



(12)发明专利申请

(10)申请公布号 CN 107808132 A

(43)申请公布日 2018.03.16

(21)申请号 201710992656.X

(22)申请日 2017.10.23

(71)申请人 重庆邮电大学

地址 400065 重庆市南岸区南山街道崇文路2号

(72)发明人 丰江帆 付阿敏 孙文正 夏英

(74)专利代理机构 重庆市恒信知识产权代理有限公司 50102

代理人 刘小红

(51)Int.Cl.

G06K 9/00(2006.01)

G06K 9/62(2006.01)

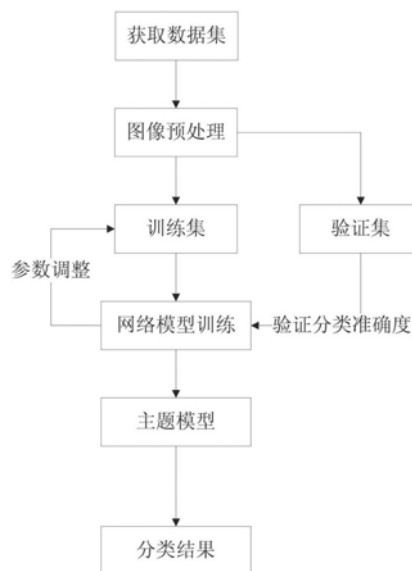
权利要求书2页 说明书4页 附图1页

(54)发明名称

一种融合主题模型的场景图像分类方法

(57)摘要

本发明请求保护一种融合主题模型的场景图像分类方法,涉及深度学习及图像分类领域。该方法包括:对数据集进行预处理,扩充已获取数据集的数量,得到符合深度学习模型处理的图像数据格式;构建符合场景图像分类的卷积神经网络模型,使用卷积神经网络对处理后的图像数据集进行预训练;使用训练集,对构建的卷积神经网络进行端到端的迭代训练,调整网络中的参数,使用验证集对训练完成的模型进行验证,对提取到的具有判别力的场景图像特征进行建模,提取特征和图像之间存在的隐藏的主题变量,得到k维向量代表的图像主题分布,k代表主题数量;每幅图像可以看做是由多个主题所组成的一个概率分布图,利用分类器实现场景图像的分类。



1. 一种融合主题模型的场景图像分类方法,其特征在于,包括以下步骤:

1)、对获取的图像数据集进行预处理操作,预处理一是扩充了数据集,二是使得处理后的图像格式符合深度学习框架的格式,将处理过的数据集的70%作为训练集,剩下的30%作为验证集;

2)、构建符合场景图像分类的卷积神经网络模型,使用卷积神经网络对处理后的训练集进行预训练,即使用卷积神经网络的卷积层和池化层提取输入图像的位置、上下文特征,利用这些学习到的图像的位置、上下文特征作为图像场景分类的基础和依据;

3)、使用训练集,对构建的卷积神经网络进行端到端的迭代训练,在训练过程中调整网络中的参数,使用验证集对训练完成的模型进行验证,网络模型提取的图像特征判断场景图像的类别;

4)、通过预训练提取到的具有判别力的场景图像特征进行建模,提取特征和图像之间存在的隐藏的主题变量,得到k维向量代表的图像主题分布;每幅图像代表多个主题所组成的一个概率分布,利用分类器实现场景图像的分类。

2. 根据权利要求1所述的融合主题模型的场景图像分类方法,其特征在于,所述步骤1)对图像数据集进行预处理操作的步骤包括裁剪、翻转,直方图均衡化,调整图像的亮度,得到更大容量的数据集,这样能够使每一个类别的场景图像得到充分的训练,调高网络的泛化能力,另一方面使得处理后的图像格式符合深度学习框架的格式,使得图像能够被网络模型读取和训练。

3. 根据权利要求1或2所述的融合主题模型的场景图像分类方法,其特征在于,所述符合场景图像分类的卷积神经网络模型,包括数据输入层、卷积计算层、池化层、全连接层、输出层,数据输入层是把预处理的数据输入到网络模型;卷积计算层采用局部连接,每个神经元看做一个过滤器,通过对窗口做滑动操作,窗口里的值与过滤器里的值对应相乘,作为下一层的输入;池化层的位置处于连续的卷积层中间,用来压缩数据和参数的数量,减少过拟合;全连接层位于卷积神经网络模型的尾部,最大可能的利用经过滑动窗口和池化后保留的少量的信息还原原来的输入信息;输出层是把保留的少量的信息输入到Softmax归一化函数,其作用是将网络的输出结果归一化成概率分布概率最高的一个即为分类结果。

4. 根据权利要求3所述的融合主题模型的场景图像分类方法,其特征在于,

所述步骤3)使用训练集,对构建的卷积神经网络进行端到端的迭代训练具体包括:的卷积神经网络前向过程的卷积是卷积核与输入图像的对应位置求积再求和的过程,求得的值作为下一次操作的输入,每次卷积核在输入的图像上移动一个位置,从上到下,从左到右覆盖一遍后得到的输出矩阵就是下一个操作的输入的特征图;在错误信号反向传播过程中,错误信号由分类器向前面的隐藏层传播;在训练过程中模型自动学习网络参数,更新权重;使用验证集对训练完成的模型进行验证,使网络模型提取的图像特征能够更准确的判断场景图像的类别。

5. 根据权利要求4所述的融合主题模型的场景图像分类方法,其特征在于,所述网络模型提取的图像特征中,其中有些特征是相似或相同的,把这些具有相似或相同的特征抽象为主题的概念;仿照统计文本分类的概率主题模型,将提取到的图像低层局部特征量化为视觉单词,通过统计整幅图像中每个视觉单词出现的频次得到图像的直方图表示,将每幅图像视为一个词频向量,每幅图像代表由多个主题组成的一个概率分布,每个主题又代表

多个视觉单词所组成的一个概率分布;将已保存的特征向量输入到主题模型中,根据实验分析不同的主题数量对分类准确率的影响,确定主题数量的个数。

6.根据权利要求1所述的融合主题模型的场景图像分类方法,其特征在于,所述卷积神经网络模型使用开源的深度学习框架TensorFlow,在该框架上搭建卷积神经网络模型,TensorFlow能够在各类机器上运行,能够同时多个CPU、GPU或者两者混合的机器上运行。

一种融合主题模型的场景图像分类方法

技术领域

[0001] 本发明属于深度学习及图像分类识别技术领域,具体是一种融合主题模型的场景图像分类方法。

背景技术

[0002] 场景图像分类,即给定一组包含多个目标类别(如山脉、河流、公路等)的场景图像,根据各个目标类别的分布关系对图像的全局语义进行分析和理解。场景图像分类不仅对整幅图像的类别有总体的认识,还对图像中各个物体、以及区域之间的上下文信息进行了分析,使得对图像的内容有了更深层次的认识,推动了机器视觉中诸如目标识别、图像检索等领域的发展,具有广泛的应用范围。

[0003] 随着智能拍照设备及计算机硬件的快速发展,使用深度学习中的深度卷积神经网络替代传统图像特征提取方法,避开了传统的模式识别算法需要进行人工特征提取的繁琐过程,具有更强大的特征提取和特征表达能力。使用深度学习算法训练的卷积神经网络模型自不仅在理论上取得了极大的突破,在实际应用中也取得了丰富的成果。

[0004] 深度学习模型属于多层神经网络模型,卷积神经网络是第一个被成功训练的深度学习模型。深层神经网络模型在训练阶段使用的是没有经过任何人为影响的原始自然图像,能够非常有效的学习具有旋转、扭曲等不变性特征的图像表示,这在很大程度上降低了手工提取特征对分类准确率的影响。卷积神经网络是由多个层次组成的可训练的多层架构,采用端到端的处理方式,把图像预处理和特征提取的过程视作一个黑盒子,通过反卷积操作或者分析分类结果的精度优化网络参数,不断提高训练的卷积神经网络模型的分类精度。

发明内容

[0005] 本发明旨在解决针对已有技术中图像的底层视觉信息与人们对图像理解的高层语义信息的不一致性而导致的底层和高层间的语义鸿沟问题。提出了一种得分类识别的效果得到进一步提高的融合主题模型的场景图像分类方法。本发明的技术方案如下:

[0006] 一种融合主题模型的场景图像分类方法,其包括以下步骤:

[0007] 1)、对获取的图像数据集进行预处理操作,这有两个好处,首先是起到扩充数据集的作用,增加了训练样本数量,其次通过预训练使得处理后的图像格式符合深度学习框架的格式,将处理过的数据集的70%作为训练集,剩下的30%作为验证集;

[0008] 2)、构建符合场景图像分类的卷积神经网络模型,使用卷积神经网络对处理后的训练集进行预训练,即通过卷积神经网络的卷积层和池化层提取输入图像的位置、上下文特征,利用这些学习到的图像的位置、上下文特征作为图像场景分类的基础和依据;

[0009] 3)、使用训练集,对构建的卷积神经网络进行端到端的迭代训练,在训练过程中调整网络中的参数,使用验证集对训练完成的模型进行验证,网络模型提取的图像特征判断场景图像的类别;

[0010] 4)、对通过预训练提取到的具有判别力的场景图像特征进行建模,提取特征和图像之间存在的隐藏的主题变量,得到k维向量代表的图像主题分布;每幅图像代表多个主题所组成的一个概率分布,利用分类器实现场景图像的分类。

[0011] 进一步的,所述步骤1)对图像数据集进行预处理操作的步骤包括裁剪、翻转,直方图均衡化,调整图像的亮度,得到更大容量的数据集,这样能够使每一个类别的场景图像得到充分的训练,调高网络的泛化能力,另一方面使得处理后的图像格式符合深度学习框架的格式,使得图像能够被网络模型读取和训练。

[0012] 进一步的,所述符合场景图像分类的卷积神经网络模型,包括数据输入层、卷积计算层、池化层、全连接层、输出层,数据输入层是把预处理的数据输入到网络模型;卷积计算层采用局部连接,每个神经元看做一个过滤器,通过对窗口做滑动操作,窗口里的值与过滤器里的值对应相乘,作为下一层的输入;池化层的位置处于连续的卷积层中间,用来压缩数据和参数的数量,减少过拟合;全连接层位于卷积神经网络模型的尾部,最大可能的利用经过滑动窗口和池化后保留的少量的信息还原原来的输入信息;输出层是把保留的少量的信息输入到Softmax归一化函数,其作用是将网络的输出结果归一化成概率分布,概率最高的一个即为分类结果。

[0013] 进一步的,所述步骤3)使用训练集,对构建的卷积神经网络进行端到端的迭代训练具体包括:的卷积神经网络前向过程的卷积是卷积核与输入图像的对应位置求积再求和的过程,求得的值作为下一次操作的输入,每次卷积核在输入的图像上移动一个位置,从上到下,从左到右覆盖一遍后得到的输出矩阵就是下一个操作的输入的特征图;在错误信号反向传播过程中,错误信号由分类器向前面的隐藏层传播;在训练过程中模型自动学习网络参数,更新权重;使用验证集对训练完成的模型进行验证,使网络模型提取的图像特征能够更准确的判断场景图像类别。

[0014] 进一步的,所述网络模型提取的图像特征中,其中有些特征是相似或相同的,把这些具有相似或相同的特征抽象为主题的概念;仿照统计文本分类的概率主题模型,将提取到的图像低层局部特征量化为视觉单词,通过统计整幅图像中每个视觉单词出现的频次得到图像的直方图表示,将每幅图像视为一个词频向量,每幅图像代表由多个主题组成的一个概率分布,每个主题又代表多个视觉单词所组成的一个概率分布;将已保存的特征向量输入到主题模型中,根据实验分析不同的主题数量对分类准确率的影响,确定主题数量的个数。

[0015] 进一步的,所述卷积神经网络模型使用开源的深度学习框架TensorFlow,在该框架上搭建卷积神经网络模型,TensorFlow能够在各类机器上运行,能够同时多个CPU、GPU或者两者混合的机器上运行。

[0016] 本发明的优点及有益效果如下:

[0017] 本发明考虑到现有图片的易获取性和计算机硬件的快速发展以及深度学习技术取得的成就,选择使用卷积神经网络模型提取图像特征比传统的人工提取特征、多特征融合等方法取得的特征更具有客观性,更具有判别力。通过调整网络参数使得网络能够学习大量的能够判别场景图像的特征,将提取到的特征保存下来;这些特征中许多特征是相似或相近的,为了避免有用的特征在分类的过程中被无意的丢弃,我们把这些具有相似或相近意义的特征抽取为主题的概念,这也弥补了低层特征直接到高层特征的差距问题,减少

了输入到分类器的特征数量,提高分类效率和分类准确度。

附图说明

[0018] 图1是本发明提供优选实施例基于深度学习实现场景图像分类的方法的流程图;

[0019] 图2为图像预处理的步骤。

具体实施方式

[0020] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、详细地描述。所描述的实施例仅仅是本发明的一部分实施例。

[0021] 本发明解决上述技术问题的技术方案是:

[0022] 图1示出了本发明的基于深度学习实现场景图像分类的方法的流程图,具体步骤如下:

[0023] (1)对图像数据集进行进行预处理,包括利用裁剪、翻转,直方图均衡化,调整图像的亮度,扩充数据集,并调整图像格式为卷积神经网络模型的可读格式,选取数据集的70%作为训练集,剩下的30%作为验证集;

[0024] (2)使用开源的深度学习框架TensorFlow,在该框架上搭建卷积神经网络模型。TensorFlow是google的深度学习框架,自带的可视化工具TensorBoard是一个非常好用的网络结构可视化工具,对于分析训练网络非常有用;TensorFlow 能够在各类机器上运行,能够同时在多个CPU、GPU或者两者混合的机器上运行,灵活性好;

[0025] 构建符合场景图像分类的卷积神经网络模型,包括数据输入层、卷积计算层、池化层、全连接层、输出层。数据输入层是把预处理的数据输入到网络模型;卷积层采用局部连接,每个神经元看做一个过滤器,通过对窗口做滑动操作,窗口里的值与过滤器里的值对应相乘,作为下一层的输入;池化层的位置处于连续的卷积层中间,用来压缩数据和参数的数量,减少过拟合;全连接层位于卷积神经网络模型的尾部,最大可能的利用经过滑动窗口和池化后保留的少量的信息还原原来的输入信息;

[0026] (4)使用训练集,对构建的卷积神经网络进行端到端的迭代训练,在训练过程中调整网络中的参数,使用验证集对训练完成的模型进行验证,使网络模型提取的图像特征能够更准确的判断场景图像的类别;在训练过程中,由于网络模型的多层性使用反向传播算法对隐含层的节点进行学习;

[0027] (5)提取到的特征中,其中有些特征是相似或相同的,我们把这些具有相似或相同的特征抽象为主题的概念;仿照统计文本分类的概率主题模型,将提取到的图像低层局部特征量化为视觉单词,通过统计整幅图像中每个视觉单词出现的频次得到图像的直方图表示,将每幅图像视为一个词频向量,每幅图像代表由多个主题组成的一个概率分布,每个主题又代表多个视觉单词所组成的一个概率分布;通过这种方法可以弥补低层特征提取直接到高层特征映射的差距问题,也避免了一些具有判别力的特征在分类的过程中被随机的遗弃,提高场景图像的分类准确率;将已保存的特征向量输入到主题模型中,根据实验分析不同的主题数量对分类准确率的影响,确定主题数量的个数;

[0028] (6)每幅图像代表多个主题所组成的一个概率分布,利用分类器实现场景图像的分类。将已经提取的主题特征向量做为分类器的输入对象,如果需要分为K 个类,可以设置

K个输出点,根据输出结果判断场景图像分类的准确度。

[0029] 图2示出了图像进行预处理的步骤,具体如下:

[0030] (1) 对图像进行缩放、翻转:图像缩放变换是按照一定的比例放大或者缩小图像,翻转变化是使图像内容沿着水平或垂直方向发生的镜像变化。为了训练出一个具有泛化能力较强的网络模型,我们需要对数据集进行扩充,对图像进行缩放和翻转变化是较为常用的两种图像增强方法,使一幅能够以不同图像内容的形式在网络中进行训练,提高网络的泛化能力和分类准确率。

[0031] (2) 对图像进行直方图均衡:直方图均衡通过提高图像中最亮和最暗部分之间的对比度,找出图像内容中亮度差距较小内容的细微差别;另外创建一个更高对比度的图像加入到训练集中,减少图像对亮度的依赖程度。

[0032] (3) 调整图像的饱和度:调整图像的饱和度是对图像色彩度的调整,增加饱和度就会增加图像色彩的纯度,降低饱和度就会减少图像色彩的纯度,当图像的饱和度降低为0时,图像就会变为灰色。

[0033] 以上这些实施例应理解为仅用于说明本发明而并不用于限制本发明的保护范围。在阅读了本发明的记载的内容之后,技术人员可以对本发明作各种改动或修改,这些等效变化和修饰同样落入本发明权利要求所限定的范围。

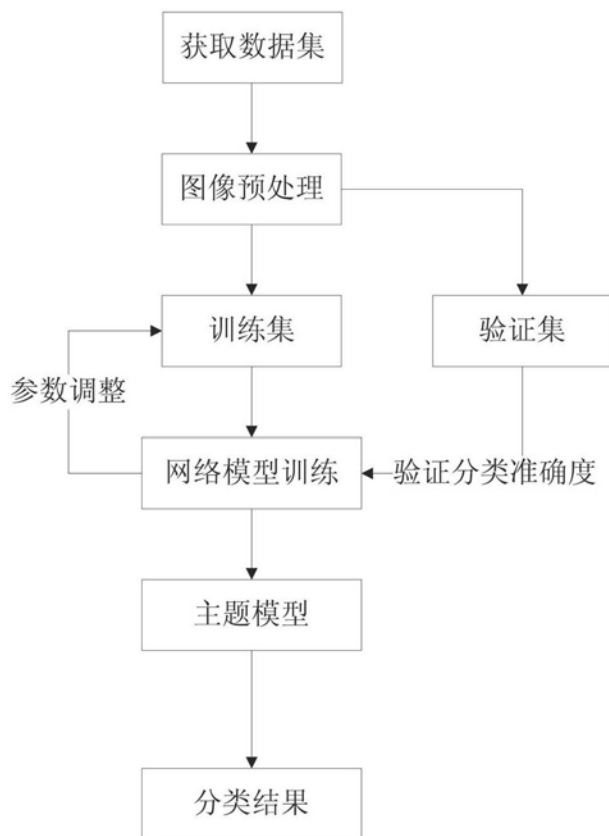


图1



图2