**基于卷积神经网络的动物识别系统的设计与实现\_第三组**

题 目： 基于卷积神经网络的动物识别系统的设计与实现

学院： 软件学院 专业： 软件工程 学生姓名： 学号：

|  |  |  |  |
| --- | --- | --- | --- |
| **一、项目背景**  1. **项目来源**  本团队在北京交通大学学习了三个月（2022年4月-2022年7月）期间参与了学校实际需求的项目“基于卷积神经网络的动物识别系统的设计与实现”。  2. **项目意义**  目前，由于计算机能力和相关理论的发展获得了重大突破，基于卷积神经网络的图像检测与识别技术已经广泛应用到人们的生产生活中。学长将深度学习的技术应用到野生动物图像识别中，优化了传统的识别方法，形成对野生动物图像更为准确的识别，为实现高效的野生动物图像识别提供了可能。不同于传统的野生动物识别，基于深度学习的野生动物识别技术可以捕获到野生动物更加细致的信息，有利于对野生动物进行更加准确的识别和研究。因此，对基于深度学习的野生动物识别和研究，可以更好的帮助社会管理者和政府全面有效的对野生动物进行保护和监管，这也正是保护和识别野生动物的关键，同时这对整个自然和社会的和谐发展具有极大的推动作用。  3. **项目的研发现状（目前同类工作中存在哪些问题，即文献综述）**  目前，动物识别系统在畜牧业等行业中，仍具有很大的使用价值，虽然在我们用于构建动物识别系统的百度Paddle-Paddle平台上也存在过许多类似的项目，但是这些项目很多仍存在以下问题：  （1）使用较为简单或者过时神经网络模型。大部分学者仍在使用经典的SVM[11]、KNN、BP算法等去构建系统的核心神经网络模型，这些经典算法也能取得较好的分类识别效果[12]，然而使用这些传统算法进行模型训练需要花费更长的时间，同时训练出来的模型较为笨重，不够轻量化，不适合模型跨平台使用。部分学者使用CNN等深度学习模型进行图像识别分类[13]，但在模型训练过程、函数优化上仍存在有较大的提升空间。  （2）数据集处理方式简陋，没有充分考虑到现实中的复杂情况。参考的许多图像识别项目仅对数据图集进行基本的剔除无效数据、裁剪图片为统一大小格式、重命名工作，对数据集的处理方式较为单一。  （3）模型预测的准确率较低。许多模型缺乏必要的优化，在少数种类的动物分类识别中，能维持较高的准确率，但当种类增多的情况下大多数仅能维持60%~80%的准确率。 | | | |
| **二、项目内容、项目目标与拟解决的关键问题**  1. **项目的内容（范围）**  基于卷积神经网络的动物识别系统是部署在AIStudio上的动物识别系统，以动物图片数据集作为卷积神经网络模型的输入，通过模型训练调整参数后得到最终模型，最终模型以动物图片为输入，动物种类作为输出，实现动物识别。  系统的主要内容包括：  （1）实验数据集处理。系统通过对压缩包进行解压，得到初始动物图片数据集，将初始数据集进行标签标注、划分，然后将数据集进行规范化处理后得到网络输入的规范数据集。  （2）网络模型搭建与训练。系统会根据已选卷积神经网络的结构组成搭建模型组网，然后实例化卷积神经网络用于训练。之后系统会默认设置超参、损失函数、优化器，然后进行资源配置，使用正则化提升精度。训练过程中对数据进行记录验证，用于最终训练。  （3）模型测试和预测。系统将测试集作为最终模型的输入，通过神经网络评价指标来对模型进行评估测试。如果测试结果良好就运行进行模型预测，即系统对输入的动物图片进行动物识别。  2. **项目目标（即预期交付物）**  设计并实现一个基于卷积神经网络的动物识别系统，系统功能包括：  （1）系统能处理实验数据集。系统将输入的数据集压缩包进行解压，提前设置好标签列表，对数据集进行标注，并生成三个数据集标注文件。然后处理自定义数据集，判断数据集的mode（训练，测试，验证），对于训练数据集定义预处理方法增加样本多样性，获取单个样本数据和标签，非RGB格式图像转化为RGB格式，将标签转换为numpy形式。系统处理完数据集后创建数据集实例作为卷积神经网络的输入。  （2）系统能搭建网络模型进行模型训练。选择合适的卷积神经网络，系统根据其结构搭建模型组网，然后实例化网络模型用于训练。系统会先给模型设置超参（一次训练选择的样本数，训练轮次，学习率）、损失函数（交叉熵）、优化器（Adam,Momentum），然后进行资源配置给系统分配gpu或者cpu。训练进行数据记录并验证数据，最后进行恢复训练得到最终模型。  （3）系统能进行训练模型测试与预测。系统将测试集的图片路径写入txt文本文件，将图片进行压缩、归一化，将创建好的文本中图片路径写入一个列表，然后加载模型将测试集作为输入，将准确率、错误率、精确率、召回率作为输出进行模型测试。通过测试结果来反映模型优劣，模型，即系统将动物图片作为最终模型的输入，输出动物类别来进行动物识别。  3. **拟解决的关键问题**  （1）如何设计数据集的处理方式是系统能处理实验数据集中的关键问题：  根据初步统计，目前数据集中已有百余种动物种类，数据量庞大，若不剔除无效数据并对数据格式进行优化，将大大影响模型训练进度，同时还应对图像数据集进行必要的旋转、翻转操作，以提高数据多样性，同时更加贴切现实应用[14]。  （2）如何选择合适的神经网络模型是系统能搭建网络模型进行模型训练中的关键问题：  解决图像分类识别问题可用的神经网络模型有很多，如何构造神经网络，设计合适的函数评估神经网络，是如何选择合适的神经网络模型，以期实现高效率、高准确率进行动物分类识别的关键。  （3）如何优化选用的神经网络模型，提高模型预测的准确率是系统能进行训练模型测试与预测中的关键问题：  为了实现达到较高的系统识别准确率的项目目标,优化神经网络模型是除选择合适的神经网络模型的另一个关键点，有效的优化能极大地提高模型预测准确率[9]。 | | | |
| **三、拟采取的解决方案及可行性分析**  1. **理论依据（可选）**  本系统是实际应用，本小组讲学习深度学习理论，paddlepaddle框架，采用VGG16以及RepVGG两种卷积神经网络组网，实现动物识别系统。  首先，自定义数据集，读取数据并创建数据集实例；  其次，对VGG16和RepVGG进行组网；  然后，设置好超参后开始训练，并尝试进行参数调优；  最后，对训练好的模型进行测试，测试完成后便可以进行部署。  2. **拟采用的技术路线**  系统数据集采用aistudio上现成的动物图片数据集，对此数据集预处理，继承paddlepaddle框架的Dataset创建自定义数据集并将加载数据集实例化，接着使用DataLoader将数据加载进paddlepaddle框架中，网络模型尝试使用VGG16与RepVGG两种，并在训练过程中对参数进行更改调优，比较两种网络在各种参数下的性能，选择性能最佳的参数与网络模型进行部署。  针对关键问题，下面给出拟采用技术方案  (1) 如何设计数据集的处理方式是系统能处理实验数据集中的关键问题，拟采用的技术路线描述如下：  VGG16以及RepVGG网络的输入为224\*224的RGB图像[1]，所以将所有图像格式化为RGB格式，将原图随机裁切为224\*224大小，为了增加训练集样本多样性，训练集中的数据将有50%的概率水平翻转。最后将图像归一化，能够提供更快的收敛和更稳固的训练。采用继承paddle.io.Dataset创建自定义数据集，将创建的数据集进行实例化，接着使用了paddle.io.DataLoader将数据加载进paddlepaddle框架中的 [2]。  (2) 如何选择合适的神经网络模型是系统能搭建网络模型进行模型训练中的关键问题，拟采用的技术路线描述如下：  VGG16具有巨大的参数数目，具有很高的拟合能力。VGG网络结构简洁，都是由小卷积核、小池化核、ReLU组合而成。VGG优点：小卷积核，小池化核，层数更深，卷积核堆叠的感受野，全连接转卷积；缺点：训练时间长，存储容量大 [1]。  RepVGG将训练推理网络结构进行独立设计，在训练时使用高精度的多分支网络学习权值，在推理时使用低延迟的单分支网络，然后通过结构重参数化将多分支网络的权值转移到单分支网络。RepVGG性能达到了SOTA，思路简单新颖。RepVGG优点：简单即快速、内存使用经济、灵活，训练时的多分支结构；缺点：模型较大，结构没有VGG简单 [3]。  (3) 如何优化选用的神经网络模型，提高模型预测的准确率是系统能进行训练模型测试与预测中的关键问题，拟采用的技术路线描述如下：  首先确定学习率，学习率是一个正标量，用于确定随机梯度下降的步长 [4]。把其他参数确定之后，尝试逐步逼近最佳的学习率。  其次是对两种损失函数的比对，损失函数是衡量模型的效果评估 [4]。第一个损失函数为交叉熵损失：在多分类的情况下公式为[5]      另外一种是负对数似然损失，由于对数似然是对概率分布求对数，概率P(x)的值为[0, 1]区间，取对数后为(-∞, 0]区间。再在前面加个符号，变成[0, ∞)区间[8]。    在优化器的对比选择上，分别采用Adam优化器与Momentum优化器，Adma吸收了Adagrad（自适应学习率的梯度下降算法）和动量梯度下降算法的优点，既能适应稀疏梯度（即自然语言和计算机视觉问题），又能缓解梯度震荡的问题[9]。而Momentum优化器Momentum梯度下降算法在与原有梯度下降算法的基础上，引入了动量的概念，使网络参数更新时的方向会受到前一次梯度方向的影响，换句话说，每次梯度更新都会带有前几次梯度方向的惯性，使梯度的变化更加平滑；Momentum梯度下降算法能够在一定程度上减小权重优化过程中的震荡问题[10]。  3. **可行性分析**  1.技术可行性分析  （1）从技术本身角度出发，基于卷积神经网络的动物识别系统，经过多年的发展，该技术已经成熟稳定，并且已经有了几个成功案例。如采用Darknet-53网络 [6]实现的畜牧业动物图像识别，以及采用MobileNet网络、MobileNet V2网络、MobileNet-Beta网络 [7]实现的动物识别算法，都是已经落地的成功的案例。  （2）从人员储备角度，本团队成员都有过该类项目经验，皆能够短时间内上手，完成项目的设计与实现。  2.组织可行性分析  （1）项目组织架构  需求：实现动物识别系统，具有动物识别的功能，能够对提供的数据集进行动物识别，保证功能齐全，保证效率，保证准确率。  人员分工如下：  韩青哲： 算法优化人员  陈泽锋：算法测试人员  贺思超：算法实现人员  陈增耀：算法设计人员  蒋涵：算法测试人员  (2)项目阶段性任务  第一阶段：完成初始Vision Transformer模型的搭建及性能测试；  时间安排：2022.1.10至2022.2.10  第二阶段：开展模型裁剪、推理加速方法的研究，阅读文献，尝试多种加速方案并且进行分析比较；  时间安排：2022.2.11至2022.3.20  第三阶段：完善最终的推理自适应加速模型；  时间安排：2022.3.21至2022.3.31  第四阶段：将模型部署到车载视觉装备，实现演示系统；  时间安排：2022.4.1至2022.4.15  第五阶段：撰写毕业论文  时间安排：2022.4.11至2022.5.1  3.社会可行性分析  （1）法律可行性：有许多已经落地的成功项目，也说明该项目在法律允许的范围内。  （2）使用可行性：可投入手机中使用，识别未见识过的动物。  4.市场可行性分析  （1）未来非接触式识别将得以大量应用  生物识别是根据人类生理特征(人脸、指纹、虹膜等)和行为特征(姿态、动作、情感等)实现身份认证的技术。常用的生物识别技术主要包括人脸识别、指纹识别、虹膜识别、行为识别以及步态识别。  随着生物识别、人工智能等技术的发展，生物识别市场规模稳步增长，2019年全球生物特征识别技术市场规模约为200亿美元。指纹识别因综合性优势市占率最高，占比达58%。2020年新冠肺炎的爆发，促使非接触生物识别市场的需求不断增加，未来，非接触式识别将得以大量应用。  （2）全球生物识别市场规模持续增长  随着生物识别、人工智能等技术的发展，生物识别市场规模稳步增长，根据信通院数据显示，2019年全球生物特征识别技术市场规模约为200亿美元。 | | | |
| 主要参考文献：  [1] Simonyan K , Zisserman A . Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. Computer Science, 2014.  [2] Yanjun Ma,Dianhai Yu,Tian Wu,Haifeng Wang. PaddlePaddle: An Open-Source Deep Learning Platform from Industrial Practice[J]. Frontiers of Data and Computing, 2019, 1(1): 105-115.  [3] Ding X , Zhang X , Ma N , et al. RepVGG: Making VGG-style ConvNets Great Again[C]// 2021.  [4] Yu T , Zhu H . Hyper-Parameter Optimization: A Review of Algorithms and Applications[J]. 2020.  [5]王峰. 基于深度学习的人脸认证方法研究[D]. 电子科技大学, 2018.  [6]拉毛杰，安见才让. 基于卷积神经网络的畜牧业动物图像识别研究[D]. 青海民族大学, 2020.  [7]袁东芝. 基于卷积神经网络的动物识别算法研究[D]. 华南理工大学, 2018.  [8] Zhu D , Yao H , Jiang B , et al. Negative Log Likelihood Ratio Loss for Deep Neural Network Classification[J]. 2018.  [9] Kingma D , Ba J . Adam: A Method for Stochastic Optimization[J]. Computer Science, 20149  [10] Ruder S . An overview of gradient descent optimization algorithms[J]. 2016.  [11] Osuna E , Freund R , Girosi F . Training svm: An application to face detection. 1997.  [12] 李红丽, 许春香, 马耀锋. 基于多核学习SVM的图像分类识别算法[J]. 现代电子技术, 2018, 41(6):4.  [13]张晓瑞. 基于卷积神经网络的多标签图像分类识别算法研究[J]. 通化师范学院学报, 2022, 43(2):75-82.  [14]丁洁云. 基于深度学习的图像分类处理[D]. 江西师范大学. | | | |
| 毕业设计（论文）进度安排： | | | |
| 序号 | 毕业设计（论文）各阶段内容 | 时间安排 | 备注 |
| 1 | 完成初始Vision Transformer模型的搭建及性能测试； | 2022.1.10至2022.2.10 |  |
| 2 | 开展模型裁剪、推理加速方法的研究，阅读文献，尝试多种加速方案并且进行分析比较； | 2022.2.11至2022.3.20 |  |
| 3 | 完善最终的推理自适应加速模型； | 2022.3.21至2022.3.31 |  |
| 4 | 将模型部署到车载视觉装备，实现演示系统； | 2022.4.1至2022.4.15 |  |
| 5 | 撰写毕业论文 | 2022.4.11至2022.5.1 |  |
| 指导教师意见： | | | |

指导教师（审核签名）： 审核日期： 年 月