Udacity毕业项目:猫狗大战开题报告

领域背景

本次我所选择的猫狗大战项目属于图像识别领域,是基于深度学习的项目,深度学习是机器学习的分支,是一种使用包含复杂结构或由多重非线性变换构成的多个处理层对数据进行高层抽象的算法,深度学习框架,尤其是基于人工神经网络的框架可以追溯到1980年福岛邦彦提出的新认知机[1]

在深度神经网络中,常见的两类问题是过拟合和过长的运算时间,对于解决此类问题,我们可以选择更加快速的GPU,增加改良神经网络的结构,使用更加适合的误差函数和激活函数等可以让我们在一定程度上降低上述问题的影响

最后就个人角度而言,我希望将深度学习方面的知识应用于现阶段的工作和学习中去尝试解决生活中面临的实际问题,因而选择了猫狗大战项目

问题陈述

在这个项目中,我们问题是通过计算机去判断所给的彩色图片中的内容是一只猫还是一只狗,并将判断结果输出,并且准确率达到kaggle Public Leaderboard 的前 10%，属于监督学习，二分类问题

数据集和输入

在该项目中,数据集来自kaggle的[Dogs vs. Cats Redux: Kernels Edition](https://www.kaggle.com/c/dogs-vs-cats-redux-kernels-edition/data),该数据集为彩色的猫狗图片，其中猫狗各12500张图片，猫狗图片比例为1:1在数据集中绝大多数的图片清晰，光线充足，多数为日常拍摄，有少量的特写，其中的图片尺寸大小不一，但图片中猫狗图像居中，需要后期对图片进行归一化处理，数据集中存在少量异常数据，在训练前需要对数据集进行清洗

我将应用该数据集去对我的深度神经网络模型进行训练和测试,使得训练后的模型能够正确的分类在测试集中的猫狗图片,并达到项目要求

解决方案声明

在这个项目中我将使用迁移学习[2]的思想,在选择迁移的模型后,重新构造模型末尾的全连接层,并调整相关参数，使得迁移后的学习模型适用于当前的实际问题

基准模型

在这个项目中，效果要求达到kaggle Public Leaderboard 前 10%，所以在将本次基准定位为logloss要低于0.06127

评估标准

在这个项目中根据kaggle的要求，使用logloss作为模型的评估标准logloss的评估公式为

在logloss中

n是测试集中的图像总数

是图像为狗的概率

如果为1则图片为狗，如果为0则图片为猫

Log（）是基数为e的对数

LogLoss 的取值范围0-1之间，其中数值越趋于0说明模型效果越好，在该项目中，要求LogLoss的值小于0.06127

项目设计

在该项目中，我将项目的通过流程分为：导入数据集，数据预处理，模型架构，测试模型，结果输出，几个部分

在导入数据集的部分中，我会将kaggle中的数据集导入项目中，并将其中用于训练的数据按照0.6:0.2:0.2的比例分割为训练集，验证集，以便选择出效果最好，泛化能力最佳的模型，训练集用于对模型进行训练，学习样本数据集，通过匹配一些参数来建立一个分类器。建立一种分类方式，验证集用于在训练过程中测试模型效果，对学习出来的模型，调整分类器的参数，验证集用于在最后验证模型是否符合项目要求，判断模型的分辨能力

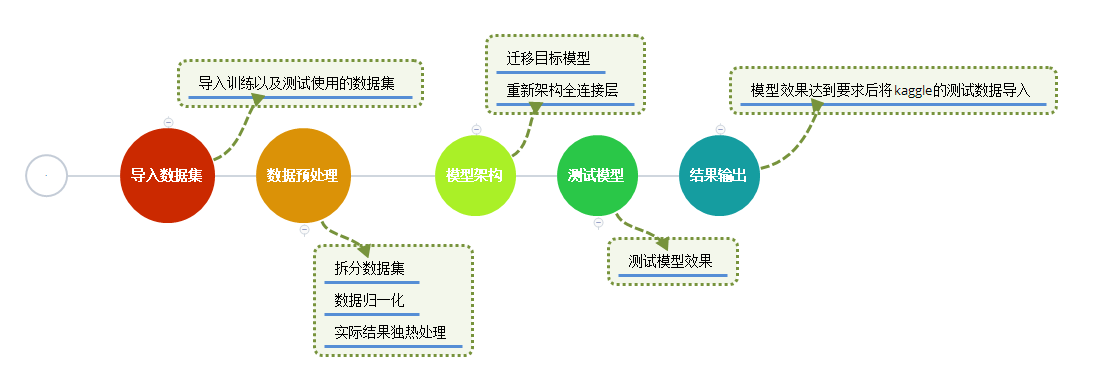
在数据集预处理的部分中我会对数据集中的图像数据进行处理，对图像进行归一化处理：具体步骤为调整图像大小，并将图像调整为具有四个维度的张量，将原始图像转换成相应的唯一标准形式，并对数据集的实际结果进行独热编码分类，图像为猫则置为零，为狗置为一，供后面的训练和测试使用

在模型架构的过程中，我将选择针对这个项目适用的模型，进行迁移学习，具体流程为导入目标迁移学习模型，消除模型末尾的全连接层，并自行建立一个池化层以及全连接层，使用softmax激活函数，使得猫狗两种判断结果各包含一个节点，然后将迁移模型最后一层的卷积层的输出直接输入自己建立的模型中，使得迁移后的模型适用于当前猫狗大战的项目

在选择所迁移的模型的时候，我们要尽量寻找分类数据大致相同，并且表现效果较好的模型进行迁移，当模型效果达不到我们所要求的标准时，可以尝试调整echo数，对训练集进行数据增强，数据集进行镜像翻转等，或者扩充我们的训练集，如果调整后依然无法达到要求，可以尝试更换迁移的模型

在模型的测试阶段，我会使用之前的的验证集对使用logloss对模型的实际效果进行测试，得到当前的模型效果，拥有测试模型是否达到了要求的kaggle Public Leaderboard 前 10%也就是logloss低于0.06127

如果在测试阶段logloss低于0.06127，达到了我们的目标结果，我们会使用kaggle的测试图像通过我们当前建立的模型，得到最后我们的结果



引用：

1. K. Fukushima., "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biol. Cybern.*, 36, 193–202, 1980
2. Jason Yosinski,1 Jeff Clune,2 Yoshua Bengio,3 and Hod Lipson4: How transferable are features in deep neural networks?(2014)