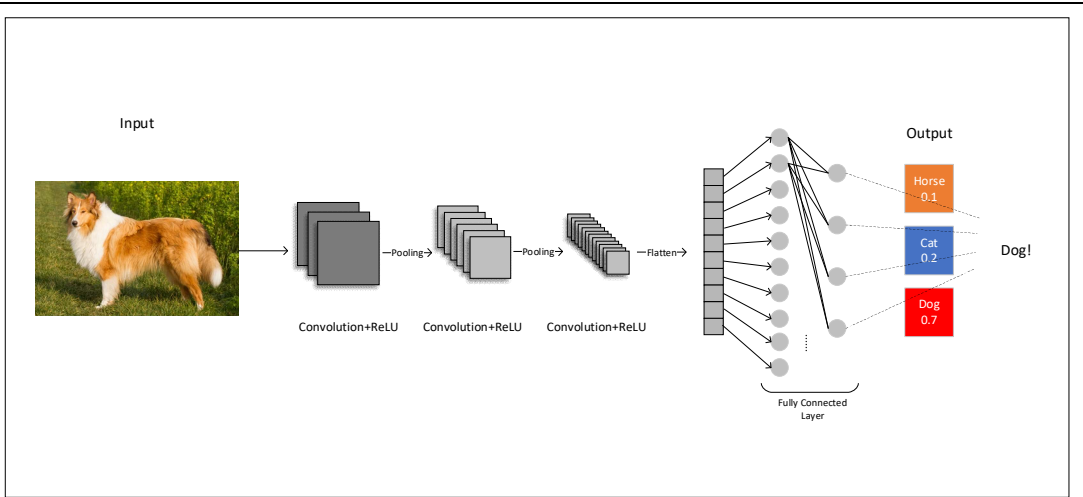


图 3 卷积神经网络

CNN 加入图像处理当中常用的卷积运算，使得神经网络能够模拟人观察图片某些特征的功能，CNN 一般由卷积层、池化层、全连接层构成，多次的卷积使得图像维度缩小，提取到图像高维特征，高维特征再经过全连接层，便能以概率的形式呈现出分类的可能。

（3）前向传播机制

假设 CNN 中第 $L-1$ 层的第 j 个神经元传向第 L 层的第 i 个神经元的权重参数用 $w_{ij}(L)$ 来表示，第 L 层的第 i 个神经元偏置用 $b_i(L)$ 来表示，自动学习率用 α 表示，整个网络的均方误差函数用 E 来表示，那么，每个神经元的参数 w ， b 都可以用前向传播公式来表示：

$$w_{ij}(L) = w_{ij}(L) - \alpha \frac{\partial E}{\partial w_{ij}(L)} \quad (1)$$

$$b_i(L) = b_i(L) - \alpha \frac{\partial E}{\partial b_i(L)} \quad (2)$$

（4）减少网络训练过拟合

神经网络容易训练过拟合从而使得识别效果差强人意，使用 Batch Normalization 技术将特征归一化可以减轻过拟合现象，并且加快学习进程。其原理是将每一批训练的输入数据进行归一化处理，求其均值和方差，算出归一化后的输入，然后在利用线性缩放和位移来调整。此外，Dropout 技术的发展也可以比较有效的缓解过拟合的发生，在一定程度上达到正则化的效果。Dropout 技术是神经网络在计算参数过程当中，以一定的概率让某一个神经元失效，从而使得模型的泛化能力更强。与机器学习当中的继承学习有着异曲同工之处，即使模拟多个不同的网络对图像进行观察，以防止某一个网络的学习程度过深而造成识别错误。

3. 进度安排

第 1 周一第 3 周 搜集资料，撰写开题报告；

第 4 周一第 5 周 论文开题；

第 6 周一第 12 周 撰写论文初稿；

第 12 周一第 15 周 修改论文；

第 16 周 论文答辩。

4. 阅读的参考文献不少于 15 篇（其中近五年外文文献不少于 3 篇）

[1] Gonzalez R C. Deep convolutional neural networks [lecture notes][J]. IEEE Signal Processing Magazine, 2018, 35(6): 79-87.

[2] LeCun Y, Bengio Y, Hinton G. Deep learning. nature (2015)[J]. May; 521 (7553): 436 10.1038/nature14539, 2015.

[3] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[J]. Advances in neural information processing systems, 2012, 25: 1097-1105.

[4] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv:1409.1556, 2014.

[5] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.

[6] Ciresan D C, Meier U, Gambardella L M, et al. Convolutional neural network committees for handwritten character classification[C]//2011 International Conference on Document Analysis and Recognition. IEEE, 2011: 1135-1139.

[7] LeCun Y, Boser B, Denker J S, et al. Backpropagation applied to handwritten zip code recognition[J]. Neural computation, 1989, 1(4): 541-551.

[8] 金连文. 手写体汉字识别的研究[D].华南理工大学,1996.

[9] 朱辉, 杨扬, 颀斌, 封筠.SVM 在小字符集手写体汉字识别中的应用研究[J].微计算机信息,2004(04):74-75.

[10] Gonzalez R C, Woods R E. Digital Image Processing. 4th ed. New York: Pearson-Prentice Hall, 2018.

[11] 周志华. 机器学习[M]. 北京：清华大学出版社，2016.

[12] 李航. 统计学习方法[M]. 北京：清华大学出版社，2012.

[13] 关保林, 巴力登. 2013. 基于改进遗传算法的 BP 神经网络手写数字识别[J]. 化工自动化及仪表, 40(6): 774-778.

[14] 梁淑芬, 胡帅花, 秦传波, 等. 2017. 基于深度学习的数字识别模块在安卓系统的

实现[J]. 五邑大学学报(自然科学版), 1(1):40-45.

[15] 曹丹. 基于 Hopfield 神经网络的脱机手写数字识别[D]. 中南大学. 2009

5. 指导教师意见

指导教师（签名）：_____

年 月 日

