



(12)发明专利申请

(10)申请公布号 CN 109978074 A

(43)申请公布日 2019.07.05

(21)申请号 201910272826.6

(22)申请日 2019.04.04

(71)申请人 山东财经大学

地址 250014 山东省济南市历下区二环东路7366号

(72)发明人 崔超然 余俊 杨文雅

(74)专利代理机构 济南圣达知识产权代理有限公司 37221

代理人 张庆骞

(51) Int. Cl.

G06K 9/62(2006.01)

G06N 3/04(2006.01)

G06N 3/08(2006.01)

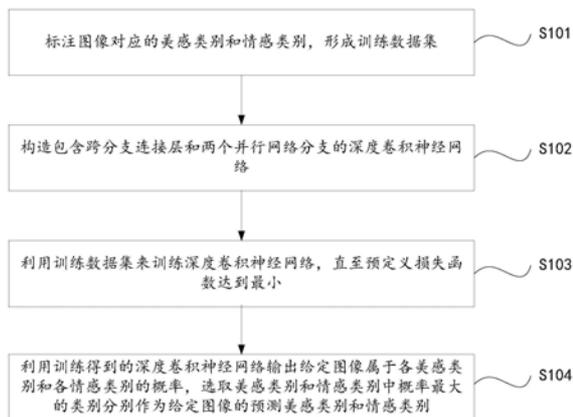
权利要求书2页 说明书8页 附图3页

(54)发明名称

基于深度多任务学习的图像美感和情感联合分类方法及系统

(57)摘要

本公开提供了一种基于深度多任务学习的图像美感和情感联合分类方法及系统。其中,该联合分类方法包括:标注图像对应的美感类别和情感类别,形成训练数据集;构造包含跨分支连接层和两个并行网络分支的深度卷积神经网络;利用训练数据集来训练深度卷积神经网络,直至预定义损失函数达到最小;利用训练得到的深度卷积神经网络输出给定图像属于各美感类别和各情感类别的概率,选取美感类别和情感类别中概率最大的类别分别作为给定图像的预测美感类别和情感类别。



1. 一种基于深度多任务学习的图像美感和情感联合分类方法,其特征在于,包括:
标注图像对应的美感类别和情感类别,形成训练数据集;

构造包含跨分支连接层和两个并行网络分支的深度卷积神经网络;

其中,两个网络分支分别负责对输入图像进行美感分类和情感分类;跨分支连接层用于连接两个网络分支中对应的卷积层组,以关联美感分类和情感分类这两个任务;深度卷积神经网络的输出代表输入图像属于各美感类别和各情感类别的概率;

利用训练数据集来训练深度卷积神经网络,直至预定义损失函数达到最小;

利用训练得到的深度卷积神经网络输出给定图像属于各美感类别和各情感类别的概率,选取美感类别和情感类别中概率最大的类别分别作为给定图像的预测美感类别和情感类别。

2. 如权利要求1所述的基于深度多任务学习的图像美感和情感联合分类方法,其特征在于,在所述深度卷积神经网络中,两个网络分支中的卷积层组数量相同均为 n ;跨分支连接层的数量为 $n-1$;第 i 个跨分支连接层将两个网络分支中第 i 个对应的卷积层组输出的图像特征图作为输入,并将这些输入的图像特征图沿着通道方向进行堆叠,将堆叠后的图像特征分别输入至两个网络分支中第 $i+1$ 个对应的卷积层组; $1 \leq i \leq n-1$; n 为大于或等于2的正整数。

3. 如权利要求2所述的基于深度多任务学习的图像美感和情感联合分类方法,其特征在于,每个卷积层组均包含一个最大池化层和至少两个连续的卷积层。

4. 如权利要求1所述的基于深度多任务学习的图像美感和情感联合分类方法,其特征在于,利用训练数据集来训练深度卷积神经网络的过程,包括:

统一训练数据集的所有图像的大小尺寸;

初始化深度卷积神经网络各层的权重,预定义损失函数;

采用随机梯度下降算法对深度卷积神经网络进行训练,确定能使得损失函数最小的网络权重;且在每次训练迭代时,从图像的随机位置处裁剪出一块固定大小图像块,并以一定概率对图像块进行水平翻转。

5. 一种基于深度多任务学习的图像美感和情感联合分类系统,其特征在于,包括:

训练数据集形成模块,其用于标注图像对应的美感类别和情感类别,形成训练数据集;

深度卷积神经网络构造模块,其用于构造包含跨分支连接层和两个并行网络分支的深度卷积神经网络;

其中,两个网络分支分别负责对输入图像进行美感分类和情感分类;跨分支连接层用于连接两个网络分支中对应的卷积层组,以关联美感分类和情感分类这两个任务;深度卷积神经网络的输出代表输入图像属于各美感类别和各情感类别的概率;

深度卷积神经网络训练模块,其用于利用训练数据集来训练深度卷积神经网络,直至预定义损失函数达到最小;

预测分类模块,其用于利用训练得到的深度卷积神经网络输出给定图像属于各美感类别和各情感类别的概率,选取美感类别和情感类别中概率最大的类别分别作为给定图像的预测美感类别和情感类别。

6. 如权利要求5所述的基于深度多任务学习的图像美感和情感联合分类系统,其特征在于,在所述深度卷积神经网络中,两个网络分支中的卷积层组数量相同均为 n ;跨分支连

接层的数量为 $n-1$ ；第 i 个跨分支连接层将两个网络分支中第 i 个对应的卷积层组输出的图像特征图作为输入，并将这些输入的图像特征图沿着通道方向进行堆叠，将堆叠后的图像特征分别输入至两个网络分支中第 $i+1$ 个对应的卷积层组； $1 \leq i \leq n-1$ ； n 为大于或等于2的正整数。

7. 如权利要求6所述的基于深度多任务学习的图像美感和情感联合分类系统，其特征在于，每个卷积层组均包含一个最大池化层和至少两个连续的卷积层。

8. 如权利要求5所述的基于深度多任务学习的图像美感和情感联合分类系统，其特征在于，所述深度卷积神经网络训练模块，包括：

尺寸统一模块，其用于统一训练数据集的所有图像的大小尺寸；

初始化模块，其用于初始化深度卷积神经网络各层的权重，预定义损失函数；

迭代训练模块，其用于采用随机梯度下降算法对深度卷积神经网络进行训练，确定能使得损失函数最小的网络权重；且在每次训练迭代时，从图像的随机位置处裁剪出一块固定大小图像块，并以一定概率对图像块进行水平翻转。

9. 一种计算机可读存储介质，其上存储有计算机程序，其特征在于，该程序被处理器执行时实现如权利要求1-4中任一项所述的基于深度多任务学习的图像美感和情感联合分类方法中的步骤。

10. 一种计算机设备，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序，其特征在于，所述处理器执行所述程序时实现如权利要求1-4中任一项所述的基于深度多任务学习的图像美感和情感联合分类方法中的步骤。

基于深度多任务学习的图像美感和情感联合分类方法及系统

技术领域

[0001] 本公开属于计算机视觉技术领域,尤其涉及一种基于深度多任务学习的图像美感和情感联合分类方法及系统。

背景技术

[0002] 本部分的陈述仅仅是提供了与本公开相关的背景技术信息,不必然构成在先技术。

[0003] 伴随计算机视觉技术的快速发展,人们不仅希望计算机能够在语义层面对图像的内容进行分析,更期望计算机能够模拟人类视觉及思维系统,产生更高层次的感知能力。作为感知理解研究中的两项代表性任务,图像的美感分类和情感分类分别旨在使计算机可以辨认人类由受到图像视觉刺激而产生的审美和情感反应。目前,图像的美感分类和情感分类技术已经被应用在图像的存储、编辑、检索等方面。例如,针对用户拍摄的关于同一物体或场景的多张候选照片,筛选最具美感的作品保存和展示,合理地降低数据的存储开销;在图像作品的创作和编辑中,分析对比候选方案的美学质量,提升作品的视觉美感;在图像检索系统中,考虑返回图像的情感倾向,为用户提供语义准确且更有感染力的检索结果。

[0004] 由于图像内容的多样性和人类感知的复杂性,自动实现对图像的美感分类和情感分类是非常具有挑战性的任务。近年来,得益于带有美感标记和情感标记的大规模图像数据集的出现,基于机器学习的方法被广泛采用。方法的核心步骤是提取在分类任务上具有良好区分能力的图像视觉特征。早期的方法主要依赖人工设计的特征,需要研究者对问题本身有深入的了解。随着深度学习在计算机视觉领域的兴起,近期的方法主要利用卷积神经网络自动地抽取特征用于图像美感和情感分类,并获得了较好的效果。

[0005] 发明人发现,现有技术通常将图像的美感分类和情感分类当作两个相互独立的任务。但直觉上,人类的美感感受和情感感受并不是孤立出现的;相反,在心理认知层面上,它们应是相互关联和相互影响的。例如,如果一幅图像能够使人们获得审美上的愉悦,那它也很可能会唤起观察者的积极情感。神经科学领域的研究也证明,人类的审美体验是一种伴随着情感状态不断升级的认知过程,反之亦然。

发明内容

[0006] 为了解决上述问题,本公开的第一个方面提供一种基于深度多任务学习的图像美感和情感联合分类方法,其通过统一的深度卷积神经网络框架,使两个任务之间可以有效地共享信息,实现对图像的美感类别和情感类别联合识别以及识别准确性和效率。

[0007] 为了实现上述目的,本公开采用如下技术方案:

[0008] 一种基于深度多任务学习的图像美感和情感联合分类方法,包括:

[0009] 标注图像对应的美感类别和情感类别,形成训练数据集;

[0010] 构造包含跨分支连接层和两个并行网络分支的深度卷积神经网络;

[0011] 其中,两个网络分支分别负责对输入图像进行美感分类和情感分类;跨分支连接

层用于连接两个网络分支中对应的卷积层组,以关联美感分类和情感分类这两个任务;深度卷积神经网络的输出代表输入图像属于各美感类别和各情感类别的概率;

[0012] 利用训练数据集来训练深度卷积神经网络,直至预定义损失函数达到最小;

[0013] 利用训练得到的深度卷积神经网络输出给定图像属于各美感类别和各情感类别的概率,选取美感类别和情感类别中概率最大的类别分别作为给定图像的预测美感类别和情感类别。

[0014] 为了解决上述问题,本公开的第二个方面提供一种基于深度多任务学习的图像美感和情感联合分类系统,其通过统一的深度卷积神经网络框架,使两个任务之间可以有效地共享信息,实现对图像的美感类别和情感类别联合识别以及识别准确性和效率。

[0015] 为了实现上述目的,本公开采用如下技术方案:

[0016] 一种基于深度多任务学习的图像美感和情感联合分类系统,包括:

[0017] 训练数据集形成模块,其用于标注图像对应的美感类别和情感类别,形成训练数据集;

[0018] 深度卷积神经网络构造模块,其用于构造包含跨分支连接层和两个并行网络分支的深度卷积神经网络;

[0019] 其中,两个网络分支分别负责对输入图像进行美感分类和情感分类;跨分支连接层用于连接两个网络分支中对应的卷积层组,以关联美感分类和情感分类这两个任务;深度卷积神经网络的输出代表输入图像属于各美感类别和各情感类别的概率;

[0020] 深度卷积神经网络训练模块,其用于利用训练数据集来训练深度卷积神经网络,直至预定义损失函数达到最小;

[0021] 预测分类模块,其用于利用训练得到的深度卷积神经网络输出给定图像属于各美感类别和各情感类别的概率,选取美感类别和情感类别中概率最大的类别分别作为给定图像的预测美感类别和情感类别。

[0022] 为了解决上述问题,本公开的第三个方面提供一种计算机可读存储介质,其通过统一的深度卷积神经网络框架,使两个任务之间可以有效地共享信息,实现对图像的美感类别和情感类别联合识别以及识别准确性和效率。

[0023] 为了实现上述目的,本公开采用如下技术方案:

[0024] 一种计算机可读存储介质,其上存储有计算机程序,该程序被处理器执行时实现上述所述的基于深度多任务学习的图像美感和情感联合分类方法中的步骤。

[0025] 为了解决上述问题,本公开的第四个方面提供一种计算机设备,其通过统一的深度卷积神经网络框架,使两个任务之间可以有效地共享信息,实现对图像的美感类别和情感类别联合识别以及识别准确性和效率。

[0026] 为了实现上述目的,本公开采用如下技术方案:

[0027] 一种计算机设备,包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序,所述处理器执行所述程序时实现上述所述的基于深度多任务学习的图像美感和情感联合分类方法中的步骤。

[0028] 本公开的有益效果是:

[0029] 本公开将多任务学习的思想运用到图像的美感分类和情感分类上来,充分利用了两个任务之间的关联特性,并设计了一个统一的深度卷积神经网络框架,通过跨分支连接

层使网络分支之间可以以交换图像特征图的方式使两个任务之间可有效地共享信息,并在训练过程中自动地学习不同任务需要哪些信息,实现对图像的美感类别和情感类别联合识别,提高了图像美感类别和情感类别的准确性。

附图说明

[0030] 构成本公开的一部分的说明书附图用来提供对本公开的进一步理解,本公开的示意性实施例及其说明用于解释本公开,并不构成对本公开的不当限定。

[0031] 图1是本公开实施例提供的一种基于深度多任务学习的图像美感和情感联合分类方法流程图。

[0032] 图2是本公开实施例提供的深度卷积神经网络示意图。

[0033] 图3是本公开实施例提供的跨分支连接层示意图。

[0034] 图4是本公开实施例提供的一种基于深度多任务学习的图像美感和情感联合分类系统结构示意图。

具体实施方式

[0035] 下面结合附图与实施例对本公开作进一步说明。

[0036] 应该指出,以下详细说明都是例示性的,旨在对本公开提供进一步的说明。除非另有指明,本文使用的所有技术和科学术语具有与本公开所属技术领域的普通技术人员通常理解的含义。

[0037] 需要注意的是,这里所使用的术语仅是为了描述具体实施方式,而非意图限制根据本公开的示例性实施方式。如在这里所使用的,除非上下文另外明确指出,否则单数形式也意图包括复数形式,此外,还应当理解的是,当在本说明书中使用术语“包含”和/或“包括”时,其指明存在特征、步骤、操作、器件、组件和/或它们的组合。

[0038] 下面结合附图1对本公开提出的基于深度多任务学习的图像美感和情感联合分类方法作详细说明。

[0039] 如图1所示,本实施例的一种基于深度多任务学习的图像美感和情感联合分类方法,包括:

[0040] S101:标注图像对应的美感类别和情感类别,形成训练数据集。

[0041] 在具体实施中,在图像美感分类问题中,将图像分为高美感和低美感两类;在图像情感分类问题中,将图像分为愉悦、敬畏、满足、兴奋、愤怒、厌恶、恐惧、悲伤共计八个基础情感类别。

[0042] 由于人们的美感和情感都是主观性很强的认知属性,存在明显的个体差异。因此,对于图像的美感类别和情感类别的标注,采用多人对同一幅图像进行共同标注的策略,之后取共识度最高的类别作为图像最终的类别。

[0043] 可以理解的,在其他的实施例中,图像美感分类和图像情感分类也可分为其他类别,本领域技术人员可以根据具体情况自行设置,在此不作详述。

[0044] S102:构造包含跨分支连接层和两个并行网络分支的深度卷积神经网络;

[0045] 其中,两个网络分支分别负责对输入图像进行美感分类和情感分类;跨分支连接层用于连接两个网络分支中对应的卷积层组,以关联美感分类和情感分类这两个任务;深

度卷积神经网络的输出代表输入图像属于各美感类别和各情感类别的概率。

[0046] 具体地,在所述深度卷积神经网络中,两个网络分支中的卷积层组数量相同均为 n ;跨分支连接层的数量为 $n-1$;第 i 个跨分支连接层将两个网络分支中第 i 个对应的卷积层组输出的图像特征图作为输入,并将这些输入的图像特征图沿着通道方向进行堆叠,将堆叠后的图像特征分别输入至两个网络分支中第 $i+1$ 个对应的卷积层组; $1 \leq i \leq n-1$; n 为大于或等于2的正整数。

[0047] 本实施例中的深度卷积神经网络如附图2所示。网络共包含两个并行分支,它们接受同一幅输入图像,并分别负责对输入图像进行美感分类和情感分类。每个网络分支的结构相同,都基于VGG16网络结构(参见Simonyan K,Zisserman A.Very deep convolutional networks for large-scale image recognition.arXiv preprint arXiv:1409.1556, 2014.)。每个网络分支由5个卷积层组、3个全连接层和1个Softmax层组成。其中,单个卷积层组内包含多个连续的卷积层和1个最大池化层,其目的是提取有效的图像特征图。全连接层对最后一个卷积层组输出的图像特征图进行多次非线性变换,将其映射为一个列向量。向量的维度等于美感类别或情感类别的个数,每一维对应一个特定的美感类别或情感类别。经过最终的Softmax层,该向量的每一维被转换为一个概率值,代表输入图像属于对应类别的概率。网络分支中每一层的具体结构和参数设定参照VGG16网络模型。

[0048] 引入跨分支连接层,将两个网络分支中对应的卷积层组进行连接,跨分支连接层的结构如附图3所示。跨分支连接层将两个卷积层组输出的图像特征图作为输入,并将它们沿着通道方向进行堆叠。假设堆叠前单个图像特征图的通道个数为 K (K 为正整数),则堆叠后图像特征图的通道个数为 $2K$ 。然后,将堆叠后的图像特征图分别输入两个卷积核大小为 1×1 的卷积层。这两个卷积层都包含 K 个卷积核,卷积核的步长和边缘填充大小分别为1和0。通过这种方式,两个卷积层将重新输出新的图像特征图,新的图像特征图的大小不变,通道个数恢复为 K ,最终分别将新的图像特征图送入一个网络分支后续的卷积层组或全连接层。从直观上看,跨分支连接层使两个网络分支之间以交换图像特征图的方式来共享信息,并方便模型在训练过程中自动地学习决定两个任务分别需要哪些信息;

[0049] 在传统的深度多任务学习方法中,不同任务通常被设置为共享较低的网络层,并在较高的网络层上维持各自的分支。在进行多任务学习训练之前,需要提前凭借经验人工地指定共享网络层。这种做法缺乏理论指导,对于共享网络层不合理的选择可能会导致方法性能的严重下滑。区别于上述方法,本实施例在所有网络层上都为不同任务设计单独的网络分支,通过跨分支连接层使网络分支之间可以以交换图像特征图的方式来共享信息,并在训练过程中自动地学习不同任务需要哪些信息,进而提高了分类准确性。

[0050] 需要说明的是,步骤101和步骤102之间的顺序可以根据本领域人员的具体情况来自行调整顺序。

[0051] S103:利用训练数据集来训练深度卷积神经网络,直至预定义损失函数达到最小。

[0052] 在具体实施中,利用训练数据集来训练深度卷积神经网络的过程,包括:

[0053] 统一训练数据集的所有图像的大小尺寸;

[0054] 初始化深度卷积神经网络各层的权重,预定义损失函数;

[0055] 采用随机梯度下降算法对深度卷积神经网络进行训练,确定能使得损失函数最小的网络权重;且在每次训练迭代时,从图像的随机位置处裁剪出一块固定大小图像块,并以

一定概率对图像块进行水平翻转。

[0056] 利用训练数据集来训练深度卷积神经网络的过程中,首先,将所有训练图像放缩到统一大小的尺寸,本实施例将图像放缩到256*256像素;然后,计算训练图像的像素平均值,并使每一幅图像分别减去该均值,该操作可以使训练图像移除共同部分,凸显训练图像的个体差异;最后,在每次训练迭代时,从减去均值的图像的随机位置处裁剪出一块固定大小图像块,并以一定概率对图像块进行水平翻转。通过这种方式,可以有效地扩充训练样本的数量,提升训练样本的多样性。本实施例选取的是224*224像素大小的图像块,每次进行水平翻转操作的概率为0.5。

[0057] 除最后一层全连接层和跨分支连接层外,每个网络分支各层的权重均采用在ImageNet数据集上预训练的VGG16模型的权重来初始化,对最后一层全连接层和跨分支连接层的权重进行随机初始化。采用交叉熵损失函数,定义在美感分类上的损失为 L_a ,在情感分类上的损失为 L_e ,即

$$[0058] \quad L_a = -y_a \log p_a - (1-y_a) \log (1-p_a)$$

$$[0059] \quad L_e = -\sum_{e=1}^8 y_e \log p_e$$

[0060] 其中, y_a 表明输入图像的真实美学类别,若图像实际为高美感图像,其值为1;否则,其值为0。 y_e 表明输入图像的真实情感类别,若图像实际属于第 e 个情感类别,其值为1;否则,其值为0。 p_a 代表网络输出的图像属于高美感类别的概率, p_e 代表网络输出的图像属于第 e 个情感类别的概率。

[0061] 进一步地,总的损失函数为 $L = L_a + \lambda L_e$ 。其中, λ 为平衡模型两类损失的超参数。在本实施例中,考虑美感分类为二分类问题,而情感分类为多分类问题,故 λ 设定的值为1/4。采用随机梯度下降算法对网络进行训练,确定能使得损失函数最小的网络权重。

[0062] S104:利用训练得到的深度卷积神经网络输出给定图像属于各美感类别和各情感类别的概率,选取美感类别和情感类别中概率最大的类别分别作为给定图像的预测美感类别和情感类别。

[0063] 在本实施例中,给定一幅图像,首先将其放缩到224*224像素,然后将图像输入训练好的网络,得到它属于各美感类别和各情感类别的概率,最后选取概率最大的类别作为图像的预测美感类别和情感类别。

[0064] 本实施例将多任务学习的思想运用到图像的美感分类和情感分类上来,充分利用了两个任务之间的关联特性,并设计了一个统一的深度卷积神经网络框架,通过跨分支连接层使网络分支之间可以以交换图像特征图的方式使两个任务之间可有效地共享信息,并在训练过程中自动地学习不同任务需要哪些信息,实现对图像的美感类别和情感类别联合识别,提高了图像美感类别和情感类别的准确性。

[0065] 下面结合附图4对本公开提出的基于深度多任务学习的图像美感和情感联合分类系统作详细说明。

[0066] 如图4所示,本实施例的一种基于深度多任务学习的图像美感和情感联合分类系统,包括:训练数据集形成模块11、深度卷积神经网络构造模块12、深度卷积神经网络训练模块13和预测分类模块14。

[0067] 其中:

[0068] 训练数据集形成模块11,其用于标注图像对应的美感类别和情感类别,形成训练数据集。

[0069] 在具体实施中,在图像美感分类问题中,将图像分为高美感和低美感两类;在图像情感分类问题中,将图像分为愉悦、敬畏、满足、兴奋、愤怒、厌恶、恐惧、悲伤共计八个基础情感类别。

[0070] 由于人们的美感和情感都是主观性很强的认知属性,存在明显的个体差异。因此,对于图像的美感类别和情感类别的标注,采用多人对同一幅图像进行共同标注的策略,之后取共识度最高的类别作为图像最终的类别。

[0071] 可以理解的,在其他的实施例中,图像美感分类和图像情感分类也可分为其他类别,本领域技术人员可以根据具体情况自行设置,在此不作详述。

[0072] 深度卷积神经网络构造模块12,其用于构造包含跨分支连接层和两个并行网络分支的深度卷积神经网络。

[0073] 其中,两个网络分支分别负责对输入图像进行美感分类和情感分类;跨分支连接层用于连接两个网络分支中对应的卷积层组,以关联美感分类和情感分类这两个任务;深度卷积神经网络的输出代表输入图像属于各美感类别和各情感类别的概率。

[0074] 具体地,在所述深度卷积神经网络中,两个网络分支中的卷积层组数量相同均为 n ;跨分支连接层的数量为 $n-1$;第 i 个跨分支连接层将两个网络分支中第 i 个对应的卷积层组输出的图像特征图作为输入,并将这些输入的图像特征图沿着通道方向进行堆叠,将堆叠后的图像特征分别输入至两个网络分支中第 $i+1$ 个对应的卷积层组; $1 \leq i \leq n-1$; n 为大于或等于2的正整数。

[0075] 本实施例中的深度卷积神经网络如附图2所示。网络共包含两个并行分支,它们接受同一幅输入图像,并分别负责对输入图像进行美感分类和情感分类。每个网络分支的结构相同,都基于VGG16网络结构(参见Simonyan K,Zisserman A.Very deep convolutional networks for large-scale image recognition.arXiv preprint arXiv:1409.1556, 2014.)。每个网络分支由5个卷积层组、3个全连接层和1个Softmax层组成。其中,单个卷积层组内包含多个连续的卷积层和1个最大池化层,其目的是提取有效的图像特征图。全连接层对最后一个卷积层组输出的图像特征图进行多次非线性变换,将其映射为一个列向量。向量的维度等于美感类别或情感类别的个数,每一维对应一个特定的美感类别或情感类别。经过最终的Softmax层,该向量的每一维被转换为一个概率值,代表输入图像属于对应类别的概率。网络分支中每一层的具体结构和参数设定参照VGG16网络模型。

[0076] 引入跨分支连接层,将两个网络分支中对应的卷积层组进行连接,跨分支连接层的结构如附图3所示。跨分支连接层将两个卷积层组输出的图像特征图作为输入,并将它们沿着通道方向进行堆叠。假设堆叠前单个图像特征图的通道个数为 K (K 为正整数),则堆叠后图像特征图的通道个数为 $2K$ 。然后,将堆叠后的图像特征图分别输入两个卷积核大小为 1×1 的卷积层。这两个卷积层都包含 K 个卷积核,卷积核的步长和边缘填充大小分别为1和0。通过这种方式,两个卷积层将重新输出新的图像特征图,新的图像特征图的大小不变,通道个数恢复为 K ,最终分别将新的图像特征图送入一个网络分支后续的卷积层组或全连接层。从直观上看,跨分支连接层使两个网络分支之间以交换图像特征图的方式来共享信息,并方便模型在训练过程中自动地学习决定两个任务分别需要哪些信息;

[0077] 在传统的深度多任务学习方法中,不同任务通常被设置为共享较低的网络层,并在较高的网络层上维持各自的分支。在进行多任务学习训练之前,需要提前凭借经验人工地指定共享网络层。这种做法缺乏理论指导,对于共享网络层不合理的选择可能会导致方法性能的严重下滑。区别于上述方法,本实施例在所有网络层上都为不同任务设计单独的网络分支,通过跨分支连接层使网络分支之间可以以交换图像特征图的方式来共享信息,并在训练过程中自动地学习不同任务需要哪些信息,进而提高了分类准确性。

[0078] 深度卷积神经网络训练模块13,其用于利用训练数据集来训练深度卷积神经网络,直至预定义损失函数达到最小。

[0079] 所述深度卷积神经网络训练模块13,包括:

[0080] 尺寸统一模块131,其用于统一训练数据集的所有图像的大小尺寸;

[0081] 初始化模块132,其用于初始化深度卷积神经网络各层的权重,预定义损失函数;

[0082] 迭代训练模块133,其用于采用随机梯度下降算法对深度卷积神经网络进行训练,确定能使得损失函数最小的网络权重;且在每次训练迭代时,从图像的随机位置处裁剪出一块固定大小图像块,并以一定概率对图像块进行水平翻转。

[0083] 利用训练数据集来训练深度卷积神经网络的过程中,首先,将所有训练图像放缩到统一大小的尺寸,本实施例将图像放缩到256*256像素;然后,计算训练图像的像素平均值,并使每一幅图像分别减去该均值,该操作可以使训练图像移除共同部分,凸显训练图像的个体差异;最后,在每次训练迭代时,从减去均值的图像的随机位置处裁剪出一块固定大小图像块,并以一定概率对图像块进行水平翻转。通过这种方式,可以有效地扩充训练样本的数量,提升训练样本的多样性。本实施例选取的是224*224像素大小的图像块,每次进行水平翻转操作的概率为0.5。

[0084] 除最后一层全连接层和跨分支连接层外,每个网络分支各层的权重均采用在ImageNet数据集上预训练的VGG16模型的权重来初始化,对最后一层全连接层和跨分支连接层的权重进行随机初始化。采用交叉熵损失函数,定义在美感分类上的损失为 L_a ,在情感分类上的损失为 L_e ,即

[0085] $L_a = -y_a \log p_a - (1-y_a) \log (1-p_a)$

[0086] $L_e = -\sum_{e=1}^8 y_e \log p_e$

[0087] 其中, y_a 表明输入图像的真实美学类别,若图像实际为高美感图像,其值为1;否则,其值为0。 y_e 表明输入图像的真实情感类别,若图像实际属于第 e 个情感类别,其值为1;否则,其值为0。 p_a 代表网络输出的图像属于高美感类别的概率, p_e 代表网络输出的图像属于第 e 个情感类别的概率。

[0088] 进一步地,总的损失函数为 $L = L_a + \lambda L_e$ 。其中, λ 为平衡模型两类损失的超参数。在本实施例中,考虑美感分类为二分类问题,而情感分类为多分类问题,故 λ 设定的值为1/4。采用随机梯度下降算法对网络进行训练,确定能使得损失函数最小的网络权重。

[0089] 预测分类模块14,其用于利用训练得到的深度卷积神经网络输出给定图像属于各美感类别和各情感类别的概率,选取美感类别和情感类别中概率最大的类别分别作为给定图像的预测美感类别和情感类别。

[0090] 在本实施例中,给定一幅图像,首先将其放缩到224*224像素,然后将图像输入训

练好的网络,得到它属于各美感类别和各情感类别的概率,最后选取概率最大的类别作为图像的预测美感类别和情感类别。

[0091] 本实施例将多任务学习的思想运用到图像的美感分类和情感分类上来,充分利用了两个任务之间的关联特性,并设计了一个统一的深度卷积神经网络框架,通过跨分支连接层使网络分支之间可以以交换图像特征图的方式使两个任务之间可有效地共享信息,并在训练过程中自动地学习不同任务需要哪些信息,实现对图像的美感类别和情感类别联合识别,提高了图像美感类别和情感类别的准确性。

[0092] 在另一实施例中,提供了一种计算机可读存储介质,其上存储有计算机程序,该程序被处理器执行时实现如图1所示的基于深度多任务学习的图像美感和情感联合分类方法中的步骤。

[0093] 在另一实施例中,提供了一种计算机设备,包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序,所述处理器执行所述程序时实现如图1所示的基于深度多任务学习的图像美感和情感联合分类方法中的步骤。

[0094] 本领域内的技术人员应明白,本公开的实施例可提供为方法、系统、或计算机程序产品。因此,本公开可采用硬件实施例、软件实施例、或结合软件和硬件方面的实施例的形式。而且,本公开可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器和光学存储器等)上实施的计算机程序产品的形式。

[0095] 本公开是参照根据本公开实施例的方法、设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器,使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

[0096] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中,使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品,该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

[0097] 这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上,使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理,从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

[0098] 本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程,是可以通过计算机程序来指令相关的硬件来完成,所述的程序可存储于一计算机可读取存储介质中,该程序在执行时,可包括如上述各方法的实施例的流程。其中,所述的存储介质可为磁碟、光盘、只读存储记忆体(Read-Only Memory,ROM)或随机存储记忆体(Random AccessMemory, RAM)等。

[0099] 以上所述仅为本公开的优选实施例而已,并不用于限制本公开,对于本领域的技术人员来说,本公开可以有各种更改和变化。凡在本公开的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本公开的保护范围之内。

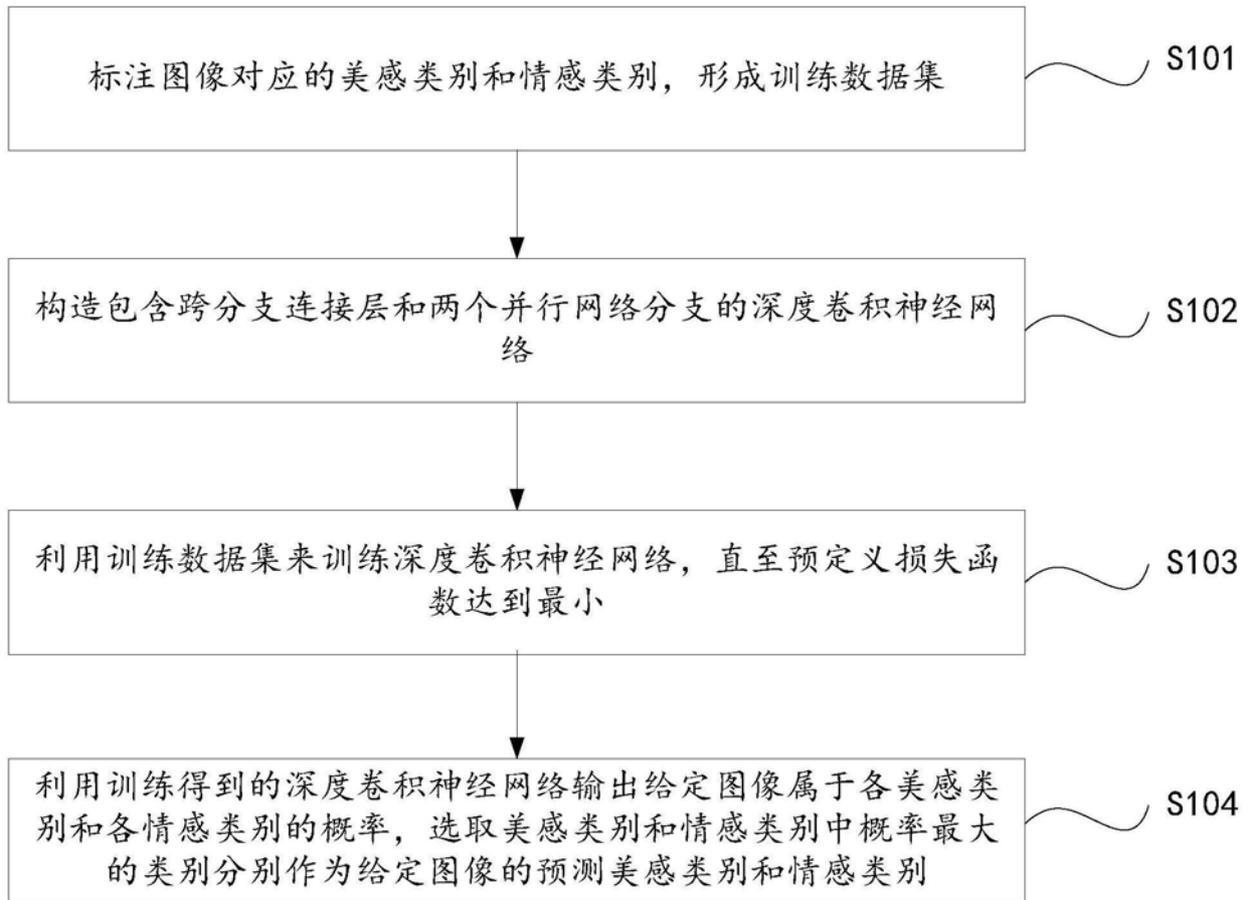


图1

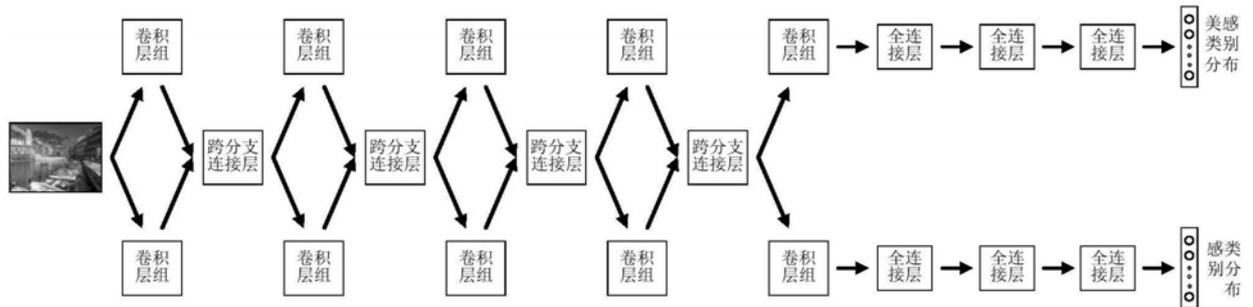


图2

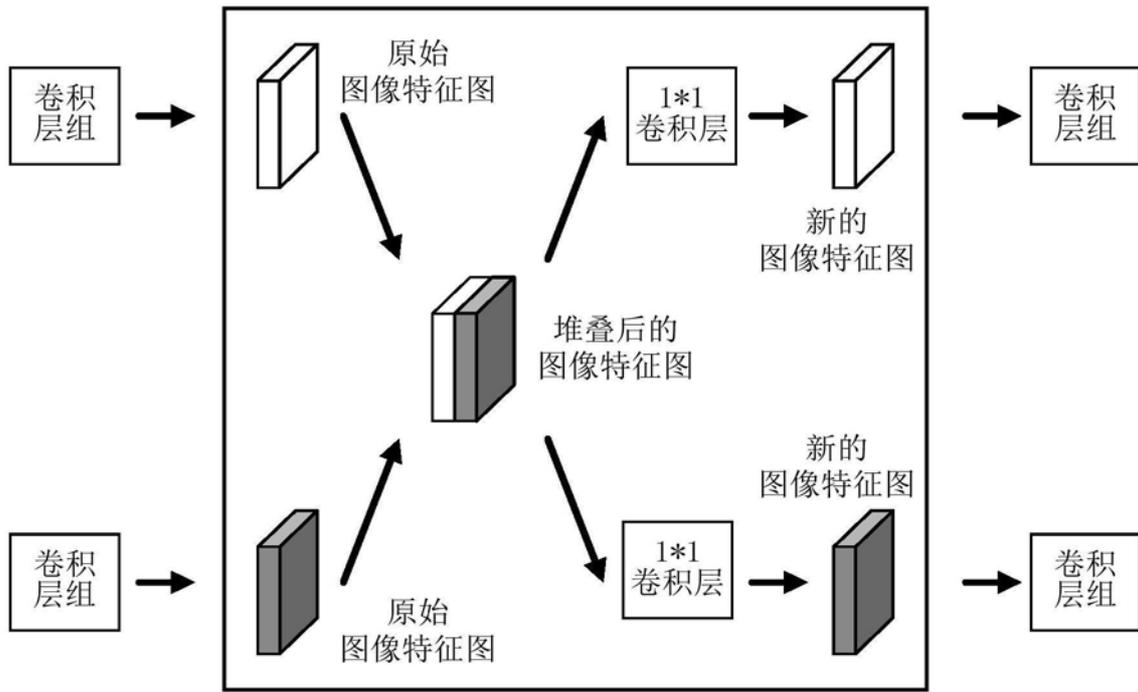


图3

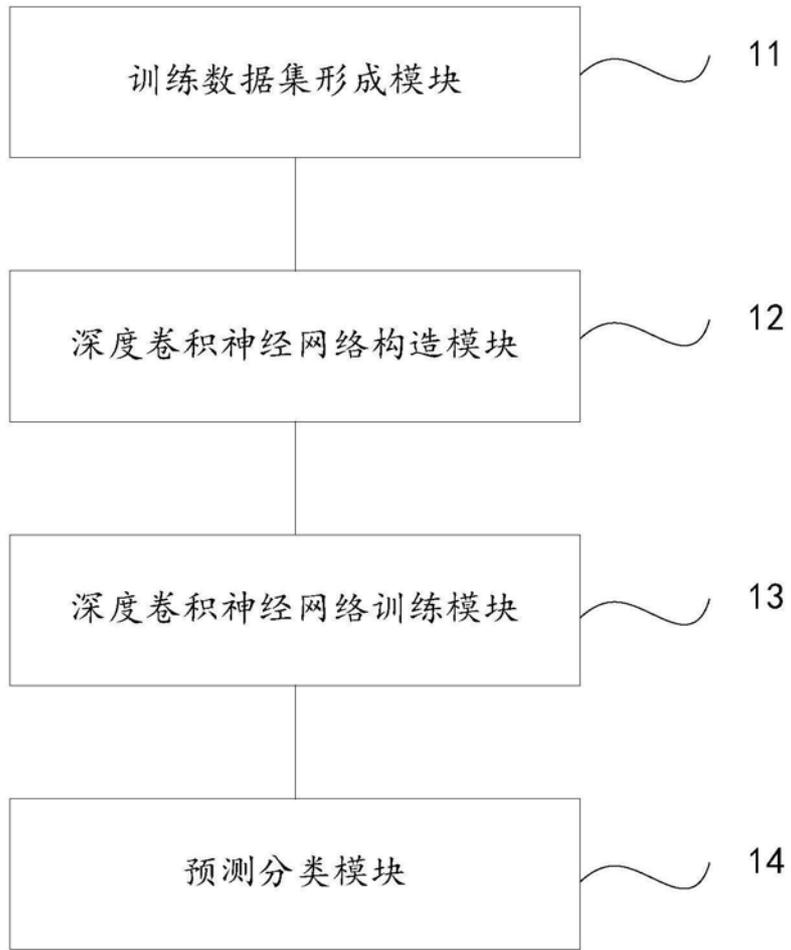


图4