



(12)发明专利

(10)授权公告号 CN 107330451 B

(45)授权公告日 2020.06.26

(21)申请号 201710456031.1

(22)申请日 2017.06.16

(65)同一申请的已公布的文献号
申请公布号 CN 107330451 A

(43)申请公布日 2017.11.07

(73)专利权人 西交利物浦大学
地址 215123 江苏省苏州市工业园区独墅湖高等教育区仁爱路111号

(72)发明人 张百灵 夏翌彰 武芳宇 吕文进

(74)专利代理机构 苏州创元专利商标事务所有
限公司 32103

代理人 范晴 丁浩秋

(51)Int.Cl.

G06K 9/62(2006.01)

G06N 3/08(2006.01)

(56)对比文件

CN 106485215 A,2017.03.08,

CN 105447529 A,2016.03.30,

CN 105117739 A,2015.12.02,

CN 104423945 A,2015.03.18,

US 2016328606 A1,2016.11.10,

US 2014376819 A1,2014.12.25,

CN 104992142 A,2015.10.21,

审查员 石松婷

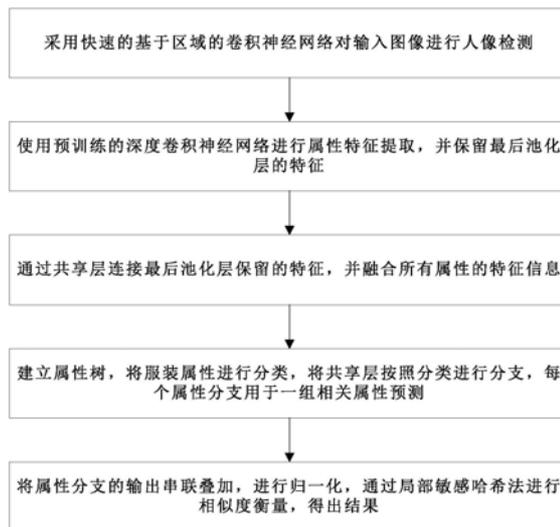
权利要求书1页 说明书7页 附图3页

(54)发明名称

基于深度卷积神经网络的服装属性检索方法

(57)摘要

本发明公开了一种基于深度卷积神经网络的服装属性检索方法,包括:采用快速的基于区域的卷积神经网络对输入图像进行人像检测;使用预训练的深度卷积神经网络进行属性特征提取,并保留最后池化层的特征;通过共享层连接最后池化层保留的特征,并融合所有属性的特征信息;建立属性树,将服装属性进行分类,将共享层按照分类进行分支,每个属性分支用于一组相关属性预测;将属性分支的输出串联叠加,进行归一化,通过局部敏感哈希法进行相似度衡量,得出结果。可以将服装属性的特征描述用于部分服装属性检测,可显著提高服装属性预测的准确率。



1. 一种基于深度卷积神经网络的服装属性检索方法,其特点在于,包括以下步骤:

S01:采用快速的基于区域的卷积神经网络对输入图像进行人像检测;

S02:使用预训练的深度卷积神经网络进行属性特征提取,并保留最后池化层的特征;

S03:通过共享层连接最后池化层保留的特征,并融合所有属性的特征信息;

S04:建立属性树,将服装属性进行分类,将共享层按照分类进行分支,每个属性分支用于一组相关属性预测;属性分支在进行属性预测的前馈过程中,共享层将副本分配给每个属性分支,而在后向传播过程中,共享层积累每个属性分支的梯度;

S05:将属性分支的输出串联叠加,进行归一化,通过局部敏感哈希法进行相似度衡量,得出结果。

2. 根据权利要求1所述的基于深度卷积神经网络的服装属性检索方法,其特征在在于,所述步骤S01具体包括:

使用选择性搜索在输入图像中生成多个感兴趣区域候选框,输入到全卷积的卷积神经网络中,在最后一个卷积层上对每个感兴趣区域求映射关系,并用感兴趣区域池化层统一大小;

通过全连接层得到特征向量,特征向量经由各自的全连接层,得到两个输出向量:一个是使用交叉熵损失函数得到的分类,另一个是每一类的边界位置回归。

3. 根据权利要求1所述的基于深度卷积神经网络的服装属性检索方法,其特征在在于,所述步骤S04中的属性分支包括类别、性别、颜色和图案。

4. 根据权利要求1所述的基于深度卷积神经网络的服装属性检索方法,其特征在在于,所述步骤S05中进行归一化后,使用主成份分析法进行降维,具体步骤如下:

S11:将原始特征数据按行排列组成矩阵X,并对X进行数据标准化,使其均值变为零;

S12:求X的协方差矩阵C,将特征值按照从大到小的顺序排序,选择其中最大的k个,然后将其对应的k个特征向量分别作为列向量组成特征向量矩阵P;

S13:通过计算 $Y=PX$,得到降维后数据Y。

5. 根据权利要求1所述的基于深度卷积神经网络的服装属性检索方法,其特征在在于,所述步骤S05中通过局部敏感哈希法进行相似度衡量包括离线建立索引和在线查找两个步骤,具体步骤如下:

S21:选取满足条件的局部敏感哈希的哈希函数,根据对查找结果的准确率确定哈希表的个数L,每个哈希表内的哈希函数的个数K,以及与哈希函数自身有关的参数;将所有数据经过哈希函数哈希到相应的桶内,构成了一个或多个哈希表;

S22:将查找数据经过哈希函数哈希得到相应的桶号,将桶号中对应的前2L个数据取出,最后计算查询数据与这2L个数据之间的相似度或距离,返回最近邻的数据。

基于深度卷积神经网络的服装属性检索方法

技术领域

[0001] 本发明涉及一种基于卷积神经网络的服装检索方法,具体地涉及一种基于深度卷积神经网络的服装属性检索方法。

背景技术

[0002] 随着互联网和服装电子商务的迅速发展,网络购物市场逐年扩大,如何利用检索技术帮助用户快速找到心仪的服装是一个非常重要的任务。然而,服装检测的识别难度很高:其一,衣服形变大,服装本身是柔性很大的物体,人不同的姿势会导致服装的形状不同;其二,在不同的光照情况和复杂的场景下,区分不同服装类型的困难也会增加;此外,衣服的设计包含了大量细节属性,比如领型、版型、颜色、装饰等等,想要区分出它们的难度很大。

[0003] 目前广泛使用的服装检索方法是利用分类器来区分各种服装属性,其成功的关键是首先对于各种类型服装的属性进行特征描述,其次选择合适的分类器来学习训练样本,从而得到模型。

[0004] 一般来说,大部分的基于手工提取特征的方法主要使用边缘梯度直方图(Histogram of Oriented Gradient,HOG)和分层梯度方向直方图(Pyramid Histogram of Oriented Gradients,PHOG)。但是这些方法的瓶颈是手工提取特征只对某一方面敏感。比如,边缘梯度直方图只对梯度敏感。因此,研究人员提出了新的方法,例如:将属性概念引入服装检索、采用细粒度的属性分类。

[0005] 另一方面,卷积神经网络已经进入到服装属性检测领域,相比于手工提取特征的方法,卷积神经网络大大提高了服装检测的性能。为了研究如何在网上商店匹配街头服装照片的实际问题,双重属性感知网络和局部共享卷积神经网络架构也被提出。例如中国专利文献CN 106250423公开了一种基于部分参数共享的深度卷积神经网络跨域服装检索方法该方法利用了深度神经网络模型中越高层的网络层的特征越具有语义信息,与训练数据所在域的关系越密切,越低层的网络层的特征越具有一般性,越与训练数据所在域无关,依此让传统的深度卷积神经网络跨域服装检索模型中的低层的网络层参数共享,而高层的网络层参数由各个域保有。使用更少参数但却获得相同效果的方法,通过利用深度神经网络模型内在的特性,显著性减少模型参数数量。但是其并没有改善服装属性预测的准确率。本发明因此而来。

发明内容

[0006] 针对上述存在的缺陷,本发明的目的是提出了一种基于深度卷积神经网络的服装属性检索方法。利用多任务学习和卷积神经网络结合的服装属性检测模型,同时对于提取的各个服装属性特征构建适当的属性树,可以将服装属性的特征描述用于部分服装属性检测,可显著提高服装属性预测的准确率。

[0007] 本发明的技术方案是:

- [0008] 一种基于深度卷积神经网络的服装属性检索方法,包括以下步骤:
- [0009] S01:采用快速的基于区域的卷积神经网络对输入图像进行人像检测;
- [0010] S02:使用预训练的深度卷积神经网络进行属性特征提取,并保留最后池化层的特征;
- [0011] S03:通过共享层连接最后池化层保留的特征,并融合所有属性的特征信息;
- [0012] S04:建立属性树,将服装属性进行分类,将共享层按照分类进行分支,每个属性分支用于一组相关属性预测;
- [0013] S05:将属性分支的输出串联叠加,进行归一化,通过局部敏感哈希法进行相似度衡量,得出结果。
- [0014] 优选的,所述步骤S01具体包括:
- [0015] 使用选择性搜索在输入图像中生成多个感兴趣区域候选框,输入到全卷积的卷积神经网络中,在最后一个卷积层上对每个感兴趣区域求映射关系,并用感兴趣区域池化层统一大小;
- [0016] 通过全连接层得到特征向量,特征向量经由各自的全连接层,得到两个输出向量:一个是使用交叉熵损失函数得到的分类,另一个是每一类的边界位置回归。
- [0017] 优选的,所述步骤S04中的属性分支包括类别、性别、颜色和图案。
- [0018] 优选的,所述步骤S04中属性分支在进行属性预测的前馈过程中,共享层将副本分配给每个属性分支,而在后向传播过程中,共享层积累每个属性分支的梯度。
- [0019] 优选的,所述步骤S05中进行归一化后,使用主成份分析法进行降维,具体步骤如下:
- [0020] S11:将原始特征数据按行排列组成矩阵X,并对X进行数据标准化,使其均值变为零;
- [0021] S12:求X的协方差矩阵C,将特征值按照从大到小的顺序排序,选择其中最大的k个,然后将其对应的k个特征向量分别作为列向量组成特征向量矩阵P;
- [0022] S13:通过计算 $Y=PX$,得到降维后数据Y。
- [0023] 优选的,所述步骤S05中通过局部敏感哈希法进行相似度衡量包括离线建立索引和在线查找两个步骤,具体步骤如下:
- [0024] S21:选取满足条件的局部敏感哈希的哈希函数,根据对查找结果的准确率确定哈希表的个数L,每个哈希表内的哈希函数的个数K,以及与哈希函数自身有关的参数;将所有数据经过哈希函数哈希到相应的桶内,构成了一个或多个哈希表;
- [0025] S22:将查找数据经过哈希函数哈希得到相应的桶号,将桶号中对应的前2L个数据取出,最后计算查询数据与这2L个数据之间的相似度或距离,返回最近邻的数据。
- [0026] 与现有技术相比,本发明的优点是:
- [0027] 将多任务学习和卷积神经网络结合作为服装属性检测模型,同时对于提取的各个服装属性特征构建适当的属性树。本发明可以将服装属性的特征描述用于部分服装属性检测。此方法可显著提高服装属性预测的准确率。

附图说明

- [0028] 下面结合附图及实施例对本发明作进一步描述:

- [0029] 图1为本发明基于深度卷积神经网络的服装属性检索方法的流程图；
- [0030] 图2为服装属性相关特征提取及构造属性树；
- [0031] 图3为服装检索示例；
- [0032] 图4是一个基于不同属性分支的服装检索的对比示例。

具体实施方式

[0033] 为使本发明的目的、技术方案和优点更加清楚明了,下面结合具体实施方式并参照附图,对本发明进一步详细说明。应该理解,这些描述只是示例性的,而并非要限制本发明的范围。此外,在以下说明中,省略了对公知结构和技术的描述,以避免不必要地混淆本发明的概念。

[0034] 如图1所示,本发明基于深度卷积神经网络的服装属性检索方法,包括以下步骤:

[0035] 第一步:采用快速的基于区域的卷积神经网络(Fast Region-based Convolutional Network, Fast RCNN)作为目标检测器,从复杂的背景图片中检测人像。目标检测具体步骤如下:

[0036] 1、使用选择性搜索(Selective Search)在一张图片中生成约2000个候选框(可能会包含人像的矩形区域),即感兴趣区域(Region of Interest, RoI)。

[0037] 2、把它们整体输入到全卷积的网络中,在最后一个卷积层上对每个感兴趣区域求映射关系,并用一个感兴趣区域池化层来统一到相同的大小。

[0038] 3、继续经过两个全连接层(Fully Connected Layers, FC)得到特征向量。特征向量经由各自的全连接层,得到两个输出向量:第一个是分类,使用交叉熵损失函数,第二个是每一类的精确边界位置回归。

[0039] 整个过程是一个实时的目标检测框架,使用多任务损失的端到端训练,大幅提升了目标检测的速度。

[0040] 第二步:基于第一步的目标检测结果,结合卷积神经网络(ResNet50及ResNet152,网络结构见表1)和多任务学习来预测服装属性,并提取属性相关的特征。为了更好的描述服装属性之间的关系,同时提高服装属性的分类效果,鉴于属性相关的先验知识,我们建立了属性树。在属性树中,服装属性被分为四类:类别,性别,颜色和图案。

| 层的名字 | 输出大小 | ResNet50 | ResNet152 |
|--------|---------|---|--|
| 卷积 1 | 112×112 | 7×7, 64, 每像素占用的字节数为 2 | |
| 卷积 2_x | 56×56 | 3×3 最大值池化, 每像素占用的字节数为 2 | |
| | | $\begin{bmatrix} 1 \times 1, & 64 \\ 3 \times 3, & 64 \\ 1 \times 1, & 256 \end{bmatrix} \times 3$ | $\begin{bmatrix} 1 \times 1, & 64 \\ 3 \times 3, & 64 \\ 1 \times 1, & 256 \end{bmatrix} \times 3$ |
| 卷积 3_x | 28×28 | $\begin{bmatrix} 1 \times 1, & 128 \\ 3 \times 3, & 128 \\ 1 \times 1, & 512 \end{bmatrix} \times 4$ | $\begin{bmatrix} 1 \times 1, & 128 \\ 3 \times 3, & 128 \\ 1 \times 1, & 512 \end{bmatrix} \times 8$ |
| 卷积 4_x | 14×14 | $\begin{bmatrix} 1 \times 1, & 256 \\ 3 \times 3, & 256 \\ 1 \times 1, & 1024 \end{bmatrix} \times 6$ | $\begin{bmatrix} 1 \times 1, & 256 \\ 3 \times 3, & 256 \\ 1 \times 1, & 1024 \end{bmatrix} \times 36$ |
| 卷积 5_x | 7×7 | $\begin{bmatrix} 1 \times 1, & 512 \\ 3 \times 3, & 512 \\ 1 \times 1, & 2048 \end{bmatrix} \times 3$ | $\begin{bmatrix} 1 \times 1, & 512 \\ 3 \times 3, & 512 \\ 1 \times 1, & 2048 \end{bmatrix} \times 3$ |
| | 1×1 | 平均池化, 1000-d 全连层, 使用交叉熵损失函数 | |

[0043] 表1 ResNet50及ResNet152网络结构

[0044] 如图2所示,具体过程如下:

[0045] 1、基于第一步中从复杂的背景图片中检测到的人像,使用预训练的卷积神经网络(ResNet50及ResNet152)进行特征提取,并保留第五个池化层的特征。

[0046] 2、通过共享层连接上一步保留的特征,并融合所有属性的特征信息。

[0047] 3、最后,共享层被分为四个服装属性分支,包括类别,性别,颜色和图案。在前馈过程中,共享层将副本分配给每个分支,而在后向传播过程中,共享层积累每个分支的梯度。

[0048] $InB_i = Out_s$ (1)

[0049] $G_s = \sum_{i=1}^4 GB_i$ (2)

[0050] 在等式(1)和等式(2)中, InB_i 是分支*i*前馈运算的输入, Out_s 是共享层的输出。 G_s 和 GB_i 是共享层的梯度和反向传播中的分支*i*的梯度。

[0051] 第三步:使用主成分分析(Principal Component Analysis,PCA)和局部敏感哈希法(Locality-Sensitive Hashing,LSH)进行相似度衡量。丰富的特征通常由高维向量表示。理想的相似度衡量通常需要满足以下四个条件:高精度、低空间复杂度、低时间复杂度和支持高维度。

[0052] 在本发明中,我们运用主成分分析法丢弃冗余特征,降低特征的维数,同时保证特征的最大协方差。在对特征进行主成分分析之后,后续的步骤会被加快。主成分分析的具体步骤如下:

[0053] 1、将原始特征数据按行排列组成矩阵X;

[0054] 2、对X进行数据标准化,使其均值变为零;

[0055] 3、求X的协方差矩阵C；

[0056] 4、将特征值按照从大到小的顺序排序，选择其中最大的k个，然后讲其对应的k个特征向量分别作为列向量组成特征向量矩阵P；

[0057] 5、通过计算 $Y=PX$ ，得到降维后数据Y；

[0058] 接下来为了检索出与查询服装图片属性相似的图片集合，我们对数据库中训练集图片建立局部敏感哈希索引，然后通过局部敏感哈希索引来查找与测试数据集图片相似的训练测试集图片，具体过程如下：

[0059] 一、离线建立索引

[0060] 首先选取满足条件的局部敏感哈希的哈希函数 (Locality-Sensitive Hashing hash function, LSH hash function)，根据对查找结果的准确率确定哈希表的个数L，每个哈希表内的哈希函数的个数K，以及与LSH hash function自身有关的参数；接下来将所有数据经过LSH hash function哈希到相应的桶内，构成了一个或多个哈希表；

[0061] 二、在线查找

[0062] 将查找数据经过LSH hash fuction哈希得到相应的桶号，将桶号中对应的前2L个数据取出，最后计算查询数据与这2L个数据之间的相似度或距离，返回最近邻的数据。

[0063] 本发明应用于具体的实例如下：

[0064] 本发明应用于一个公开的服装属性数据集，该数据集包括26种服装属性。实验开展方式为：基于卷积神经网络的MATLAB工具箱 (MatCovNet)，每16张样本图片进行一次随即梯度下降运算，学习率为0.00001。数据集被随机分为独立的三部分，训练集，验证集和测试集。其中训练集用来拟合模型，验证集用来估测模型收敛情况，而测试集用来评估模型性能。三部分分别占总样本的70%，10%和20%。

[0065] 我们分别将收敛的152层的残差网 (ResNet152) 和50层的残差网 (ResNet50) 作为多任务卷积神经网络的初始化状态，从共享层之后的多任务卷积神经网络层参数被随机初始化。表2为采用本发明技术方案的服装属性预测准确率，可以看到，在ResNet152和多任务学习组合下，本发明的技术方案获得了较高的服装属性预测准确率。表3给出了在同多任务学习组合下，ResNet152和ResNet50的时间复杂度。

| | ResNet152 +MTL | ResNet50 +MTL | Before CRF | After CRF |
|------------------|---------------------------|--------------------------|-------------------|------------------|
| [0066] 种类 | 62.50 | 61.00 | 48.46 | 54.91 |
| 性别 | 88.73 | 87.28 | 82.29 | 81.07 |
| 黑色 | 85.51 | 83.77 | 84.91 | 84.07 |

| | | | | | |
|--------|--------|-------|-------|-------|-------|
| [0067] | 蓝色 | 94.78 | 93.04 | 90.23 | 90.70 |
| | 棕色 | 92.46 | 91.01 | 85.93 | 87.15 |
| | 蓝绿色 | 93.62 | 93.91 | 85.28 | 86.12 |
| | 灰色 | 81.74 | 81.45 | 79.11 | 78.93 |
| | 绿色 | 96.23 | 95.36 | 88.64 | 89.49 |
| | 紫色 | 95.94 | 95.65 | 80.42 | 82.85 |
| | 红色 | 96.52 | 96.23 | 89.95 | 92.29 |
| | 白色 | 82.90 | 83.19 | 75.56 | 76.40 |
| | 黄色 | 96.81 | 96.52 | 84.35 | 83.41 |
| | 大于两种颜色 | 91.30 | 92.46 | 77.90 | 79.02 |
| | 花卉图案 | 97.36 | 97.95 | 66.78 | 84.72 |
| | 图形 | 99.14 | 98.85 | 87.43 | 92.10 |
| | 格子图案 | 96.17 | 96.76 | 68.08 | 76.21 |
| | 固体 | 95.33 | 96.00 | 82.76 | 90.70 |
| | 斑点 | 99.12 | 97.64 | 67.43 | 76.87 |
| | 条纹 | 95.30 | 96.55 | 71.07 | 78.27 |
| | 平均值 | 91.66 | 91.30 | 78.77 | 82.38 |

[0068] 表2本发明下的服装属性预测准确率

| | | | |
|--------|-----------|----------------------|---------------------|
| [0069] | | ResNet152+MTL | ResNet50+MTL |
| | | (ms/f) | (ms/f) |
| | 训练 | 322.3 | 82.1 |
| | 测试 | 45.6 | 18.1 |

[0070] 表3多任务卷积神经网络的时间复杂度

[0071] 在本发明中,多任务卷积神经网络的每个分支用于一组相关属性预测,图3为服装检索示例。部分或者所有分支的输出被串联叠加起来,通过Z得分法(Z-score)归一化,主成份分析法降维。最后,将局部敏感哈希法用于搜索前30个最相似的服装。

[0072] 表4为检索具有某组属性的服装的准确率,表5为以不同分支为基础的服装属性检索准确率。

| | | | | | |
|--------|------------|--------------|--------------|-----------------|--------------------|
| [0073] | | 类别 | 类别+性别 | 类别+性别+颜色 | 类别+性别+颜色+图案 |
| | 种类 | 30.67 | 30.04 | 27.88 | 27.92 |
| | 性别 | 72.33 | 77.18 | 74.08 | 73.3 |
| | 黑色 | 82.94 | 82.90 | 84.16 | 83.93 |
| | 蓝色 | 86.74 | 86.72 | 86.51 | 87.11 |
| | 棕色 | 67.71 | 67.45 | 71.6 | 70.7 |
| | 蓝绿色 | 86.29 | 86.52 | 87.54 | 87.2 |
| | 灰色 | 84.77 | 84.63 | 85.61 | 85.51 |
| | 绿色 | 89.01 | 89.05 | 89.08 | 89.02 |
| | 紫色 | 67.85 | 68.41 | 71.42 | 70.96 |

| | | | | | |
|--------|--------|-------|-------|-------|-------|
| [0074] | 红色 | 90.24 | 90.24 | 90.24 | 90.27 |
| | 白色 | 90.17 | 90.19 | 90.24 | 90.15 |
| | 黄色 | 91.12 | 91.17 | 91.3 | 91.18 |
| | 大于两种颜色 | 69.47 | 68.52 | 72.39 | 71.92 |
| | 花卉图案 | 91.06 | 91.17 | 91.28 | 91.32 |
| | 图形 | 84.67 | 84.55 | 85.04 | 84.96 |
| | 格子图案 | 90.45 | 90.16 | 90.04 | 90.47 |
| | 固体 | 93.42 | 93.25 | 92.88 | 93.29 |
| | 斑点 | 89.32 | 89.3 | 89.18 | 89.48 |
| | 条纹 | 75.24 | 75.13 | 74.78 | 76.19 |
| | 平均值 | 89.67 | 89.78 | 89.68 | 90.31 |

[0075] 表4基于属性的服装检索准确率

| | | | | | |
|--------|----|-------|-------|-------|-------|
| [0076] | | 种类 | 性别 | 颜色 | 图案 |
| | 种类 | 30.67 | 72.33 | 82.94 | 86.74 |
| | 性别 | 30.04 | 77.18 | 82.90 | 86.72 |
| | 颜色 | 27.88 | 74.08 | 84.16 | 86.51 |
| | 图案 | 27.92 | 73.3 | 83.93 | 87.11 |

[0077] 表5以不同分支为基础的服装属性检索准确率

[0078] 图4是一个基于不同属性分支的服装检索的对比示例。当使用类别分支特征进行检索时,第二十四张检索结果为女性。然而当我们将类别和性别分支同时用于服装检索时,检索结果发生变化,在前三十张检索结果中没有女性。

[0079] 表6列出了本发明整个系统的运行时间,主要包括三个不同的模块。当系统在中央处理器,Intel (R) Xeon (R) E5-1650v3,运行时,每张图片的处理时间为101.5毫秒。

| | | | | | |
|--------|------------------|------------------|------------------|----------------|--------------|
| [0080] | | Fast RCNN | 多任务卷积神经网络 | 局部敏感哈希法 | 总计 |
| | 运行时间 ms/f | 74.6 | 18.1 | 8.8 | 101.5 |

[0081] 表6系统运行的时间复杂度

[0082] 应当理解的是,本发明的上述具体实施方式仅仅用于示例性说明或解释本发明的原理,而不构成对本发明的限制。因此,在不偏离本发明的精神和范围的情况下所做的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。此外,本发明所附权利要求旨在涵盖落入所附权利要求范围和边界、或者这种范围和边界的等同形式内的全部变化和修改例。

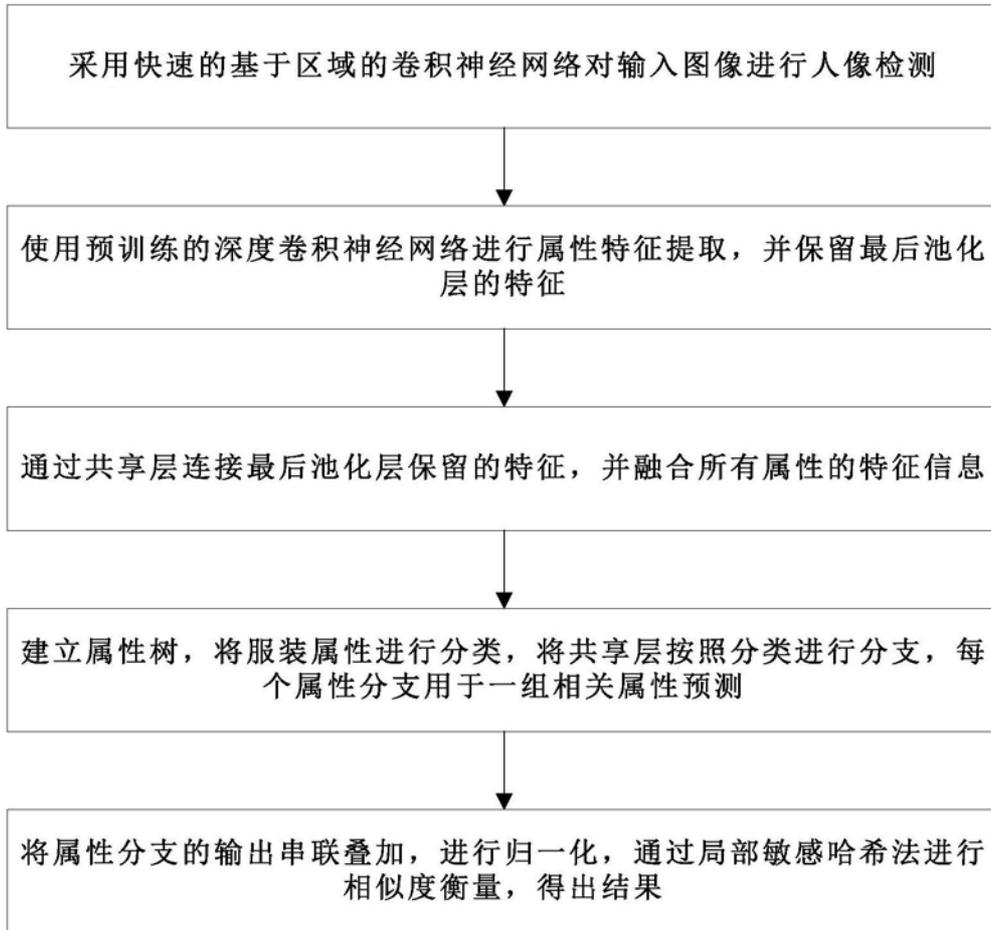


图1

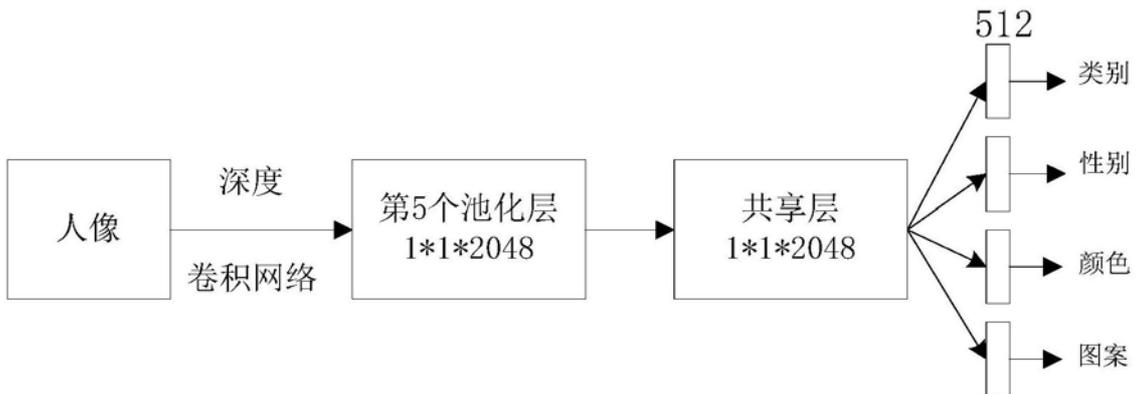


图2



图3



图4