

猫狗大战（开题报告）

项目概览

猫狗大战（Dogs vs Cats）是源自 kaggle 上 2013 年的一个竞赛项目。这个项目主要是为了使用机器学习算法进行图片分类，尤其是深度学习。项目中提供了一个数据集，包括 25000 个已标定的数据和 12500 个未标定数据。通过这 25000 个已标定数据来构建相应的预测算法模型，并使用这个算法模型对 12500 个未标定的数据进行预测，并将预测的结果提交到 kaggle 上，以得到一个综合评分，以评价预测算法模型的优良水平。

近些年，随着深度学习的突破性进展，整个机器学习领域也发展到了新的高度，尤其是卷积神经网络（CNN）在图片识别方面，其识别的准确性已经超过人类的识别水平。卷积神经网络可追溯到上个世纪 80 年代，Fukushima 首次提出 neocognitron 模型，可以看作是卷积神经网络的第一个实现。经过若干年的发展，直到 2006 年深度学习的理论被提出后，卷积神经网络的表征能力得到了关注，并随着数值计算设备的更新开始快速发展。自 2012 年的 AlexNet 开始，卷积神经网络多次成为 ImageNet 大规模视觉识别竞赛（ImageNet Large Scale Visual Recognition Challenge, ILSVRC）的优胜算法，包括 2013 年的 ZFNet，2014 年的 VGGNet 和 GoogleNet，2015 年的 ResNet 等。

问题说明

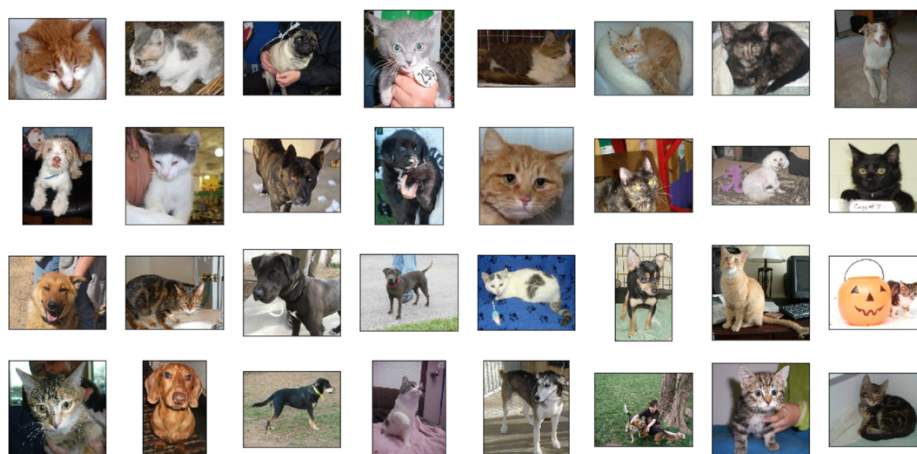
猫狗大战提供的图片源自真实拍摄，分辨率差异较大。图片中猫和狗的颜色丰富，类型多样，姿态迥异，同时还有复杂的背景，极大的增加了识别的难度。

本项目将使用基于深度学习的卷积神经网络（CNN）来构建预测算法模型，通过 25000 个已标定数据（train data）来训练神经网络，然后使用训练好的神经网络模型来预测未标注的数据（test data），来完成预测准确度排名进入到 kaggle 上猫狗大战 Public Leaderboard 排名前 10% 的项目目标。

数据研究

从 kaggle 的 dogs_vs_cats 的数据集中可以发现，数据集的 train 和 test 是进行了分类的，但是 train 数据集中的验证集是没有单独分类的，需要在训练模型的时候自行生成。另外，train 数据集的数据标注也是包含在文件名中的，需要进行特征提取。

为了更好的对数据有直观的了解，将内存中的图片可视化出来。如下：



通过上图可以看出，训练所用的数据，尽管有各种复杂的背景，尽管猫和狗的姿势和形态也比较多样性，但整体来看图片都比较清晰，算是比较理想的数据集。

算法与方法

为了完成项目的分类需求，最常用的就是使用卷积神经网络（CNN）来进行分类，较为著名的 CNN 网络，有 AlexNet、VGGNet、ResNet、Xception 和 Inception 等。这里我们使用 Resnet50 作为我们的神经网络，基本上可以得到不错的评分结果。鉴于项目的时间要求和算力方面考虑，这里通过迁移学习来达成项目的目标。

本项目中使用的 dogs_vs_cats 与 ImageNet 整体上来看，相似度还是比较高的，且 dogs_vs_cats 只提供了 25000 个训练数据（含验证数据）和 12500 个测试数据，样本偏小，非常适合迁移学习。