

|  |  |
| --- | --- |
| 报告名称： | 基于卷积神经网络算法的图像识别实践 |
| 学生学院： | 计算机科学与技术学院/人工智能学院 |
| 教学班级： |  |
| 学生学号： |  |
| 学生姓名： |  |
| 课程名称： |  |
| 完成日期： | 2022年4月30日 |
| 报告成绩： |  |

基于卷积神经网络算法的图像识别实践

摘要：

在21世纪，图像识别技术已经极大的改变了我们的生活方式。目前，全世界许多知名科技公司已经开始了在领域的布局。图像识别可以通过分类并提取重要特征而排除多余的信息来识别图像中的物体。图像质量的好坏直接影响图像识别算法的准确度，因此在图像识前，需要进行图像预处理。图像预处理的步骤为数字化、几何变换和图像增强。

**关键词：图像识别; 卷积神经网络; 图像预处理**

1. 数据来源

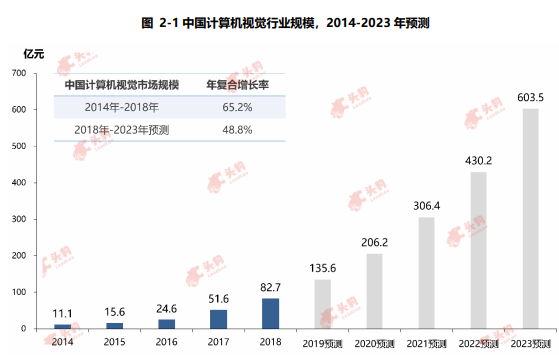
1.1收集数据

图像识别，顾名思义，是指为了实现识别各种不同模式的目标和对象，利用计算机对图像进行处理、分析和理解的一种技术，是深度学习算法的一种实践应用。目前图像识别技术一般分为人脸识别与商品识别，人脸识别主要运用在安全检查、身份核验与移动支付中；商品识别主要运用在商品流通过程中，尤其是无人货架、智能零售柜等无人零售领域。

1.2 行业背景

目前，全世界许多知名科技公司已经开始了在领域的布局。Facebook签下的人工智能专家Yann LeCun提出的LeNet为代表的卷积神经网络，在应用到各种不同的图像识别任务时都取得了很好的效果，被认为是通用图像识别系统的代表之一；Google 通过模拟神经网络通过对上百万份 YouTube 视频的学习自行掌握了猫的关键特征，这是机器在没有人帮助下自己读懂了猫的概念。值得一提的是，负责这个项目的Andrew NG专家目前就职于百度研究院，其一个重要的研究方向就是图像识别。这也能看出国内科技公司对图像识别技术的重视程度。

从行业规模来看，随着图像识别准确率显著突破,图像识别行业将进入快速发展阶段,预测到2023年中国计算机视觉行业规模达603.5亿元。



1. 问题分析

2.1问题描述

当我们人类看到一个东西时，大脑会迅速把看到的东西和记忆中相同或相类的东西进行匹配，从而识别这个物体。那么如何使得计算机执行这个操作从而识别它呢？

2.2 数据分析目标

图像的传统识别流程分为四个步骤：图像采集→图像预处理→特征提取→图像识别。图像识别软件国外代表的有康耐视等，国内代表的有图智能、海深科技等。另外在地理学中指将遥感图像进行分类的技术。

计算机通过分类并提取重要特征而排除多余的信息来识别图像中的物体。

1. 图像预处理

3.1 图像预处理的目的

图像识别中，图像质量的好坏直接影响识别算法的准确度，因此在图像识前，需要进行预处理。图像预处理的主要目的是消除图像中无关的信息，恢复有用的真实信息，从而增强有关信息的可靠性、最大限度地简化数据。

3.2 图像预处理的步骤

一般的预处理流程为：数字化->几何变换->图像增强。

3.3 图像预处理–数字化

一幅原始图像的灰度值是空间变量的连续函数。通过对图像灰度采样并加以量化，可以得到计算机能够处理的数字图像。为了使数字图像能重建原来的图像,使得M（点阵行数）、N（点阵列数）和b（灰度等级）尽可能的大,这样重建图像的质量就越好。并且当取样周期等于或小于原始图像中最小细节周期的一半时，重建图像的频谱等于原始图像的频谱，因此重建图像与原始图像可以完全相同。

3.4 图像预处理–几何变换

几何变换用于修正图像采集系统的系统误差和仪器位置的随机误差。比如拍摄器姿态和高度变化引起的误差，这种误差难以用模型表示出来，所以一般通过把被观测的图和已知正确几何位置的图相比较，用图中一定数量的控制点解双变量多项式函数组而达到变换的目的。

3.5 图像预处理–图像增强

通过对图像中的信息选择性的加强或抑制，来达到改善图像的视觉效果，或者将图像转变为更适合于机器处理的形式，以便于数据抽取或识别。图像增强技术有多种方法，对数变换、密度分层和直方图均衡等都可用于改变图像灰调和突出细节。不过我们实际应用时往往要用不同的方法，反复进行试验才能达到满意的效果。

1. 模型求解

首先，我们需要导入数据集，我们可以使用MNIST数据集、ImageNet数据集或者PASCAL数据集。这些数据集都是公开的，支持全球的开发者进行下载训练模型数据集。我们将数据集下载解压缩后会发现这些文件并不是标准的图像格式。也就是有一个训练图片集，一个训练标签集，一个测试图片集，一个测试标签集；这些文件是压缩文件，解压出来，我们看到的是二进制文件。

接下来要进行数据的处理，我们可以参考Iris数据的处理方式进行，即首先将label进行one-hot处理，之后使用TensorFlow 2.0自带的data API进行打包，这样很方便组合成train与label的配对数据集。

一个简单的卷积神经网络模型是由卷积层、池化层、dropout层以及作为分类的全连接层构成，同时每一层之间使用relu激活函数做分割。全连接层的作用是对卷积层所提取的特征做最终分类。这里我们首先使用flat函数，将提取计算后的特征值平整化，之后的2个全连接层起到特征提取和分类的作用。最终做出分类。

|  |
| --- |
| #输入的初始化  input\_XS = tf.keras.Input([28, 28, 1])  #卷积层，通过设定卷积核数据、卷积核大小、padding方式和激活函数初始化了整个卷积类。  conv = tf.keras.layers.Conv2D(32, 3, padding=" SAME", activation=tf.nn.relu)(input\_xs)  conv = tf.keras.layers.BatchNormalization()(conv)  conv = tf.keras.layers.Conv2D(64, 3, padding="SAME", activation=tf.nn.relu)(conv)  conv = tf.keras.layers.MaxPool2D(strides=[1, 1])(conv)  conv = tf.keras.layers.Conv2D(128, 3, padding="SAME", activation=tf.nn.relu)(conv)  flat = tf.keras.layers.Flatten()(conv)  #全连接层，作用：分类  dense = tf.keras.layers.Dense(512, activation=tf.nn.relu)(flat)  logits = tf.keras.layers.Dense(10, activation=tf.nn.softmax)(dense)  model = tf.keras.Model(inputs=input\_XS, outputs=logits)  print(model.summary()) |

1. 总结

图像识别是一种机器学习方法，其设计类似于人类大脑的功能。通过这种方法，计算机可以识别图像中的视觉元素。通过依靠大型数据库计算机可以理解图像，并制定相关的标签和类别。

图像识别技术的实现上有两个难题，一是图像需要处理的数据量太大，导致成本很高，效率很低。另一个是图像在数字化的过程中很难保留原有的特征，导致图像处理的准确率不高。卷积神经网络可以较好的解决以上两个问题。CNN将复杂问题简化，把大量参数降维成少量参数，再做处理。并且使用类似视觉的方式保留了图像的特征，当图像做翻转，旋转或者变换位置时，它也能有效的识别出来是类似的图像。

卷积神经网络是传统神经网络的改进。主要包含数据输入层、卷积计算层、ReLU激励层、池化层及全连接层。数据输入层主要是对原始图像数据进行预处理。在卷积层中，一个神经元只与邻层部分的神经元连接，包含若干个特征平面，每个特征平面由多个神经元按矩阵形式排列组成，同一特征平面的神经元共享权值（卷积核）。 池化用于压缩网络参数和数据大小，降低过拟合。全连接层，即两层之间所有神经元权重连接，通常全连接层在卷积神经网络尾部，跟传统的神经网络神经元的连接方式相同。

图像识别技术应用相当广泛。随着科技的进步，人类对图像识别技术的认识也会更加深刻。在21世纪，我们无法想象离开了图像识别技术以后我们的生活会变成什么样。图像识别技术是人类现在以及未来生活必不可少的一项技术。

参考文献

1. 王晓华. TensorFlow 2.0卷积神经网络实战,清华大学,2020.
2. 言有三.深度学习之图像识别：核心技术与案例实战,机械工业出版社,2019.
3. 卷积神经网络（CNN）在图像识别的最新进展

<https://zhuanlan.zhihu.com/p/38681805>

