文章编号：2095-2163(2021)07-0162-04 中图分类号：TP317.4 文献标志码：A

基于卷积神经网络 VGG 的猫狗识别

王 彬，张正平，贾明俊，陆安江，卢学敏
（贵州大学 大数据与信息工程学院，贵阳 550025）

摘 要：随着大数据时代的到来，数据挖掘、图像处理等已经成为了一个热门研究方向。本文的研究目的是自动识别猫狗类型，采用的是基于数据挖掘的猫狗自动识别技术。本文将位于全方位下拍摄的具有外貌复杂的猫狗图像运用卷积神经网络训练。本实验挑选前沿的深度学习框架 pytorch 以及计算能力强大的 GPU，使用深度神经网络 VGG16，分别对猫狗图像进行网络训练与测试。实验显示使用 VGG16 网络模型进行识别的准确率非常高，在猫狗类型识别问题上具有突出优势。

关键词：数据挖掘；VGG16；分类；图像处理；深度学习

VGG dog and cat recognition based on convolutional neural network

WANG Bin, ZHANG Zhengping, JIA Mingjun, LU Anjiang, LU Xuemin
(College of Big Data and Information Engineering, Guizhou University, Guiyang 550025, China)

【Abstract】 With the advent of the era of big data, data mining and image processing has become the hot research direction. The purpose of this paper is to automatically identify the cat and dog types using the cat and dog automatic recognition technology based on deep learning. In this paper, convolution neural network is used to train the cat and dog images with complex appearance. In this experiment, the researchers select the advanced deep learning framework pytorch and the powerful GPU, and use the deep neural network VGG16 to train and test the cat and dog images. Experiments show that the accuracy of VGG16 network model is very high, and it has outstanding advantages in cat and dog type recognition.

【Key words】 data mining; VGG16; classification; image processing; deep learning

0 引 言

在数据挖掘中，对数据挖掘模型进行分析和设计主要是通过对算法模型的研究来付诸实现的，其中卷积池化方法是最基本的。数据挖掘包括的内容十分广泛，主要有：数据分析理论、数据预测理论、数据安全理论、数据侦测理论和数据追踪理论。数据挖掘中比较有名的数据库中的知识发现（Knowledge Discover in Database，KDD），迄至目前也仍是人工智能和数据库领域的热点研究问题。研究可知，数据挖掘的含义是指从数据库的各种数据中找出潜在的、并有隐藏价值的信息的过程。数据挖掘是一种决策支持过程，并主要基于人工智能、机器学习、模式识别、统计学、数据库、可视化技术等，高度自动化地分析企业的数据，做出归纳性的推理[1]，从中挖掘出潜在的模式，帮助决策者调整市场策略，降低风险，做出正确的决策。

1 研究背景与意义

1.1 研究背景

数据挖掘源起自上世纪 80 年代关于人工智能的投资项目夭折，从而改变战略投身实际应用。这是一种时兴的，面向商业应用的人工智能研究。选择数据挖掘这一术语，展现了与统计、精算、长期从事预言模型的经济学家之间没有技术的重复部分。数据挖掘技术主要包含 3 个部分：算法和技术；数据；建模能力[2]。对此拟做阐释分述如下。

（1）机器学习。机器学习是计算机科学和人工智能衍生的产物。机器学习分为 2 种学习方式：自组织学习（如神经网络）和从例子中总结出规律（如决策树）。

（2）统计。统计包括：预言算法（回归）、抽样、基于经验的设计等，现在也开始支持数据挖掘。

（3）决策支持系统。

（4）数据仓库。

（5）OLAP（联机分析处理）、DataMart（数据集市）、多维数据库等将数据仓库、OLAP、数据挖掘等技术融合在一起，即构成企业决策分析环境。

1.2 研究意义

目前，正处于一个大数据时代，无论是云计算、社交网络，还是物联网、移动互联网和智慧城市，都

作者简介：王 彬（1996-），男，硕士研究生，主要研究方向：图像处理；张正平（1964-），男，博士，教授，主要研究方向：无线电技术、电信技术、信息处理与控制；陆安江（1978-），男，博士，教授，主要研究方向：优化通信与信息系统。

收稿日期：2021-03-15

与大数据息息相关。大数据已然成为有着特定意义的专有名词，而不只是说数据量庞大。21 世纪的新一代信息技术有了突破创新以及广泛实用性普及，例如云计算、移动互联网和物联网等，这意味着人类社会正以高速跨进大数据时代。越来越多的行业对大数据的发展与应用都抱有积极态度，而更多的用户为了提升自身的工作效率也开始去尝试或考虑如何运用大数据类的解决方案。随着数据化的逐步推进，继传统企业三大竞争策略陆续问世后，大数据已然成为企业可以运营的第四种全新战略[3-4]。值得一提的是，社会只需要该项技术注重相关关系，摒弃因果。简单来讲就是，只要知道是什么就好，而不用了解其背后原因。故而传统惯例就此被颠覆，人们对现实世界的理解以及面对事物的判决思维方式也被质疑挑战，基于此将会换个层面、角度思考问题，从而对传统决策产生了极其重大的影响。因此人们就可从众多繁杂数据中提取需要的，有用的、关注的信息。这也就是数据挖掘的意义所在。

2　猫狗分类器设计

2.1　模型选择

研究中，给出一个 VGG 结构图，如图 1 所示。在此过程中，主要使用的是共有 16 层模型的 VGG16，该模型需要的是 224 * 224 * 3 维度的输入数据[5]。VGG 模型在 2014 年的 ILSVRC 竞赛中因表现优异而荣获第二，虽然 GoogLeNet 占据了第一名的位置，但是在多个迁移学习任务的较量中，VGG 模型的表现都比 GoogLeNet 好。此外，VGG 模型是从图像中提取 CNN 特征的首选算法。只不过 140 多兆的参数量却是其主要缺点，所需存储空间较大[6]。综合上述分析可知，VGG 模型的研究价值很大。所以本次研究选用了 VGG16 模型。

2.2　平台搭建

实验室的电脑是深度学习工作站，有着高性能的 GPU。所使用的 Ubuntu18.04 64 bit。安装软件是 Anaconda，基于 anaconda 安装 pytorch 的框架。下载 kaggle 比赛中所使用的数据集[7]。

2.3　猫狗分类器

2.3.1　图片的导入和预览

输入的图片需要分辨率为 224 * 224，如图 2 所示。为此使用 transform.CenterCrop(224) 对原始图片进行裁剪。载入的图片训练集合为 20 000 个和验证集合为 5 000 个，原始图片全部为训练集合，需要拆分出一部分验证集合，输出的 Label，1 代表狗，0 代表猫[8]。

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input(224*224 RGB image)					
conv3-64	conv3-64 LRN	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 conv1-256	conv3-256 conv3-256 conv3-256	conv3-256 conv3-256 conv3-256 conv3-256
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
FC-4096					
FC-4096					
FC-1000					
softmax					

图 1　模型的选择

Fig. 1　Model selection

2.3.2　迁移模型

研究中使用的训练集的图片都为 224 * 224 * 3，要想对猫与狗的图片识别效果更佳，那么迁移过来的 VGG16 模型就要去适应新的要求，因此本次研究就将 VGG16 全连接层的最后一部分做出了调整改动并且对参数进行了重新训练。但即便是训练整个全连接层的全部参数，计算机的运行耗时也并不

少,所以本文只是训练了全连接层的最后一层,在节约时间的同时也能得到很好的效果。这里给出了训练的部分截图如图 3 所示。

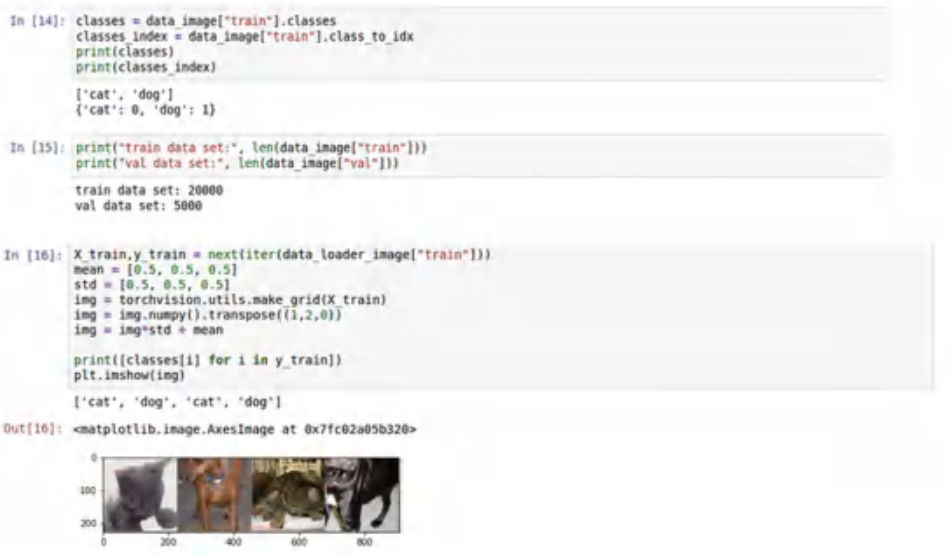

图 2 数据导入

Fig. 2 Data import

图 3 迁移模型

Fig. 3 Migration model

2.3.3 测试模型

网上的实例 CPU 进行训练,因为速度很慢,只对 100 张图片进行训练的演示,进行 1 次训练的 *Loss* 为 0.350 1, *Accuracy* 准确率为 94%。验证集的 *Loss* 为 0.9151, *Accuracy* 准确率为 88%。这只是 100 张图片的 1 次训练,更多的图片以及多次的训练可能会得到一个更好的结果[9]。测试中无需考虑此问题,因为研究中使用的训练很快。本次研究最终得到的测试结果见图 4。

研究中,在 5 000 个验证集中随机地选择了 16 张图片进行测试,可以看到,第一个是狗,虽然只有鼻子和嘴,但还是做到了正确的识别。测试的 16 张图片中,肉眼看到的是—狗、猫、猫、狗、狗、猫、猫、狗、狗、猫、猫、猫、猫、猫、猫和狗。有的图片虽然连面部都没有,但是 VGG 模型的测试结果都是正确的。由此可见,在大量训练下的准确率是很高的[10]。

图 4　模型测试

Fig. 4　Model test

3　结束语

本文所涉及的模型系统主要用到了数据挖掘中的图像处理理论—分类：这是数据挖掘中最为基本和比较成熟的一个分支，着重于研究大数据的分析和图像处理问题，其基本的分析和综合方法是卷积池化。主要对于图像特征进行提取，一层层卷积、一层层池化之后，再将每层都要进行优化。最后将每个图片信息降到一维的维度。由连接层输出。本文在实际测试过程中也发现一些问题，没有预处理数据集，对一些反光、模糊等的图片识别有误差。还有理论没有很好地结合实际。同时，对于代码移植方面暴露出的问题，也要在后续研究中进一步加以完善和改进。

参考文献

[1] 任剑岚. 数据挖掘技术应用案例的分析[J]. 信息通信，2012(6)：164.

[2] 周晟. 挖掘.com 公司—数据挖掘技术和.com 公司[J]. 软件世界，2001(6)：132-134.

[3] 维克托·迈尔-舍恩伯格，肯尼思·库克耶. 大数据时代[M]. 杭州：浙江人民出版社，2013.

[4] 马毅. 商业银行邂逅大数据：挑战与竞争战略演进[J]. 征信，2014，32(2)：75-78.

[5] 冯国徽. 基于卷积神经网络 VGG 模型的小规模图像分类[D]. 兰州：兰州大学，2018.

[6] 汤鹏杰，谭云兰，许恺晟，等. 基于 GoogLeNet 多阶段连带优化的图像描述[J]. 井冈山大学自然版，2016，37(5)：47-57.

[7] LEI Tao，ZHANG Yu，WANG S I，et al. Simple recurrent units for highly parallelizable recurrence[C]//Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. Brussels，Belgium：EMNLP，2018：4470-4481.

[8] 王鹏伟. 基于多尺度理论的图像分割方法研究[D]. 合肥：中国科学技术大学，2007.

[9] TOMAŠEV N，MLADENIC D. Class imbalance and the curse of minority hubs[J]. Knowledge-Based Systems，2013，53：157-172.

[10] 李华胜，杨桦，袁保宗. 人脸识别系统中的特征提取[J]. 北京交通大学学报，2001，25(2)：18-21.

(上接第 161 页)

[9] XIAO P，XING H，XIAOFENG M. Privacy preserving towards continuous query in location-based services[J]. Journal of computer research and development，2010，1：18.

[10] XUE M，KALNIS P，PUNG H K. Location diversity：Enhanced privacy protection in location based services[C]//International Symposium on Location- and Context-Awareness. Berlin/Heidelberg：Springer，2009：70-87.

[11] HAN B，LIU L，OMIECINSKI E. Road-network aware trajectory clustering：Integrating locality，flow，and density[J]. Mobile Computing，IEEE Transactions on，2015，14(2)：416-429.

[12] GUSTAV Y H，WANG Y，KAMENYI D M，et al. Query privacy preservation in location based services on road networks[J]. Journal of Computational Information Systems，2014，10(12)：5255-5263.

[13] 黎孟. 基于语义多样性的位置隐私保护算法研究[D]. 长沙：湖南大学，2018.

[14] KAMENYI D M，WANG Yong，ZHANG Fengli，et al. Authenticated privacy preserving for continuous query in location based services[J]. Journal of Computational Information Systems，2013，9(24)：9857-9864.

[15] Figshare[EB/OL]. [2018-12-06]. https://figshare.com/articles/Urban_Road_Network_Data/2061897.

[16] 潘晓，肖珍，孟小峰. 位置隐私研究综述[J]. 计算机科学与探索，2007，1(3)：268-281.