

基于传统机器学习与深度学习的图像分类算法对比分析

刘华祠

(长江大学电子信息学院,湖北 荆州,434023)

摘 要:深度学习作为机器学习的一个分支,具有强大的功能和灵活性,文章对机器学习:K 邻近,贝叶斯,支持向量机,与深度学习:卷积神经网络,递归神经网络在大样本手写数字识别与小样本图像场景分类效果上进行了对比与分析,实验结果表明:传统机器学习在小样本数据集上具有较好的解决效果,深度学习框架在大样本数据集上具有较高的识别精度。

关键词:机器学习;深度学习;图像分类

中图分类号:TP391.41 **文献标识码:**A

Comparative Analysis of Image Classification Algorithms Based on Traditional Machine Learning and Deep Learning

LIU Hua-ci

(School of Electronics and Information, Yangtze University, Jingzhou 434023, China)

Abstract: Deep learning, as a branch of machine learning, has powerful functions and flexibility. This paper is about machine learning: K-near, Bayesian, support vector machine, and deep learning: convolution neural networks, recursive neural networks in large The sample handwritten digit recognition and the small sample image scene classification effect are compared and analyzed. The experimental results show that the traditional machine learning has a good solution effect on the small sample data set, and the deep learning framework has a wide range on the large sample data set. Applicability.

Key words: machine learning; deep learning; image classification

随着人工智能的来临,深度学习在各大领域占据着最为主要的成分。传统的图像识别技术是以浅层次结构模型为主,需要人为对图像进行预处理,导致降低图像识别的准确率^[1]。为了提高图像识别精度,深度学习模型结构被提出,如:DBN,GAN,CNN,RNN等。但是,这些深度学习模型结构等都需要大规模数据集的训练。传统的机器学习图像识别模型,K 邻近(KNN)、贝叶斯网络、支持向量机(SVM)等,仍然具备各种优势。本文通在两种数据集上对比分析,分别采用传统的机器学习和深度学习的方法验证精度和时间,得出机器学习与深度学习各自的优缺点。通过实验,分析深度学习和传统的机器学习的优点与不足。

1 传统机器学习

1.1 本文特征提取方法

传统机器学习方法中,除了分类方法之外,提取图像特征向量也是一个重要的研究课题。常用的特征有:

SIFT 特征^[2],Harris 角点检测^[3]等,对于图像特征的提取具有较好的效果。本文采用图像的直方图方法提取图像的特征向量,为了防止过拟合,在提取特征向量引入 L2 正则化提高模型的泛化能力^[4],采用如图 1 所示的方法进行分类识别。

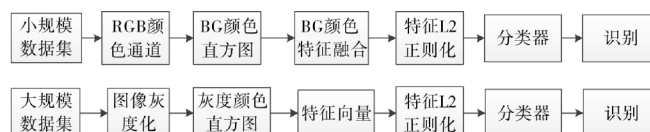


图 1 传统机器学习识别方法

1.2 K 邻近算法

KNN 通过计算待测样本数据与训练数据中不同类别数据点间的相似度进而进行待测样本分类^[5]。KNN 算法常用的距离测量公式为欧式距离。

$$d(X,Y)=\sum_{i=1}^n \|x_i-y_i\|^2 \quad (1)$$

算法的基本步骤为:

step1: 计算测试集特征向量与训练集特征向量之间的距离。

step2: 按照距离的远近进行排序。

step3: 选取距离最近的 K 个点。

step4: 计算前 K 个点所在类别的出现频率。

step5: 统计前 K 个点中出现频率最高的类别, 作为测试集的分类类别。

1.3 贝叶斯

贝叶斯分类方法是基于概率分类模型, 贝叶斯方法的分类性能主要取决于估计的条件概率项的准确性。当训练数据稀缺时, 这些条件概率项的估计精度较低, 从而导致朴素贝叶斯方法的分类性能下降^[6]。实验中采用多项式分布的朴素贝叶斯模型的分类方法:

$\hat{\theta}_{yi} = \frac{N_{yi} + \alpha}{N_{y_i} + \alpha n}$, 其中 $\hat{\theta}_{yi}$ 表示标签 i 出现的样本属于类别 y 的概率。

算法流程如下:

step1: 设 $x \in \{a_1, a_2, \dots, a_n\}$ 为一个测试集, a 为 x 的特征向量。

step2: 有类别集合 $C = \{y_1, y_2, \dots, y_n\}$ 。

step3: 计算 $P(y_1|x), P(y_2|x), \dots, P(y_n|x)$

step4: 如果 $P(y_k|x) = \max\{P(y_1|x), P(y_2|x), \dots, P(y_n|x)\}$, 则

$x \in y_k$ 。

1.4 支持向量机

支持向量机(SVM)通过构造超平面在高维空间中实现原空间中的非线性决策。SVM 是基于统计学习理论提出的一种解决小样本分类问题的学习算法, 能克服传统的人工神经网络等算法的欠和过学习的不足^[7]。SVM 采用了由简单到复杂的分类过程, 具有很好的泛化能力。

其相关算法步骤为:

step1: 将多分类问题转变成凸优化问题。

求取 $\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^R \zeta_i$ 的最优解, 其中 C 为惩罚系数, 采用默认值 1.0, 使得

$$\begin{cases} y_i(w \cdot x_i + b) \geq 1 - \zeta_i \\ \zeta_i \geq 0 \end{cases}, i=1, \dots, n \quad (7)$$

step2: 定义拉格朗日目标函数:

$$L(w, b, a, \zeta) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^n a_i (w \cdot x_i + b) - 1 + \zeta_i \quad \begin{cases} \frac{\partial L}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^n a_i y_i x_i \\ \frac{\partial L}{\partial b} = 0 \Rightarrow b = \sum_{i=1}^n a_i y_i \end{cases} \quad (8)$$

代入上式, 并求其对偶问题:

$$\min \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N a_i a_j y_i y_j K(x_i, x_j) - \sum_{i=1}^N a_i \quad (9)$$

使得: $\sum_{i=1}^N a_i y_i = 0, 0 \leq a_i \leq C, i=1, 2, \dots, N$

step3: 求最优解 a^*

step4: 求 $b^*, b^* = y_j - \sum_{i=1}^N a_i^* y_i K(x_i, x_j)$

$$\text{求 } w^*, w^* = \sum_{i=1}^N a_i^* y_i x_i \quad (10)$$

step5: 得到最优分类函数 $f(x) = \text{sign}(w^* \cdot x + b^*)$ (11)

2 深度学习

2.1 卷积神经网络

卷积神经网络一般由数据输入层、卷积层、池化层、全连接层、输出层组成^[8]。卷积操作过程: 将矩阵相对应的元素与卷积核对应的元素相乘, 然后把所有相乘后的元素相加, 把相加后的结果赋值给中心像素。深度学习中卷积层能自动提取图像的特征, 故深度学习避免繁琐的提取某一固定特征的要求。卷积操作过程如图 2 所示:

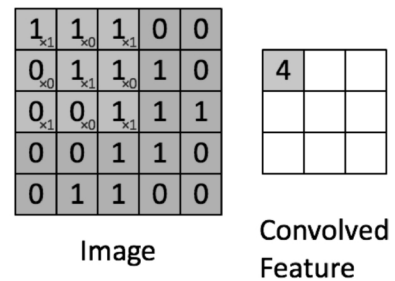


图 2 卷积运算

本文使用的网络包含 2 个卷积层、2 个池化层、1 个全连接层。为了减小网络的计算量和提取图像细微的特征, 第一层的卷积采用 5×5 的卷积核, 第二层的卷积采用 3×3 的卷积核; 池化层采用 2×2 的池化核; 采样过程取步长为 2, 边界填充模式为 'same' 方式。本次卷积神经网络实验中: 损失函数为交叉熵损失, 激活函数为 relu, 分类函数为 softmax。交叉熵损失函数为:

$$L_s = -\frac{1}{n} \sum_{i=1}^n [y_i \ln a + (1 - y_i) \ln (1 - a)] \quad (12)$$

卷积网络模型如图 3 所示。

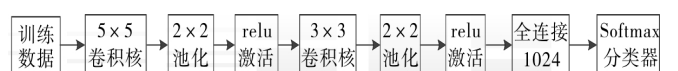


图 3 卷积网络模型

2.2 递归神经网络

递归神经网络(RNN)中常用的模型为 LSTM(长短期记忆网络),其模型结构如图 4(a)所示。LSTM 神经单元共有三个门:输入门,遗忘门,输出门,其结构单元如图 4(b)所示。通过对门的控制,LSTM 就可以决定输出的信息变化。构建这种神经网络结构的目的是为了解决较长序列在训练过程中由于链式求导产生的梯度消失或者梯度爆炸问题^[9]。

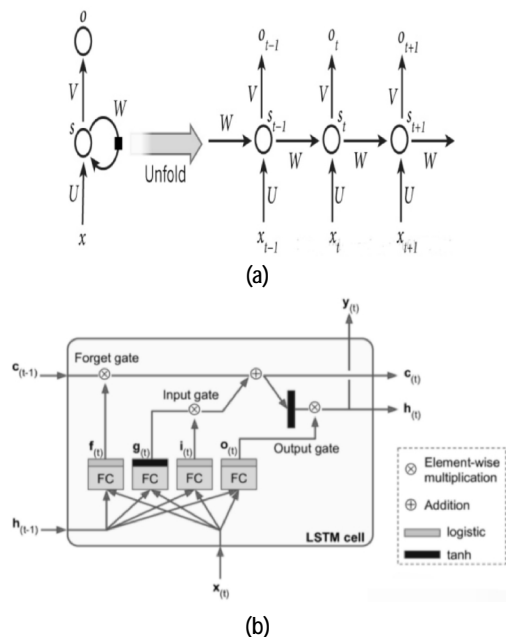


图 4 递归神经网络

本次实验中采用递归神经网络模型中的 LSTM 网络,其隐藏层神经单元个数为 100。输入为图像的宽(W),序列数为图像的高(H),损失函数为交叉熵损失函数。使用动态 RNN 处理时序数据,分类函数为 softmax。网络模型如图 5 所示:

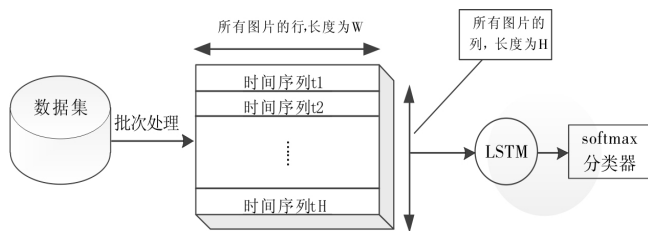


图 5 递归神经网络模型

3 实验分析

大样本数据集采用 mnist 手写数字图片数据集^[10],共有 10000 张 28×28 格式的图片,分别是 0-9 的手写体数字,各类数字图片的张数如图 6 所示,实验中选取 2000 张作为测试集,8000 张作为训练集。



图 6 mnist 手写数字图片

小样本数据集采用 COREL1000 图片集^[11],共有 1000 张 384×256 和 256×384 格式的图片,实验过程中全部调整为 96×96 大小。分别是部落,海岸,建筑,车辆,恐龙,大象,花朵,骏马,山峰,食品十类,每类 100 张图片,如图 7 所示。分别对这十类图片进行标记,依次标记为 0-9,实验中选取 200 张作为测试集,800 张作为训练集。



图 7 COREL1000 图片

本次实验所运行的环境为:ubuntu 16.04LTS, python 的版本为 Python 3.5, 处理器型号为 Intel(R) Core (TM) i3-4150 CPU @ 3.50GHz 4 核,运行内存为 6G,深度学习框架无 GPU 版的 tensorflow,其版本为 1.8.0。

当采用小样本 COREL1000 数据集时,各个分类模型的分类效果见表 1。

表 1 COREL1000 分类结果

分类器参数	精度	时间
KNN	0.50	47s
Bayes	0.66	24s
SVM	0.85	1m2s
CNN	0.82(11 次)	2m06s
LSTM	0.76(20 次)	5m28s

当采用大样本 mnist 数据集时,各个分类模型的分类效果见表 2。

表 2 mnist 分类结果

分类器参数	精度	时间
KNN	0.86	49m03s
Bayes	0.77	3m19s
SVM	0.88	27m45s
CNN	0.97(11 次)	21m15s
LSTM	0.94(20 次)	57m58s

4 结束语

通过实验数据对比分析可知,在小样本数据分类上,机器学习在图像分类识别上具有简单高效的优点,并能解决高像素图片的快速分类。在大样本数据上,深度学习模型能进行精准的分类识别,并且具备较强的鲁棒性。在传统机器学习模块中,贝叶斯分类器模型能够快速的做出判断,但不具备较好的分类能力;SVM 具备较强分类能力,但计算耗时。深度学习中卷积神经

网络具备较好的提取图像特征的能力,但模型训练复杂度较高;递归神经网络在图像分类识别方面明显低于卷积神经网络,但更适合处理序列问题。在小规模数据集上,传统的机器学习模型更具备分类优势,在大规模数据以及识别精度上,深度学习模型更具备更好的能力。

参考文献:

- [1] 郑远攀,李广阳,李晔.深度学习在图像识别中的应用研究综述[J].计算机工程与应用,2019,55(12):20-36.
- [2] 李鹤宇,王青.一种具有实时性的 SIFT 特征提取算法[J].宇航学报,2017(8):856-871.
- [3] HARRIS C, STEPHEN M.A combined corner and edge detect or [C]//Proceedings of the Fourth Alvery Vision Conference. Manchester:The University of Sheffield Printing Unit, 1988: 207-217.
- [4] 李建伟,郑永果.基于 CNN 的动物识别研究.软件导刊,

2019,18(3):30-33.

- [5] 文武,李培强.基于 K 中心点和粗糙集的 KNN 分类算法[J].计算机工程与设计,2018,39(11):97-102.
- [6] 杨晓花,高海云.基于改进贝叶斯的书目自动分类算法[J].Computer Science,2018,45(8):203-207.
- [7] 杨梅芳,石义龙.基于 2DPCA+PCA 与 SVM 的人脸识别[J].信息技术,2018.
- [8] 林恒青.基于深度卷积神经网络的脱机手写汉字识别系统的设计与实现[J].湖北理工学院学报,2019,35(02):31-34.
- [9] 汤梦,朱杰.一种基于 LSTM 的合成语音自然度评价方法的研究[J].信息技术,2019(05):41-44.
- [10] 黄睿,陆许明,邹依林.基于 TensorFlow 深度学习手写体数字识别及应用[J].电子技术应用,2018,44(10):12-16.
- [11] Gong M, Hao Y, Mo H, et al. Naturally combined shape-color moment invariants under affine transformations [J]. Computer Vision and Image Understanding, 2017:S1077314217301376.

(上接第 11 页)

4 结束语

本文设计并实现了一款基于 LL1 文法的解释器,给出了模块及详细的流程设计,并详细分析了各个模块组件的实现原理及方法,经测试,该解释器能够正确有效的工作,并且能够准确的识别出用户程序中的各种错误。笔者下一步的工作是进一步完善该解释器的用户接口和使用 JIT(动态编译技术)来进一步加快块解释器的运行速度。

参考文献:

- [1] Aho A V, Lam M S, Sethi R, et al. Compilers: Principles, Techniques, and Tools (2nd Edition)[M]. Addison Wesley,2006.
- [2] Cooper K, Linda Torczon, 库珀, et al. 编译器设计 [M]. 人民

邮电出版社,2013.

- [3] 张素琴.编译原理.第2版[M].清华大学出版社,2005.
- [4] Deremer F L. Lexical analysis [J]. Lecture Notes in Computer Science, 1974, 21(2):109-120.
- [5] 邹昌伟. C 编译器剖析[M]. 清华大学出版社,2016.
- [6] London R L, Shaw M, Wulf W A. Abstraction and Verification in Alphard: A Symbol Table Example[M]. Alphard: Form and Content. 1978.
- [7] Appel. Modern Compiler Implementation In Java [M]. 高等教育出版社,2003.
- [8] 徐玲,韦巍.基于 ANTLR 的 TTCN-3 编译器的面向对象设计[J].控制工程,2012(s1):122-125.
- [9] 陈庆.UEFI 系统脚本语言与解释器的研究与实现[D].华中科技大学,2011.