



(12)发明专利申请

(10)申请公布号 CN 107862344 A

(43)申请公布日 2018.03.30

(21)申请号 201711245804.8

(22)申请日 2017.12.01

(71)申请人 中南大学

地址 410083 湖南省长沙市岳麓区麓山南路932号

(72)发明人 郭克华 唐达济

(74)专利代理机构 长沙正奇专利事务有限责任公司 43113

代理人 马强 王娟

(51)Int.Cl.

G06K 9/62(2006.01)

权利要求书2页 说明书7页 附图4页

(54)发明名称

一种图像分类方法

(57)摘要

本发明公开了一种图像分类方法,选择M张具有类代表的原始图片;对所述原始图片进行预处理,将M张原始图片扩充为M组训练图片,每组包含N张图片;对M组图片进行tensorflow训练,得到训练模型;对新的图片,根据所述训练模型对所述新的图片进行图像类别测试。本发明以图片数据为应用对象,对一张原始图片进行图像变换处理生成多张图片,不需要用户提供成千上万的训练图片数据,只需一张原始图片就能自动化的生成1152张图片,该方法不仅能节省大量原始图片数据的采集成本,同时也能保证深度学习的训练模型的需要,并且能对多种图片分类取得一个很好的识别效果。



1. 一种图像分类方法,其特征在于,包括以下步骤:
 - 1) 选择M张具有类代表的原始图片;
 - 2) 对所述原始图片进行预处理,将M张原始图片扩充为M组训练图片,每组包含N张图片;
 - 3) 对M组图片进行tensorflow训练,得到训练模型;
 - 4) 对新的图片,根据所述训练模型对所述新的图片进行图像类别测试。
2. 根据权利要求1所述的图像分类方法,其特征在于,将M张原始图片扩充为M组训练图片的具体实现过程包括:
 - 1) 对所述原始图片进行翻转,将一张原始图片扩充为两张翻转图片;
 - 2) 对所述翻转图片的R、G、B进行扫描,指定一个度degree,给出随机噪声noise的区间 $[-degree, degree]$,分别对翻转图片的R、G、B进行噪声添加,即 $R+noise$ 、 $G+noise$ 、 $B+noise$;若 $R+noise$ 、 $G+noise$ 、 $B+noise$ 小于0,将该小于0的元素赋值为0; $R+noise$ 、 $G+noise$ 、 $B+noise$ 大于255,将该大于255的元素赋值为255,从而得到加噪后的图片;
 - 3) 获取所述加载后的图片的宽度Hortical、长度Vertical和颜色Color,分别从宽度和长度按照一个比例o进行压缩变换,生成压缩图片;其中比例o为正整数;
 - 4) 设置亮度调节度,利用下式分别处理压缩图片的像素R、G、B三分量: $pixel = p[i] + degree$, pixel为像素颜色集,该像素颜色集包含{R、G、B}三个像素颜色分量;i的值分别为0,1,2,p[0]代表B,p[1]代表G,p[2]代表R;若pixel小于0,pixel赋值为0,pixel大于255,pixel赋值为255,分别取亮度调节度为正数和负数的两种情况对压缩图片进行变换,通过pixel的改变得到不同的亮度调节图片;
 - 5) 对所述亮度调节图片进行灰度调节,得到灰度调节图片;
 - 6) 遍历所述灰度调节图片的所有像素点,对每个像素点进行如下曝光调节,得到曝光图片;
若 $p[0] < 128$, $p[0] = (\text{byte}) (p[0] \wedge 0xFF)$;
若 $p[1] < 128$, $p[1] = (\text{byte}) (p[1] \wedge 0xFF)$;
若 $p[2] < 128$, $p[2] = (\text{byte}) (p[2] \wedge 0xFF)$;
 - 7) 将所述曝光图片转化为RGB位图格式,获取其宽度width和高度height,设置一个色调调节值hue,利用所述色调调节值hue对每个像素点的RGB进行色调调节,得到颜色调节图片;
 - 8) 以 90° 为步长,对所述颜色调节图片进行三次旋转,得到旋转图片。
3. 根据权利要求2所述的图像分类方法,其特征在于,度degree为整数且 $0 \leq degree \leq 255$ 。
4. 根据权利要求2所述的图像分类方法,其特征在于,从宽度和长度按照一个比例o对压缩图片进行压缩变换后得到的像素值为 $(Hortical * o + m++, Vertical * o + n++, Color)$;其中 $m++, n++$ 为0到o之间的整数。
5. 根据权利要求2所述的图像分类方法,其特征在于,灰度调节的具体实现过程为:亮度调节图片像素点的color值, $color1 = \text{oldBitmap.GetPixel}(x, y)$, $color2 = \text{oldBitmap.GetPixel}(x+1, y+1)$;分别从红、绿、蓝三个通道进行调节,选择0到255之间的灰度调节度degree, $R = \text{Math.Abs}(color1.R - color2.R + degree)$; $G = \text{Math.Abs}(color1.G -$

color2.G+degree); $B=\text{Math.Abs}(\text{color1.B}-\text{color2.B}+\text{degree})$, Math.Abs 函数代表返回函数内的绝对值, 通过该函数保证 R、G、B 的值在 0 到 255 之间, 若有小于 0 的元素, 则将该元素值赋值为 0, 若有大于 255 的元素, 则将该元素值赋值为 255; color1 代表图片调节前像素点 (x, y) 的颜色, color 代表像素点 (x, y) 附近像素点 $(x+1, y+1)$ 的颜色; $\text{oldBitmap.GetPixel}$ 函数用于获取压缩图片中的颜色值, 是由 R、G、B 组成的集合, $\text{color1.R}-\text{color2.R}$ 代表对像素点水平和竖直方向进行逐一差分。

6. 根据权利要求 2 所述的图像分类方法, 其特征在于, 所述色调调节值 hue 为介于 -180 到 180 之间的浮点数。

7. 根据权利要求 6 所述的图像分类方法, 其特征在于, 利用所述色调调节值 hue 对每个像素点的 RGB 进行色调调节, 利用所述色调调节值 hue 对每个像素点的 RGB 进行色调调节的公式为: $\text{hsl}=\text{HSL.FromRgb}(p[2], p[1], p[0])$, $\text{hsl.Hue}+=\text{hue}$; 其中 hsl 是工业界的颜色标准, 其三个缩写字母分别代表图像的色相、饱和度和明度, 函数 $\text{HSL.FromRgb}(R, G, B)$ 用来实现图像红绿蓝值到色调值的转化过程, hsl.Hue 代表当前图像的色调值, hue 代表所调节色调的变化值。

一种图像分类方法

技术领域

[0001] 本发明涉及图像处理领域,特别是一种图像分类方法。

背景技术

[0002] 近年来,信息技术的迅速发展,加大了机器学习、深度学习的对数据的采集、处理和存储依赖。无论是政府还是研究机构都耗费巨资投入大数据技术研究,目标是通过大数据技术和深度学习实现感知、认知和预测支持的结合。深度学习往往需要成千上万的样本数据进行训练,为了保证深度学习的效果,传统的深度学习依赖于大量数据作为样本进行训练,而大量数据的采集成本非常高。而如何获取经济有效的训练数据样本成为深度学习的一个重要问题。因此,为了保证深度学习的效果,并且不需要大量的训练数据采集,这里提出一种基于数据扩充的深度学习方法。该方法通过对一张原始数据图片进行翻转、加噪、压缩、亮度、灰度、曝光、变色、旋转8个处理生成1152张图片。然后将扩充后的图片数据导入Tensorflow中进行训练得到模型,并用其他图片进行识别。实验结果表明,该扩充后的数据图片通过训练能够保证深度学习的有效性。

[0003] 目前,对深度学习样本的研究主要还是基于对深度学习模型的改进、样本特征的提取和卷积网络、深度信度网络的改进。在[1]一文中,作者提出了一类新的小样本深度学习模型:UGES反向传导模型。其基本思路是:在保留深层结构的同时,压缩需要学习参数的数量。在[2]一文中,作者在人脸识别中提出了一种层对层的卷积神经网络训练方法解决小样本的过拟合问题,并达到了91%的准确率。在[3]一文中,作者针对小超声图像数据集的肿瘤分类问题,提出了一个特征学习和分类的方法(DPN-MKL)。实验结果表明该方法在小样本数据的肿瘤分类中超出了传统用的DL方法。

[0004] 针对如何减少深度学习的数据训练样本成本并且保证深度学习效果的研究,要么是对训练样本进行特征标注提取,要么对学习过程进行方法改进。这些研究都带来了其他方面成本(比如标注成本)的增加,并且样本数量并不能够有非常显著的减少,目前这一块所做的研究工作并不多,也没有成熟的方法和工具支持图片数据的扩充过程。在深度学习中,训练数据越好,训练效果越好,需要用来学习的训练样本量很大,但是有时无法提供成千上万的数据图片用来学习,训练数据不足或高成本成为制约深度学习的一大瓶颈。

发明内容

[0005] 本发明所要解决的技术问题是,针对现有技术不足,提供一种图像分类方法,节省大量原始图片数据的采集成本,同时保证深度学习的训练模型的需要,并且能对多种图片分类取得很好的识别效果。

[0006] 为解决上述技术问题,本发明所采用的技术方案是:一种图像分类方法,包括以下步骤:

[0007] 1) 选择M张具有类代表的原始图片;

[0008] 2) 对所述原始图片进行预处理,将M张原始图片扩充为M组训练图片,每组包含N张

图片;

[0009] 3) 对M组图片进行tensorflow训练,得到训练模型;

[0010] 4) 对新的图片,根据所述训练模型对所述新的图片进行图像类别测试。

[0011] 将M张原始图片扩充为M组训练图片的具体实现过程包括:

[0012] 1) 对所述原始图片进行翻转,将一张原始图片扩充为两张翻转图片;

[0013] 2) 对所述翻转图片的R、G、B进行扫描,指定一个度degree,给出随机噪声noise的区间 $[-degree, degree]$,分别对翻转图片的R、G、B进行噪声添加,即 $R+noise$ 、 $G+noise$ 、 $B+noise$;若 $R+noise$ 、 $G+noise$ 、 $B+noise$ 小于0,将该小于0的元素赋值为0; $R+noise$ 、 $G+noise$ 、 $B+noise$ 大于255,将该大于255的元素赋值为255,从而得到加噪后的图片;

[0014] 3) 获取所述加载后的图片的宽度Hortical、长度Vertical和颜色Color,分别从宽度和长度按照一个比例o进行压缩变换,生成压缩图片;其中比例o为正整数;

[0015] 4) 设置亮度调节度,对 $RGB=256^2*红+256^1*绿+256^0*蓝=65536*红+256*绿+蓝$ (红,绿,蓝是0~255之间的数,数值越小,亮度越低,数值越大,亮度越高。)想对每个像素进行选择,就得先选中它,可以用亮度改变它。本亮度调节采用最常用的对图像像素点的R、G、B三个分量同时进行增加(减少)某个值,达到调整亮度的目的。利用下式分别处理压缩图片的像素R、G、B三分量: $pixel=p[i]+degree$,pixel为像素颜色集,里面包含{R、G、B}三个像素颜色分量;i的值分别为0,1,2,p[0]代表B,p[1]代表G,p[2]代表R;若pixel小于0,pixel赋值为0,pixel大于255,pixel赋值为255,分别取亮度调节度为正数和负数的两种情况对压缩图片进行变换,主要是通过pixel的改变得到不同的亮度调节图片;

[0016] 5) 对所述亮度调节图片进行灰度调节,得到灰度调节图片;

[0017] 6) 遍历所述灰度调节图片的所有像素点,对每个像素点进行如下曝光调节,得到曝光图片;

[0018] 若 $p[0]<128$, $p[0]=(byte)(p[0]^0xFF)$;

[0019] 若 $p[1]<128$, $p[1]=(byte)(p[1]^0xFF)$;

[0020] 若 $p[2]<128$, $p[2]=(byte)(p[2]^0xFF)$;

[0021] 7) 将所述曝光图片转化为RGB位图格式,获取其宽度width和高度height,设置一个色调调节值hue,利用所述色调调节值hue对每个像素点的RGB进行色调调节,得到颜色调节图片;

[0022] 8) 以 90° 为步长,对所述颜色调节图片进行三次旋转,得到旋转图片。

[0023] 度degree为整数且 $0 \leq degree \leq 255$ 。

[0024] 从宽度和长度按照一个比例o对压缩图片进行压缩变换后得到的像素值为(Hortical*o+m++,Vertical*o+n++,Color);其中m++,n++为0到o之间的整数。

[0025] 灰度调节的具体实现过程为:亮度调节图片像素点的color值,color1=oldBitmap.GetPixel(x,y),color2=oldBitmap.GetPixel(x+1,y+1);分别从红、绿、蓝三个通道进行调节,选择0到255之间的灰度调节度degree, $R=Math.Abs(color1.R-color2.R+degree)$; $G=Math.Abs(color1.G-color2.G+degree)$; $B=Math.Abs(color1.B-color2.B+degree)$,Math.Abs函数代表返回函数内的绝对值,通过该函数保证R、G、B的值在0到255之间,若有小于0的元素,则将该元素值赋值为0,若有大于255的元素,则将该元素值赋值为255;color1代表图片调节前像素点(x,y)的颜色,color代表像素点(x,y)附近像素点(x+1,

y+1)的颜色;oldBitmap.GetPixel函数用于获取压缩图片中的颜色值,是由R、G、B组成的集合,color1.R-color2.R代表对像素点水平和垂直方向进行逐一差分。

[0026] 所述色调调节值hue为介于-180到180之间的浮点数。

[0027] 利用所述色调调节值hue对每个像素点的RGB进行色调调节,通过HSL.FromRgb(p[2],p[1],p[0])将图像的RGB值转化为HSL的值。然后通过hsl.Hue+=hue对图像的色调进行一定的调节,hue为色调调节的大小。利用所述色调调节值hue对每个像素点的RGB进行色调调节的公式为:hsl=HSL.FromRgb(p[2],p[1],p[0]),hsl.Hue+=hue;其中hsl是工业界的颜色标准,其三个缩写字母分别代表图像的色相(H)、饱和度(S)和明度(L),函数HSL.FromRgb(R、G、B)用来实现图像红绿蓝值到色调值的转化过程,hsl.Hue代表当前图像的色调值,hue代表所调节色调的变化值。

[0028] 与现有技术相比,本发明所具有的有益效果为:本发明以图片数据为应用对象,对一张原始图片进行图像变换处理生成多张图片,不需要用户提供成千上万的训练图片数据,只需一张原始图片就能自动化的生成1152张图片,该方法不仅能节省大量原始图片数据的采集成本,同时也能保证深度学习的训练模型的需要,并且能对多种图片分类取得一个很好的识别效果。

附图说明

[0029] 图1为本发明方法流程图;

[0030] 图2为本发明图片垂直翻转示意图,苹果把柄朝右,经过翻转变成把柄朝左;

[0031] 图3为本发明图片加噪示意图,右侧苹果比左侧苹果经过加噪有更多影响视觉效果噪声点;

[0032] 图4为本发明图片的压缩变化图,右侧苹果比左侧苹果经过压缩像素之间的清晰度降低;

[0033] 图5为本发明图片的亮度调节变化,中间苹果比最左侧苹果亮度更高,最右侧的苹果比最左侧的苹果亮度更低;

[0034] 图6为本发明图片的灰度调节变化,右侧的苹果比左侧的苹果相比灰度值发生了改变;

[0035] 图7为本发明图片的曝光调节变化,右侧的苹果比左侧的苹果相比曝光值发生了改变;

[0036] 图8为本发明图片的颜色调节变化,中间的苹果比左侧的苹果相比要白很多,最右侧的苹果则是介于黑色和白色苹果之间;

[0037] 图9为本发明图片的旋转,最左侧是原始正放的苹果图片,从左往右,苹果依次旋转90度得到3张新图片;

[0038] 图10为本发明数据扩充图;

[0039] 图11为本发明图像识别的5类水果(苹果、榴莲、葡萄、香蕉、西瓜)原始图片;

[0040] 图12为本发明水果识别测试效果图;

[0041] 图13为本发明做图像识别的5类动物(猫、狗、狮子、豹子、老虎)原始图片;

[0042] 图14为本发明动物识别测试效果图;

[0043] 图15为本发明用来做图像识别的5类汽车品牌(奔驰、宝马、奥迪、本田、马自达)原

始图片；

[0044] 图16为本发明汽车识别测试效果图。

具体实施方式

[0045] 图像的分类是深度学习的一个热门研究,一个训练好的图像分类模型离不开好的分类训练样本数据。基于tensorflow深度学习样本需要5组训练样本,且每组训练样本的数量至少成百上千张才能保证训练效果。选择5张具有类代表的原始图片,这里的类代表指的是具有该类别事物明显特征的图片,可以是水果集、动物集、花儿集、汽车集等任意常用tensorflow图像识别的数据集。其中这5张图片属于同一个大类别集合,且每张原始图片具有较好的辨识度,以便于后续的图像处理和数据扩充。

[0046] 基于深度学习训练样本的数据扩充,本发明从一张原始训练图片入手,通过图像处理最终将一张训练图片扩充成为1152张图片用于tensorflow训练。具体实现方法步骤如下:

[0047] 步骤1:选择5张具有类代表的原始图片;

[0048] 步骤2:根据原始图片分别从图像的翻转、加噪、压缩、亮度调节、灰度调节、曝光调节、变色和旋转8个方面对图像进行变换。

[0049] 步骤3:将5张原始图片扩充为5组训练图片,每组图片1152张;

[0050] 步骤4:将5组图片进行tensorflow训练,得到训练模型;

[0051] 步骤5:用新的测试图片依据训练好的模型进行图像识别测试。

[0052] 其具体流程如附图中的图1。实验表明该技术方案对图片扩充后的样本进行训练具有明显的图像识别效果。

[0053] 图像样本数据的扩充是该方案的一个重要技术点,主要目的是能从客观事物的图像中提取有效信息,通过图像处理不断生成新的图片,并能让转化后的图片够被机器理解学习。在基本保持原始图片的有效信息的同时,对原始图片进行变换,作为新的训练样本被tensorflow学习。根据原始图片的像素分布、亮度、颜色、灰度、曝光、噪声、压缩等方面对图片进行变换生成新的训练图片。下面分别从以下8个方面对图像进行变换处理。

[0054] 步骤1:翻转

[0055] 对于图像而言图像的垂直翻转后不改变图像的特征提取但生成了新的图片,比如书本本来向右,垂直翻转为向左,机器还可以辨识出书的基本特征。通过对原始图片像素位置的对称变换FlipX,而保留了图片的原始特征。这样就将1张图片生成另一张新的图片,对图片数据进行了2倍的扩充。图片翻转效果见图2。翻转生成的图片用于第二步加噪处理图片的输入。

[0056] 步骤2:加噪

[0057] 图像中各种妨碍人们对其信息接受的因素即可称为图像噪声,噪声在理论上是不可预测,只能用概率统计方法来认识的随机误差,因此对于大量的图片数据而言,图像噪声是多维随机的,可以通过人为对图片添加噪声的方式生成新的图片,从而对翻转后的图片进行扩充。在32位的RGB格式图像中,噪声的最小值为0,最大值为255。首先对原始图像的R、G、B进行扫描保存,然后指定一个度degree(degree为整数且degree大于等于0,小于等于255)给出随机噪声noise的区间 $[-degree, degree]$,然后分别对图像的R、G、B进行噪声添

加,即 $R+noise$ 、 $G+noise$ 、 $B+noise$ 。若 R 、 G 、 B 中经过加噪后值小于0,将该小于0的元素赋值为0;若 R 、 G 、 B 中经过加噪后值大于255,将该大于255的元素赋值为255。通过加噪的方式可以得到新的样本数据,对数据进行2倍扩充。图片加噪效果见图3。加噪生成后的图片用于第三步压缩处理图片的输入。

[0058] 步骤3:压缩

[0059] 图像压缩是指以较少的比特有损或无损的表示原来的像素矩阵技术,去除冗余数据,由原二维像素阵列变换为一个统计上无关联的数据集合。通过对图像进行压缩可以基本保留原始图片的特征并且生成新的训练图片。首先对原始图像的宽度 $Hortical$ 、长度 $Vertical$ 和颜色 $Color$ 进行获取,然后分别从宽度和长度按照一个比例 o (o 为压缩比例,为正整数)进行压缩变换 $SetPixel(Hortical*o+m++,Vertical*o+n++,Color)$,其中 $m++$, $n++$ 为0到 o 之间的整数。通过图像压缩的方式生成新的图片,从而对数据进行2倍扩充。图片压缩效果见图4。压缩生成后的图片用于第四步亮度调节处理图片的输入。

[0060] 步骤4:亮度调节

[0061] 图像的像素具有相应的亮度,亮度的值介于0和255之间,靠近0的像素亮度较低,靠近255的像素较高,其余部分属于中间调。不同的图片具有不同的像素亮度,这里通过对图像进行亮度调节可以基本保持原始图片的特征并生成新的样本数据。首先扫描原始图片的宽度和高度,将图片转化成32位RGB的位图流格式,设置一个亮度调节度 $degree$, $degree$ 的值为-255到255之间的任意整数,若 $degree$ 为负数,表示将像素亮度降低, $degree$ 为正数,表示将像素亮度增高。分别处理像素 R 、 G 、 B 三分量 p , $pixel=p[i]+degree$, i 的值分别为0,1,2, $p[0]$ 代表 B , $p[1]$ 代表 G , $p[2]$ 代表 R 。若 $pixel$ 小于0, $pixel$ 赋值为0, $pixel$ 大于255, $pixel$ 赋值为255。为了模拟亮度调节的降低和增高,分别取 $degree$ 为正数和负数两种情况对原始图片进行变换。这样可以由一张原始图片得到一张亮度降低和一张亮度增高的图片,新得到的图片具有原始图片的基本特征。通过亮度调节对数据进行了3倍扩充。图片亮度调节效果见图5。亮度调节生成后的图片用于第五步灰度处理图片的输入。

[0062] 步骤5:灰度调节

[0063] 图像的灰度是指将图像的白色与黑色之间按对数关系分为256阶。一张图片由红色、绿色、蓝色三个通道组成,用不同的灰度色阶来表示“红、绿、蓝”在图像中的比重,从最暗黑色到最亮的白色灰度。不同图片具有不同的灰度值,灰度的调节可以基本保持原图的基本特征并生成新的图片。首先获取像素点的 $color$ 值, $color1=oldBitmap.GetPixel(x,y)$, $color2=oldBitmap.GetPixel(x+1,y+1)$ 。然后灰度的调节分别从红、绿、蓝三个通道进行调节,选择0到255之间的灰度调节度 $degree$, $r=Math.Abs(color1.R-color2.R+degree)$; $g=Math.Abs(color1.G-color2.G+degree)$; $b=Math.Abs(color1.B-color2.B+degree)$ 。最后保证 r 、 g 、 b 的值在0到255之间若有小于0的元素其值赋值为0,若有大于255的元素,其值赋值为255。这样,一张原始图片经过灰度调节可以生成一张新的图片,对图片数据进行2倍扩充。图片灰度调节效果见图6。灰度调节生成后的图片用于第六步曝光处理图片的输入。

[0064] 步骤6:曝光调节

[0065] 对同一事物的拍摄,不同的通光时间(快门速度)和通光面积(光圈大小)可以得到不同的图片,但都保留了事物的基本特征。这里,可以通过对原始图片的曝光调节生成新的

图片,从而扩充样本数据用于深度学习。首先将图片扫描为32为RGB位图流格式data,并获取图片宽度width和高度height。然后设置一个偏移量。遍历所有像素点,对每个像素点进行如下曝光调节:

[0066] 若 $p[0] < 128$, $p[0] = (\text{byte}) (p[0] \wedge 0xFF)$;

[0067] 若 $p[1] < 128$, $p[1] = (\text{byte}) (p[1] \wedge 0xFF)$;

[0068] 若 $p[2] < 128$, $p[2] = (\text{byte}) (p[2] \wedge 0xFF)$;其中 $p[0]$ 、 $p[1]$ 、 $p[2]$ 分别代表B、G、R。通过上述方式,可以将一张图片进行曝光调节后生成一张新的图片,对数据进行了2倍扩充。图片曝光调节效果见图7。曝光生成后的图片用于第七步颜色处理图片的输入。

[0069] 步骤7:颜色调节

[0070] 图像每个像素点的颜色有多种呈现状态,包括色调、饱和度、亮度等,颜色特征是在图像检索中应用最为广泛的视觉特征之一。客观事物本身可能具有不同的视觉颜色效果,例如一个红色的苹果、黄色的苹果、和青色的苹果,都会被识别为一个苹果。这里我们主要对原始图片的色调进行一些变换生成不同颜色的图片,而保持原始图片的纹理、形状等特征。其中像素的色调介于-180和180之间的浮点数。首先将图片转化为32为RGB的位图格式,获取其宽度width和高度height。然后设置一个色调调节值hue,其中hue介于-180到180之间的浮点数。接下来对每个像素点的RGB进行色调调节, $hsl = \text{HSL.FromRgb}(p[2], p[1], p[0])$, $hsl.Hue += hue$ 。其中hsl是一种工业界的颜色标准, $p[0]$ 、 $p[1]$ 、 $p[2]$ 分别代表B、G、R。在颜色调节时可以对hue赋不同的值来表现事物不同的色调,这里对hue分别赋值两次,因此可以由一张原始图片新得到2张不同颜色的图片数据,对数据进行了3倍扩充。图片颜色调节效果见图8。颜色调节生成后的图片用于第八步旋转处理图片的输入。

[0071] 步骤8:旋转

[0072] 图像旋转是指图像以某一点为中心旋转一定角度,形成一幅新的图像过程。旋转前后的点离中心的位置不变,这种几何变换没有改变图像每个像素的RGB,只是改变了像素的位置,保留了原始图片的基本特征。这里分别以90度一次对图片进行旋转,可以生成90度旋转图片、180度旋转图片和270度旋转图片。即由一张原始图片得到3张新的扩充图片,对数据进行了4倍扩充。图片旋转调节效果见图9。

[0073] 通过上述8个步骤的图像处理,一张原始图片经过翻转扩充2倍,加噪扩充2倍,压缩扩充2倍,亮度扩充3倍,灰度调节扩充2倍,曝光调节扩充2倍,变色扩充3倍,旋转扩充4倍,一共扩充 $2 \times 2 \times 3 \times 2 \times 2 \times 3 \times 4 = 1152$ 倍。即通过这些图像处理手段生成了1152倍具有原始图片基本特征而又与原始图片不完全一样的图片用于深度学习的样本训练。图片处理效果见图10。从数量上满足了深度学习的样本需要。

[0074] 将前面选择的5张具有明显类特征的原始图片按照上述方法进行数据扩充,得到5组1152张图片,将这5组图片进行tensorflow深度学习,得到训练好的pb模型。最后选取多组测试图片集对训练好的模型进行测试。实验结果表明,采用该数据扩充方法在图像分类识别中具有通用性,能有效对测试图片集进行分类,达到和大样本采集数据一样的学习效果,具体实验数据结果见附图11-附图16。

[0075] 本发明根据深度学习需要大量样本数据进行训练,通过一套特殊图像处理方法,在保留原始图片基本特征的基础上,对图片进行了1152倍的扩充,然后基于tensorflow对扩充的样本数据进行学习得到训练模型,最后用一组测试图片集进行测试,发现该模型对

图片分类具有很好的识别效果。该技术方案提出了一种基于数据扩充的深度学习方法,极大降低了深度学习对大量样本数据采集的成本,该图像处理方法已经通过用C#编程实现了图像的自动化生成,通过该程序自动化的将一张图片5分钟内生成1152张保持原始图片基本特征而又可以用于tensorflow训练的数据图片,在保证训练样本数量和质量的同时,节省了相关人员的时间和精力。



图1



图2



图3

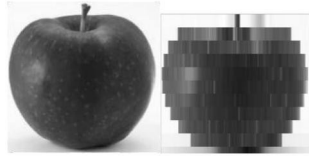


图4



图5



图6



图7



图8

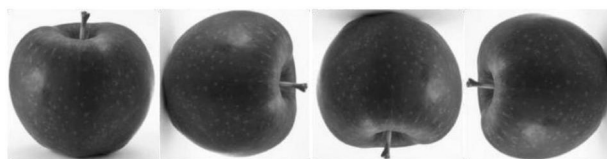


图9

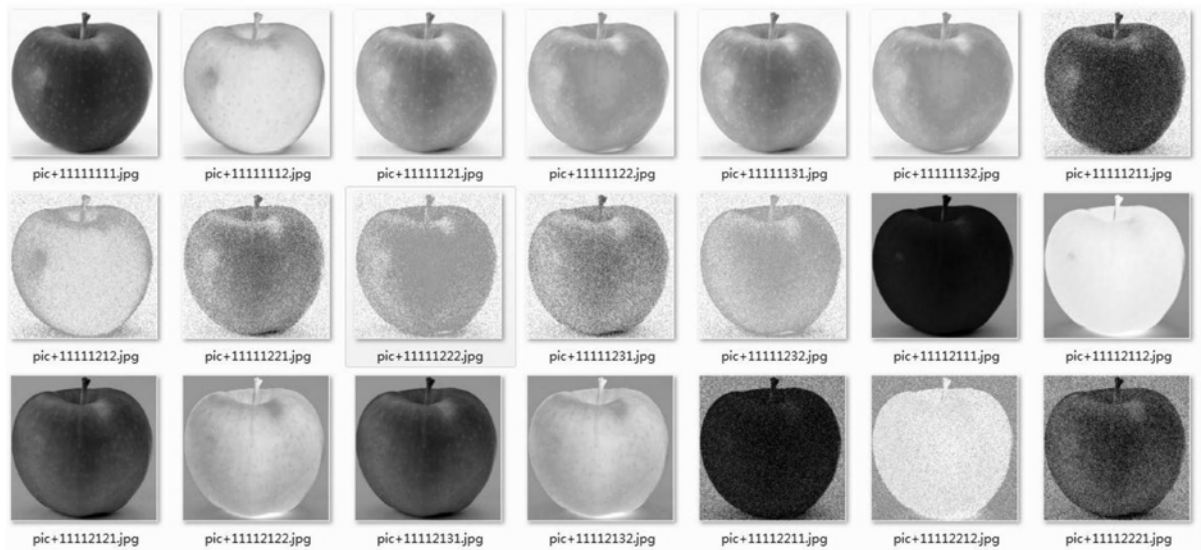


图10



图11

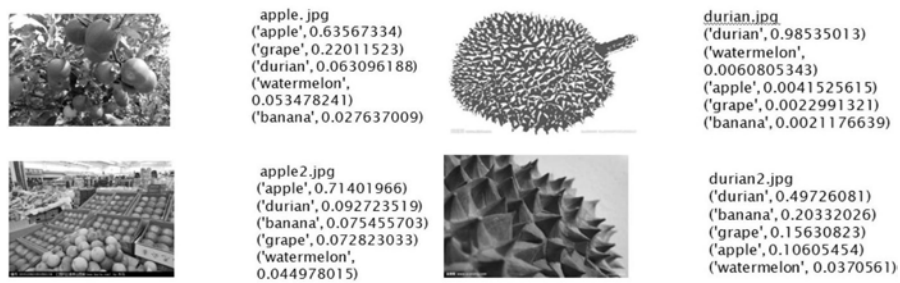


图12



图13



图14



图15



图16