

# (12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织  
国际局



(10) 国际公布号

WO 2020/199468 A1

(43) 国际公布日  
2020 年 10 月 8 日 (08.10.2020)

(51) 国际专利分类号:

G06K 9/32 (2006.01)

(21) 国际申请号:

PCT/CN2019/102186

(22) 国际申请日:

2019 年 8 月 23 日 (23.08.2019)

(25) 申请语言:

中文

(26) 公布语言:

中文

(30) 优先权:

201910269283.2 2019 年 4 月 4 日 (04.04.2019) CN

(71) 申请人: 平安科技 (深圳) 有限公司 (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN];

中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(72) 发明人: 庞焱 (PANG, Ye); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

王义文 (WANG, Yiwen); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong

518000 (CN)。王健宗 (WANG, Jianzong); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(74) 代理人: 深圳市沃德知识产权代理有限公司 (普通合伙) (SHENZHEN WORLD INTELLECTUAL PROPERTY AGENCY (GENERAL PARTNERSHIP)); 中国广东省深圳市福田区园岭街道八卦四路 10 号中浩大厦 1528-1530 室于志光, Guangdong 518000 (CN)。

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,

(54) Title: IMAGE CLASSIFICATION METHOD AND DEVICE, AND COMPUTER READABLE STORAGE MEDIUM

(54) 发明名称: 图像分类方法、装置及计算机可读存储介质

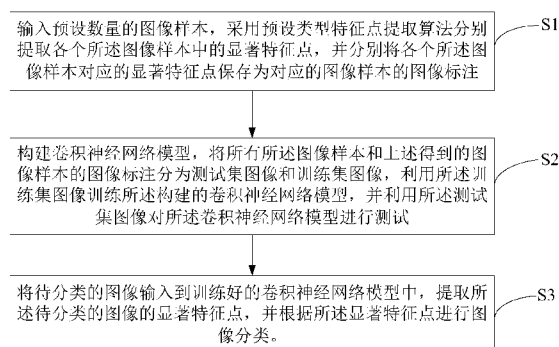


图 1

- S1 Input a preset number of image samples, extract a salient feature point in each of the image samples by using a preset type of feature point extraction algorithm, and respectively save the salient feature point corresponding to each of the image samples as an image label of the corresponding image sample
- S2 Construct a convolutional neural network model, divide all the image samples and the obtained image labels of the image samples into test set images and training set images, train the constructed convolutional neural network model by using the training set images, and test the convolutional neural network model by using the test set images
- S3 Input an image to be classified into the trained convolutional neural network model, extract a salient feature point of the image to be classified, and perform image classification according to the salient feature point

(57) Abstract: An image classification method. The method comprises: inputting a preset number of image samples, extracting a salient feature point in each of the image samples by using a preset type of feature point extraction algorithm, and respectively saving the salient feature point corresponding to each of the image samples as an image label of the corresponding image sample (S1); constructing a convolutional neural network model, dividing all the image samples and the obtained image labels of the image samples into test set images and training set images, training the constructed convolutional neural network model by using the training set images, and testing the convolutional neural network model by using the test set images (S2); and inputting an image to be classified into the trained convolutional neural network model, extracting a salient feature point of the image to be classified, and performing image classification according to the salient feature point (S3). Also provided are an image classification device and a computer readable storage medium. A convolutional neural network model can be trained to learn a scale invariant feature transform algorithm to extract a feature point of an image, so as to implement image classification.

SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,  
US, UZ, VC, VN, ZA, ZM, ZW。

- (84) 指定国 (除另有指明, 要求每一种可提供的地区  
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,  
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,  
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,  
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,  
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,  
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,  
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告 (条约第21条(3))。

(57) 摘要: 一种图像分类方法, 该方法包括: 输入预设数量的图像样本, 采用预设类型特征点提取算法分别提取各个所述图像样本中的显著特征点, 并分别将各个所述图像样本对应的显著特征点保存为对应的图像样本的图像标注 (S1); 构建卷积神经网络模型, 将所有所述图像样本和上述得到的图像样本的图像标注分为测试集图像和训练集图像, 利用所述训练集图像训练所述构建的卷积神经网络模型, 并利用所述测试集图像对所述卷积神经网络模型进行测试 (S2); 将待分类的图像输入到训练好的卷积神经网络模型中, 提取所述待分类的图像的显著特征点, 并根据所述显著特征点进行图像分类 (S3)。还提出一种图像分类装置以及一种计算机可读存储介质。可以训练卷积神经网络模型学习尺度不变特征变换算法提取图像特征点, 以实现图像的分类。

## 图像分类方法、装置及计算机可读存储介质

本申请要求于2019年04月04日提交中国专利局、申请号为201910269283.2、发明名称为“图像分类方法、装置及计算机可读存储介质”的中国专利申请的  
5 的优先权，其全部内容通过引用结合在申请中。

### 技术领域

本申请涉及人工智能技术领域，尤其涉及一种图像分类方法、装置及计算机可读存储介质。

### 背景技术

近几年来，随着卷积神经网络在图像处理领域的逐渐成熟，如何得到准确可靠的训练模型成为人们研究的热点。目前卷积神经网络的错误率从AlexNet的16.4%到现在广为使用的ResNet的3.6%，理论上错误率已经低于人  
15 眼错误率的5.1%。但是，在实际场景应用中，由于环境的复杂度增加，图像更容易受到光照，角度，像素等的影响，错误率会比理论上大大增加。

经典ResNet采用残差网络具备更少的参数，并为图像处理带来了更高的准确率和50层的卷积模型。然而进行模型训练时，目前的卷积神经网络仅考虑自身的分类损失，同时在图像特征提取过程中更容易受到视角变化、噪声  
20 的不良影响。因此，为了优化分类效果，需要CNN网络能够学习到更好的特征，然后用于分类。

### 发明内容

本申请提供一种图像分类方法、装置及计算机可读存储介质，其主要目的  
25 的在于提供一种优选的图像分类方案。

为实现上述目的，本申请提供的一种图像分类方法，包括：

输入预设数量的图像样本，采用预设类型特征点提取算法分别提取各个所述图像样本中的显著特征点，并分别将各个所述图像样本对应的显著特征  
点保存为对应的图像样本的图像标注；

30 构建卷积神经网络模型，将所有所述图像样本和上述得到的图像样本的图像标注分为测试集图像和训练集图像，利用所述训练集图像训练所述卷积神经网络模型，并利用所述测试集图像对所述卷积神经网络模型进行测试；

将待分类的图像输入到训练好的卷积神经网络模型中，提取所述待分类的图像的显著特征点，并根据所述显著特征点进行图像分类。

35 此外，为实现上述目的，本申请还提供一种图像分类装置，该装置包括存储器和处理器，所述存储器中存储有可在所述处理器上运行的图像分类程序，所述图像分类程序被所述处理器执行时实现如下步骤：

输入预设数量的图像样本，采用预设类型特征点提取算法分别提取各个所述图像样本中的显著特征点，并分别将各个所述图像样本对应的显著特征

点保存为对应的图像样本的图像标注；

构建卷积神经网络模型，将所有所述图像样本和上述得到的图像样本的图像标注分为测试集图像和训练集图像，利用所述训练集图像训练所述构建的卷积神经网络模型，并利用所述测试集图像对所述卷积神经网络模型进行测试；

将待分类的图像输入到训练好的卷积神经网络模型中，提取所述待分类的图像的显著特征点，并根据所述显著特征点进行图像分类。

此外，为实现上述目的，本申请还提供一种计算机可读存储介质，所述计算机可读存储介质上存储有图像分类程序，所述图像分类程序可被一个或者多个处理器执行，以实现如上所述的图像分类方法的步骤。

本申请提出的图像分类方法、装置及计算机可读存储介质输入预设数量的图像样本，采用预设类型特征点提取算法分别提取各个所述图像样本中的显著特征点，并分别将各个所述图像样本对应的显著特征点保存为对应的图像样本的图像标注；构建卷积神经网络模型，将所有所述图像样本和上述得到的图像样本的图像标注分为测试集图像和训练集图像，利用所述训练集图像训练所述构建的卷积神经网络模型，并利用所述测试集图像对所述卷积神经网络模型进行测试；将待分类的图像输入到训练好的卷积神经网络模型中，提取所述待分类的图像的显著特征点，并根据所述显著特征点进行图像分类。因此，可以训练卷积神经网络模型学习尺度不变特征变换算法提取图像特征点，并且根据对图像特征点的提取实现图像的分类。

## 附图说明

图1为本申请一实施例提供的图像分类方法的流程示意图；

图2为本申请一实施例提供的图像分类装置的内部结构示意图；

图3为本申请一实施例提供的图像分类装置中图像分类程序的模块示意图。

本申请目的的实现、功能特点及优点将结合实施例，参照附图做进一步说明。

## 具体实施方式

应当理解，此处所描述的具体实施例仅仅用以解释本申请，并不用于限定本申请。

本申请提供一种图像分类方法。参照图 1 所示，为本申请一实施例提供的图像分类方法的流程示意图。该方法可以由一个装置执行，该装置可以由软件和/或硬件实现。

在本实施例中，所述图像分类方法包括：

S1、输入预设数量的图像样本，采用预设类型特征点提取算法分别提取各个所述图像样本中的显著特征点，并分别将各个所述图像样本对应的显著特征点保存为对应的图像样本的图像标注。

本申请所述预设类型特征点提取算法为尺度不变特征变换 (SIFT, Scale Invariant Feature Transform) 算法。所述 SIFT 主要用来对图像进行关键点提取, 提取后的关键点可以用来做特征匹配等其他操作。

详细的, 所述步骤 S1 包括:

5 逐一选择所述图像样本或者批量选择所述图像样本, 在一个图像样本被选择后, 针对该图像样本执行如下步骤:

1.1、利用该图像样本制作高斯金字塔: 将该图像样本进行下采样得到  $N$  个不同尺寸的图像, 再对每张图像使用  $k$  组不同大小的高斯核进行高斯模糊, 得到  $N * K$  个不同尺寸和模糊程度的图像, 其中,  $N$  和  $K$  均为正整数;

10 1.2、将高斯金字塔转化为 DoG 金字塔, 即差分金字塔: 在高斯金字塔中, 每层的  $K$  张照片按照顺序两两相减, 得到差分图像, 最后得到  $N * (K-1)$  张差分图像, 组成 DoG 金字塔。

1.3、对所述差分金字塔进行多尺度空间极值点检测, 得到该图像样本的候选特征点。

15 尺度空间的目的是检测在尺度变化时, 图像中仍然稳定的特征。

原始图像样本  $I(x, y)$  的尺度空间定义为:  $L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$ , 其中  $L(x, y, \sigma)$  代表该图像样本的尺度空间,  $G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$ ,  $\sigma$  为像素的尺度参数, 当  $\sigma$  连续变化,  $G(x, y, \sigma)$  构成可变尺度的 2 维高斯函数。

20 本申请在进行多尺度空间极值点检测时, 遍历所述 DoG 金字塔中每一张图像的像素, 比较当前像素点周围的 8 个像素, 以及上下两张图像的 18 个像素, 共 26 个像素, 以确保在尺度空间和二维图像空间都检测到极值点。一个点如果在尺度空间本层以及上下两层的 26 个邻域中是最大或最小值时, 就认为该点是图像在该尺度下的其中一个候选特征点。

25 DoG 金字塔的每组图像中的第一层和最后一层是无法进行比较取得极值的。为了满足尺度变换的连续性, 在所述 DoG 金字塔的每组图像的顶层使用高斯模糊生成 3 幅图像, 高斯金字塔每组有  $S+3$  层图像, DoG 金字塔的每组有  $S+2$  组图像。

30 取  $S=3$ , 也就是每组有 3 层, 则  $k = 2^{\frac{1}{S}} = 2^{\frac{1}{3}}$ , 即高斯金字塔每组有  $(S-1)*3$  层图像, DoG 金字塔每组有  $(S-2) * 2$  层图像。在 DoG 金字塔的第一组有两层尺度分别是  $\sigma, k\sigma$ , 第二组有两层的尺度分别是  $2\sigma, 2k\sigma$ , 由于只有两项是无法比较取得极值的 (只有左右两边都有值才能有极值)。无法比较取得极值, 需要继续对每组的图像进行高斯模糊, 使得尺度形成  $\sigma, k\sigma, k^2\sigma, k^3\sigma, k^4\sigma$ , 这样就可以选择中间的三项  $k\sigma, k^2\sigma, k^3\sigma$ 。对应的下一组由上一组降采样得到的三项是  $2k\sigma, 2k^2\sigma, 2k^3\sigma$ , 其首项  $2k\sigma = 2 \cdot 2^{\frac{1}{3}}\sigma = 2 \cdot 2^{\frac{4}{3}}\sigma$ , 刚好与上一组的最后一项  $k^3\sigma = 2^{\frac{3}{3}}\sigma$  的尺度连续起来。

1.4 对得到的该图像样本的候选特征点进行定位, 并去掉低对比度的候选特征点和不稳定的边缘候选特征点, 得到该图像样本的显著特征点。

上述得到的图像样本的候选特征点，需要进行定位，以确定所述候选特征点的位置。本申请对尺度空间 $G(x, y, \sigma)$ 进行二次泰勒展开： $D(x) = D + \frac{\partial D^T}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial x^2} x$ ，然后进行求导，令 $D(x) = 0$ ，得到该图像样本的候选特征点的位置： $\hat{x} = -\frac{\partial^2 D^{-1}}{\partial x^2} \frac{\partial D}{\partial x}$ 。

进一步地，本申请对所述候选特征点进行筛选，去掉低对比度的候选特征点和不稳定的边缘候选特征点，方法如下：

首先，本申请将所述候选特征点带入到所述 $D(x)$ 函数中，利用二次泰勒展开式进行最小二乘拟合，得到 $D(\hat{x}) = D + \frac{\partial D^T}{\partial x} \hat{x}$ 。设对比度的预设阈值为 $T=0.03$ ，如果 $D(\hat{x}) \geq 0.03$ ，则保留该候选特征点，否则作为低对比度候选特征点丢弃。

进一步地，本申请计算曲面在该候选特征点的主曲率，进而去除不稳定的边缘候选特征点。候选特征点的 $D(x)$ 函数的主曲率与 $2 \times 2H$ 的特征值成正比： $H$ 为候选特征点邻域对应位置的差分矩阵 $H = \begin{bmatrix} D_{xx} & D_{yx} \\ D_{xy} & D_{yy} \end{bmatrix}$ 。其中 $D_{xx}$ 、 $D_{xy}$ 、 $D_{yx}$ 、 $D_{yy}$ 是候选特征点邻域对应位置的差分求得的。为了避免求具体的值，可以使用 $H$ 特征值得比例。设 $\alpha = \lambda_{max}$ 为 $H$ 的最大特征值， $\beta = \lambda_{min}$ 为 $H$ 的最小特征值，则： $Th(H) = D_{xx} + D_{yy} = \alpha + \beta$ ， $Det(H) = D_{xx} D_{yy} - D_{xy}^2 = \alpha \cdot \beta$ 。其中， $Th(H)$ 为矩阵 $H$ 的迹， $Det(H)$ 为矩阵 $H$ 的行列式。设 $\gamma = \frac{\alpha}{\beta}$ 表示最大特征值和最小特征值的比值，则 $\frac{Th(H)^2}{Det(H)} = \frac{(\alpha + \beta)^2}{\alpha \beta} = \frac{(\gamma \beta + \beta)^2}{\gamma \beta^2} = \frac{(\gamma + 1)^2}{\gamma}$ 。上式的结果与两个特征值的比例有关，和具体的大小无关，当两个特征值相等时其值最小，并且随着 $\gamma$ 的增大而增大。因此为了检测主曲率是否在某个阈值 $T_\gamma$ 下，只需检测 $\frac{Th(H)^2}{Det(H)} > \frac{(T_\gamma + 1)^2}{T_\gamma}$ ，如果上式成立，则剔除该候选特征点，否则保留该候选特征点。

1.5 计算所述显著特征点的主方向。

本申请利用关键点邻域像素的梯度方向分布特性为每个显著特征点指定方向参数，使算子具备旋转不变性。

每个点 $L(x, y)$ 处梯度的模值 $m(x, y)$ 和方向 $\theta(x, y)$ 公式如下：

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2}$$

$$\theta(x, y) = \alpha \tan 2((L(x+1, y) - L(x-1, y)) / (L(x, y+1) - L(x, y-1)))$$

其中， $L$ 为每个显著特征点所在的尺度， $\alpha = \lambda_{max}$ 为 $H$ 的最大特征值， $H$ 为候选特征点邻域对应位置的差分矩阵。

至此，图像的显著特征点已经提取完毕，每个显著特征点有三个信息：位置、尺度、方向，由此可以确定一个SIFT特征区域。

1.6 构造所述显著特征点的描述子，并根据所述显著特征点的描述子，将

所述显著特征点保存为该图像样本的图像标注。

所述显著特征点在确定了位置、尺度和方向之后，下一步就是根据所述显著特征点周围的局部特征构造所述显著特征点的描述子，其生成过程如下：

①旋转所述显著特征点的坐标，使坐标轴旋转为所述显著特征点的主方向上，确保旋转不变性。

②在旋转后的图像上，以所述显著特征点为中心取  $8 \times 8$  大小的窗口，设所述显著特征点位于该窗口的中央位置，所述显著特征点周围的方格代表所述显著特征点的邻域范围，每个小方格代表一个像素，在每个像素上计算梯度，箭头的方向和长度分别表示像素的方向和模值。

③将所述窗口平均分为  $4 \times 4$  的子窗口，再利用高斯模糊方法增加与所述显著特征点邻近域的权重值、降低所述显著特征点较远的权重值，最后计算每个区域的  $0^\circ$ 、 $45^\circ$ 、 $90^\circ$ 、 $135^\circ$ 、 $180^\circ$ 、 $225^\circ$ 、 $270^\circ$ 、 $315^\circ$  方向的梯度累加值，以得到所述显著特征点的描述子的特征向量。这时的 SIFT 特征向量已经取出来对旋转和尺度变化带来的影响。最后进行灰度值的归一化，消除光线的影响。

根据上述算法提取到图像样本的显著特征点构造的描述子，将所述显著特征点保存为该图像样本的图像标注。

S2、构建卷积神经网络模型，将所有所述图像样本和上述得到的图像样本的图像标注分为测试集图像和训练集图像，利用所述训练集图像训练所述构建的卷积神经网络模型，并利用所述测试集图像对所述卷积神经网络模型进行测试。

本申请较佳实施例中，所述预设模型为 ResNet50 模型。

本申请较佳实施例将图像样本和图像对应的图像标注分为测试集和训练集，其中，图像样本的 70% 作为训练集图像，30% 作为测试集图像。

本申请较佳实施例中，利用所述图像样本和上述得到的图像样本的图像标注训练所述构建的卷积神经网络模型的步骤包括：

2.1 对训练集图像进行批量标准化。

网络训练过程中 CNN 内部节点的分布变化称为内部协变量的转移，通常认为它是影响网络训练速度的重要原因。为减少内部协变量的转移，提出了批量标准化 (Batch Normalization, BN)，它通过在每个非线性变换引入标准化和移位步骤，有效地解决了内部协变量转移的问题，从而大大加快了网络的收敛速度。本申请较佳实施例利用下述公式对训练集图像进行批量标准化：

$$y = f(BN(W_x))$$

其中， $x$  是输入； $y$  是输出； $W$  是卷积矩阵； $f$  是 ReLU 函数。

2.2 将所述训练集图像输入至由输入层、卷积层 (Convolution, Conv)、激活函数和输出层构成的残差网络结构的 CNN 模型。

所述训练集图像经过输入层后，由卷积层提取其特征图，再经过激活函数进行非线性变换，然后从输出层输出，并进入下一个神经单元。本申请所

述激活函数包括 tanh 函数和 sigmoid 函数，如下：

$$f(x) = \tanh(x)$$

$$f(x) = (1 + e^{-x})^{-1}$$

其中，由输入层、卷积层、激活函数和输出层组成的 CNN 模型可表达为：

$$y = f(Wx + b)$$

其中， $x$  是输入； $y$  是输出； $W$  是卷积矩阵； $b$  是卷积偏置； $f$  是 ReLU 函数。

2.3 利用图像集合，构建标准 50 层深度残差网络 (ResNet50) 网络结构。

深度残差网络(ResNet)出现其实就是为了解决网络深度变深以后的性能退化问题。因此，在本案例中采用 ResNet50 作为 CNN 模型，不仅是具有更深层的结构，也增强了模型非线性能力，能够更好的拟合数据。标准的 ResNet50 中网络分成 5 部分，分别是：conv1，conv2\_x，conv3\_x，conv4\_x，conv5\_x，首先有个输入 7x7x64 的卷积，然后经过 3+4+6+3=16 个 building block，每个 block 为 3 层，所以有 16x3=48 层，最后有个 fc 层(用于分类)，所以 1+48+1=50 层。

在标准的 ResNet 中，Conv5\_x 会输出分类，然而 CNN 中，输入图像通过和三个可训练的滤波器和可加偏置进行卷积，产生三个特征映射图，然后特征映射图中每组的四个像素再进行求和，加权值，加偏置，通过一个 Sigmoid 函数得到三个特征映射图。这些映射图再进过滤波，最终，这些像素值被光栅化，并连接成一个向量输入到传统的神经网络，得到输出。本案例中将特征映射作为显著特征点输出。

计算显著特征点损失 (Loss) 与分类损失 (Loss)，并将其用于进行反向传播计算误差值，更新权重训练 CNN 模型。

所述 Loss 的计算公式： $\tilde{\ell}(y, z) = -\log\left(\frac{e^{zy}}{\sum_{j=1}^m e^{z_j}}\right) = \log(\sum_{j=1}^m e^{z_j}) - zy$ 。  $y$  为图像所属于的类别编号，是真实的标签值； $z$  为图像所对应的数据，这个数据也是输入 softmax 层的数据，这个数据是 ResNet 输出的数据，即特征映射和分类器传出。根据 loss 调整各层参数，不断迭代直至收敛。最终得到训练好的 ResNet50 模型。

S3、将待分类的图像输入到训练好的卷积神经网络模型中，提取所述待分类的图像的显著特征点，并根据所述显著特征点进行图像分类。

本申请还提供一种图像分类装置。参照图 2 所示，为本申请一实施例提供的图像分类装置的内部结构示意图。

在本实施例中，图像分类装置 1 可以是 PC (Personal Computer，个人电脑)，也可以是智能手机、平板电脑、便携计算机等终端设备。该图像分类装置 1 至少包括存储器 11、处理器 12，通信总线 13，以及网络接口 14。

其中，存储器 11 至少包括一种类型的可读存储介质，所述可读存储介质包括闪存、硬盘、多媒体卡、卡型存储器 (例如，SD 或 DX 存储器等)、磁

性存储器、磁盘、光盘等。存储器 11 在一些实施例中可以是图像分类装置 1 的内部存储单元，例如该图像分类装置 1 的硬盘。存储器 11 在另一些实施例中也可以是图像分类装置 1 的外部存储设备，例如图像分类装置 1 上配备的插接式硬盘，智能存储卡（Smart Media Card, SMC），安全数字（Secure Digital, SD）卡，闪存卡（Flash Card）等。进一步地，存储器 11 还可以既包括图像分类装置 1 的内部存储单元也包括外部存储设备。存储器 11 不仅可以用于存储安装于图像分类装置 1 的应用软件及各类数据，例如图像分类程序 01 的代码等，还可以用于暂时地存储已经输出或者将要输出的数据。

处理器 12 在一些实施例中可以是一中央处理器（Central Processing Unit, CPU）、控制器、微控制器、微处理器或其他数据处理芯片，用于运行存储器 11 中存储的程序代码或处理数据，例如执行图像分类程序 01 等。

通信总线 13 用于实现这些组件之间的连接通信。

网络接口 14 可选的可以包括标准的有线接口、无线接口（如 WI-FI 接口），通常用于在该装置 1 与其他电子设备之间建立通信连接。

可选地，该装置 1 还可以包括用户接口，用户接口可以包括显示器（Display）输入单元比如键盘（Keyboard），可选的用户接口还可以包括标准的有线接口、无线接口。可选地，在一些实施例中，显示器可以是 LED 显示器、液晶显示器、触控式液晶显示器以及 OLED（Organic Light-Emitting Diode，有机发光二极管）触摸器等。其中，显示器也可以适当的称为显示屏或显示单元，用于显示在图像分类装置 1 中处理的信息以及用于显示可视化的用户界面。

图 2 仅示出了具有组件 11-14 以及图像分类程序 01 的图像分类装置 1，本领域技术人员可以理解的是，图 1 示出的结构并不构成对图像分类装置 1 的限定，可以包括比图示更少或者更多的部件，或者组合某些部件，或者不同的部件布置。

在图 2 所示的装置 1 实施例中，存储器 11 中存储有图像分类程序 01；处理器 12 执行存储器 11 中存储的图像分类程序 01 时实现如下步骤：

步骤一、输入预设数量的图像样本，采用预设类型特征点提取算法分别提取各个所述图像样本中的显著特征点，并分别将各个所述图像样本对应的显著特征点保存为对应的图像样本的图像标注。

本申请所述预设类型特征点提取算法为尺度不变特征变换（SIFT，Scale Invariant Feature Transform）算法。所述 SIFT 主要用来对图像进行关键点提取，提取后的关键点可以用来做特征匹配等其他操作。

详细的，所述步骤 S1 包括：

逐一选择所述图像样本或者批量选择所述图像样本，在一个图像样本被选择后，针对该图像样本执行如下步骤：

1.1、利用该图像样本制作高斯金字塔：将该图像样本进行下采样得到 N 个不同尺寸的图像，再对每张图像使用 k 组不同大小的高斯核进行高斯模糊，得到 N\*K 个不同尺寸和模糊程度的图像，其中，N 和 K 均为正整数；

1.2、将高斯金字塔转化为 DoG 金字塔,即差分金字塔:在高斯金字塔中,每层的 K 张照片按照顺序两两相减,得到差分图像,最后得到  $N * (K-1)$  张差分图像,组成 DoG 金字塔。

1.3、对所述差分金字塔进行多尺度空间极值点检测,得到该图像样本的候选特征点。

尺度空间的目的是检测在尺度变化时,图像中仍然稳定的特征。

原始图像样本  $I(x,y)$  的尺度空间定义为:  $L(x,y,\sigma) = G(x,y,\sigma) * I(x,y)$ , 其中  $L(x,y,\sigma)$  代表该图像样本的尺度空间,  $G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$ ,  $\sigma$  为像素的尺度参数,当  $\sigma$  连续变化,  $G(x,y,\sigma)$  构成可变尺度的 2 维高斯函数。

本申请在进行多尺度空间极值点检测时,遍历所述 DoG 金字塔中每一张图像的像素,比较当前像素点周围的 8 个像素,以及上下两张图像的 18 个像素,共 26 个像素,以确保在尺度空间和二维图像空间都检测到极值点。一个点如果在尺度空间本层以及上下两层的 26 个邻域中是最大或最小值时,就认为该点是图像在该尺度下的其中一个候选特征点。

DoG 金字塔的每组图像中的第一层和最后一层是无法进行比较取得极值的。为了满足尺度变换的连续性,在所述 DoG 金字塔的每组图像的顶层使用高斯模糊生成 3 幅图像,高斯金字塔每组有  $S+3$  层图像,DoG 金字塔的每组有  $S+2$  组图像。

取  $S=3$ ,也就是每组有 3 层,则  $k = 2^{\frac{1}{S}} = 2^{\frac{1}{3}}$ ,即高斯金字塔每组有  $(S-1)*3$  层图像,DoG 金字塔每组有  $(S-2) * 2$  层图像。在 DoG 金字塔的第一组有两层尺度分别是  $\sigma, k\sigma$ ,第二组有两层的尺度分别是  $2\sigma, 2k\sigma$ ,由于只有两项是无法比较取得极值的(只有左右两边都有值才能有极值)。无法比较取得极值,需要继续对每组的图像进行高斯模糊,使得尺度形成  $\sigma, k\sigma, k^2\sigma, k^3\sigma, k^4\sigma$ ,这样就可以选择中间的三项  $k\sigma, k^2\sigma, k^3\sigma$ 。对应的下一组由上一组降采样得到的三项是  $2k\sigma, 2k^2\sigma, 2k^3\sigma$ ,其首项  $2k\sigma = 2 \cdot 2^{\frac{1}{3}}\sigma = 2 \cdot 2^{\frac{4}{3}}\sigma$ ,刚好与上一组的最后一项  $k^3\sigma = 2^{\frac{3}{3}}\sigma$  的尺度连续起来。

1.4 对得到的该图像样本的候选特征点进行定位,并去掉低对比度的候选特征点和不稳定的边缘候选特征点,得到该图像样本的显著特征点。

上述得到的图像样本的候选特征点,需要进行定位,以确定所述候选特征点的位置。本申请对尺度空间  $G(x,y,\sigma)$  进行二次泰勒展开:  $D(x) = D + \frac{\partial D^T}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial x^2} x$ , 然后进行求导,令  $D(x) = 0$ , 得到该图像样本的候选特征点的位置:  $\hat{x} = -\frac{\partial^2 D^{-1}}{\partial x^2} \frac{\partial D}{\partial x}$ 。

进一步地,本申请对所述候选特征点进行筛选,去掉低对比度的候选特征点和不稳定的边缘候选特征点,方法如下:

首先,本申请将所述候选特征点带入到所述  $D(x)$  函数中,利用二次泰勒展开式进行最小二乘拟合,得到  $D(\hat{x}) = D + \frac{\partial D^T}{\partial x} \hat{x}$ 。设对比度的预设阈值为

$T=0.03$ ，如果  $D(\hat{x}) \geq 0.03$ ，则保留该候选特征点，否则作为低对比度候选特征点丢弃。

进一步地，本申请计算曲面在该候选特征点的主曲率，进而去除不稳定的边缘候选特征点。候选特征点的  $D(x)$  函数的主曲率与  $2 \times 2H$  的特征值成正比

5 比  $H$  为候选特征点邻域对应位置的差分矩阵  $H = \begin{bmatrix} D_{xx} & D_{yx} \\ D_{xy} & D_{yy} \end{bmatrix}$ 。其中  $D_{xx}$   $D_{xy}$ ，

$D_{yx}$ ， $D_{yy}$  是候选特征点邻域对应位置的差分求得的。为了避免求具体的值，可以使用  $H$  特征值得比例。设  $\alpha = \lambda_{max}$  为  $H$  的最大特征值， $\beta = \lambda_{min}$  为  $H$  的最小特征值，则： $Th(H) = D_{xx} + D_{yy} = \alpha + \beta$ ， $Det(H) = D_{xx} + D_{yy} - D_{xy}^2 = \alpha \cdot \beta$ 。其中， $Th(H)$  为矩阵  $H$  的迹， $Det(H)$  为矩阵  $H$  的行列式。设  $\gamma = \frac{\alpha}{\beta}$  表示

10 最大特征值和最小特征值的比值，则  $\frac{Th(H)^2}{Det(H)} = \frac{(\alpha + \beta)^2}{\alpha \beta} = \frac{(\gamma \beta + \beta)^2}{\gamma \beta^2} = \frac{(\gamma + 1)^2}{\gamma}$ 。

上式的结果与两个特征值的比例有关，和具体的大小无关，当两个特征值相等时其值最小，并且随着  $\gamma$  的增大而增大。因此为了检测主曲率是否在某个阈值  $T_\gamma$  下，只需检测  $\frac{Th(H)^2}{Det(H)} > \frac{(T_\gamma + 1)^2}{T_\gamma}$ ，如果上式成立，则剔除该候选特征点，否则保留该候选特征点。

15 1.5 计算所述显著特征点的主方向。

本申请利用关键点邻域像素的梯度方向分布特性为每个显著特征点指定方向参数，使算子具备旋转不变性。

每个点  $L(x, y)$  处梯度的模值  $m(x, y)$  和方向  $\theta(x, y)$  公式如下：

$$m(x, y) = \sqrt{(L(x + 1, y) - L(x - 1, y))^2 + (L(x, y + 1) - L(x, y - 1))^2}$$

$$\theta(x, y) = \alpha \tan 2((L(x + 1, y) - L(x - 1, y)) / (L(x, y + 1) - L(x, y - 1)))$$

20 其中， $L$  为每个显著特征点所在的尺度， $\alpha = \lambda_{max}$  为  $H$  的最大特征值， $H$  为候选特征点邻域对应位置的差分矩阵。

至此，图像的显著特征点已经提取完毕，每个显著特征点有三个信息：位置，尺度、方向，由此可以确定一个 SIFT 特征区域。

25 1.6 构造所述显著特征点的描述子，并根据所述显著特征点的描述子，将所述显著特征点保存为该图像样本的图像标注。

所述显著特征点在确定了位置、尺度和方向之后，下一步就是根据所述显著特征点周围的局部特征构造所述显著特征点的描述子，其生成过程如下：

① 旋转所述显著特征点的坐标，使坐标轴旋转为所述显著特征点的主方向上，确保旋转不变性。

30 ② 在旋转后的图像上，以所述显著特征点为中心取  $8 \times 8$  大小的窗口，设所述显著特征点位于该窗口的中央位置，所述显著特征点周围的方格代表所述显著特征点的邻域范围，每个小方格代表一个像素，在每个像素上计算梯度，箭头的方向和长度分别表示像素的方向和模值。

③ 将所述窗口平均分为  $4 \times 4$  的子窗口，再利用高斯模糊方法增加与所述

显著特征点邻近域的权重值、降低所述显著特征点较远的权重值，最后计算每个区域的  $0^\circ$ 、 $45^\circ$ 、 $90^\circ$ 、 $135^\circ$ 、 $180^\circ$ 、 $225^\circ$ 、 $270^\circ$ 、 $315^\circ$  方向的梯度累加值，以得到所述显著特征点的描述子的特征向量。这时的 SIFT 特征向量已经取出来对旋转和尺度变化带来的影响。最后进行灰度值的归一化，消除光线的影响。

根据上述算法提取到图像样本的显著特征点构造的描述子，将所述显著特征点保存为该图像样本的图像标注。

步骤二、构建卷积神经网络模型，将所有所述图像样本和上述得到的图像样本的图像标注分为测试集图像和训练集图像，利用所述训练集图像训练所述构建的卷积神经网络模型，并利用所述测试集图像对所述卷积神经网络模型进行测试。

本申请较佳实施例中，所述预设模型为 ResNet50 模型。

本申请较佳实施例将图像样本和图像对应的图像标注分为测试集和训练集，其中，图像样本的 70% 作为训练集图像，30% 作为测试集图像。

本申请较佳实施例中，利用所述图像样本和上述得到的图像样本的图像标注训练所述构建的卷积神经网络模型的步骤包括：

2.1 对训练集图像进行批量标准化。

网络训练过程中 CNN 内部节点的分布变化称为内部协变量的转移，通常认为它是影响网络训练速度的重要原因。为减少内部协变量的转移，提出了批量标准化 (Batch Normalization, BN)，它通过在每个非线性变换引入标准化和移位步骤，有效地解决了内部协变量转移的问题，从而大大加快了网络的收敛速度。本申请较佳实施例利用下述公式对训练集图像进行批量标准化：

$$y = f(BN(W_x))$$

其中， $x$  是输入； $y$  是输出； $W$  是卷积矩阵； $f$  是 ReLU 函数。

2.2 将所述训练集图像输入至由输入层、卷积层 (Convolution, Conv) 激活函数和输出层构成的残差网络结构的 CNN 模型。

所述训练集图像经过输入层后，由卷积层提取其特征图，再经过激活函数进行非线性变换，然后从输出层输出，并进入下一个神经单元。本申请所述激活函数包括 tanh 函数和 sigmoid 函数，如下：

$$f(x) = \tanh(x)$$

$$f(x) = (1 + e^{-x})^{-1}$$

其中，由输入层、卷积层、激活函数和输出层组成的 CNN 模型可表达为：

$$y = f(Wx + b)$$

其中， $x$  是输入； $y$  是输出； $W$  是卷积矩阵； $b$  是卷积偏置； $f$  是 ReLU 函数。

2.3 利用图像集合，构建标准 50 层深度残差网络 (ResNet50) 网络结构。

深度残差网络 (ResNet) 出现其实就是为了解决网络深度变深以后的性能退化问题。因此，在本案例中采用 ResNet50 作为 CNN 模型，不仅是具有更

深层的结构，也增强了模型非线性能力，能够更好的拟合数据。标准的 ResNet50 中网络分成 5 部分，分别是：conv1，conv2\_x，conv3\_x，conv4\_x，conv5\_x，首先有个输入 7x7x64 的卷积，然后经过 3+4+6+3=16 个 building block，每个 block 为 3 层，所以有 16x3=48 层，最后有个 fc 层(用于分类)，所以 1+48+1=50 层。

在标准的 ResNet 中，Conv5\_x 会输出分类，然而 CNN 中，输入图像通过和三个可训练的滤波器和可加偏置进行卷积，产生三个特征映射图，然后特征映射图中每组的四个像素再进行求和，加权值，加偏置，通过一个 Sigmoid 函数得到三个特征映射图。这些映射图再进过滤波，最终，这些像素值被光栅化，并连接成一个向量输入到传统的神经网络，得到输出。本案例中将特征映射作为显著特征点输出。

计算显著特征点损失 (Loss) 与分类损失 (Loss)，并将其用于进行反向传播计算误差值，更新权重训练 CNN 模型。

所述 Loss 的计算公式： $\tilde{\ell}(y, z) = -\log\left(\frac{e^{zy}}{\sum_{j=1}^m e^{z_j}}\right) = \log(\sum_{j=1}^m e^{z_j}) - z_y$ 。y

为图像所属于的类别编号，是真实的标签值；z 为图像所对应的数据，这个数据也是输入 softmax 层的数据，这个数据是 ResNet 输出的数据，即特征映射和分类器传出。根据 loss 调整各层参数，不断迭代直至收敛。最终得到训练好的 ResNet50 模型。

步骤三、将待分类的图像输入到训练好的卷积神经网络模型中，提取所述待分类的图像的显著特征点，并根据所述显著特征点进行图像分类。

可选地，在其他实施例中，图像分类程序还可以被分割为一个或者多个模块，一个或者多个模块被存储于存储器 11 中，并由一个或多个处理器（本实施例为处理器 12）所执行以完成本申请，本申请所称的模块是指能够完成特定功能的一系列计算机程序指令段，用于描述图像分类程序在图像分类装置中的执行过程。

例如，参照图 3 所示，为本申请图像分类装置一实施例中的图像分类程序的程序模块示意图，该实施例中，图像分类程序可以被分割为特征点提取模块 10、模型训练模块 20、分类模块 30，示例性地：

特征点提取模块 10 用于：输入预设数量的图像样本，采用预设类型特征点提取算法分别提取各个所述图像样本中的显著特征点，并分别将各个所述图像样本对应的显著特征点保存为对应的图像样本的图像标注。

可选地，所述采用预设类型特征点提取算法分别提取各个所述图像样本中的显著特征点，并分别将各个所述图像样本对应的显著特征点保存为对应的图像样本的图像标注，包括：

逐一选择所述图像样本或者批量选择所述图像样本，在一个图像样本被选择后，针对该图像样本执行如下步骤：

利用该图像样本制作高斯金字塔：将该图像样本进行下采样得到 N 个不同尺寸的图像，再对每张图像使用 k 组不同大小的高斯核进行高斯模糊，得

到  $N * K$  个不同尺寸和模糊程度的图像，其中， $N$  和  $K$  均为正整数；

将高斯金字塔转化为差分金字塔：在高斯金字塔中，每层的  $K$  张照片按照顺序两两相减，得到差分图像，最后得到  $N * (K-1)$  张差分图像，组成所述差分金字塔；

5 对所述差分金字塔进行多尺度空间极值点检测，得到该图像样本的候选特征点；

对得到的该图像样本的候选特征点进行定位，并去掉候选特征点中的低对比度的特征点和不稳定的边缘特征点，得到该图像样本的显著特征点；

10 计算所述显著特征点的主方向，及构造所述显著特征点的描述子，并根据所述显著特征点的描述子，将所述显著特征点保存为该图像样本的图像标注。

可选地，所述对得到的该图像样本的候选特征点进行定位，并去掉其中的低对比度的特征点和不稳定的边缘特征点，得到该图像样本的显著特征点，包括：

15 将该图像样本  $I(x,y)$  的尺度空间定义为： $L(x,y,\sigma) = G(x,y,\sigma) * I(x,y)$ ，其中  $L(x,y,\sigma)$  代表该图像样本的尺度空间， $G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$ ， $\sigma$  为像素的尺度参数，当  $\sigma$  连续变化， $G(x,y,\sigma)$  构成可变尺度空间的 2 维高斯函数；

对尺度空间  $G(x,y,\sigma)$  进行二次泰勒展开： $D(x) = D + \frac{\partial D^T}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial x^2} x$ ，并进行求导，令  $D(x) = 0$ ，得到该图像样本的候选特征点的位置：

$$20 \quad \hat{x} = -\frac{\partial^2 D^{-1}}{\partial x^2} \frac{\partial D}{\partial x} ;$$

将所述候选特征点带入到所述  $D(x)$  函数中，利用二次泰勒展开式进行最小二乘拟合，得到  $D(\hat{x}) = D + \frac{\partial D^T}{\partial x} \hat{x}$ ；

如果  $D(\hat{x}) \geq$  预设阈值，则保留该候选特征点，否则将该候选特征点作为低对比度特征点丢弃；

25 将候选特征点的  $D(x)$  函数的主曲率与  $2 \times 2$   $H$  的特征值成正比： $H$  为候选特征点邻域对应位置的差分矩阵， $H = \begin{bmatrix} D_{xx} & D_{yx} \\ D_{xy} & D_{yy} \end{bmatrix}$ ，其中， $D_{xx}$ ， $D_{xy}$ ， $D_{yx}$ ， $D_{yy}$  是候选特征点邻域对应位置的差分求得；

设  $\alpha = \lambda_{max}$  为  $H$  的最大特征值， $\beta = \lambda_{min}$  为  $H$  的最小特征值，则： $Th(H) = D_{xx} + D_{yy} = \alpha + \beta$ ， $Det(H) = D_{xx} + D_{yy} - D_{xy}^2 = \alpha \cdot \beta$ ，其中， $Th(H)$  为矩阵  $H$  的迹， $Det(H)$  为矩阵  $H$  的行列式；

30 设  $\gamma = \frac{\alpha}{\beta}$  表示最大特征值和最小特征值的比值，则  $\frac{Th(H)^2}{Det(H)} = \frac{(\alpha + \beta)^2}{\alpha \beta} = \frac{(\gamma \beta + \beta)^2}{\gamma \beta^2} = \frac{(\gamma + 1)^2}{\gamma}$ ；

当  $\frac{Th(H)^2}{Det(H)} > \frac{(T_\gamma + 1)^2}{T_\gamma}$  时，剔除该候选特征点，否则保留该候选特征点，其中， $T_\gamma$  为预设阈值。

可选地，所述显著特征点的主方向为：

$$\theta(x, y) = \alpha \tan 2((L(x+1, y) - L(x-1, y)) / (L(x, y+1) - L(x, y-1)))$$

其中，L 为每个特征点所在的尺度， $\alpha = \lambda_{max}$  为 H 的最大特征值，H 为候选特征点邻域对应位置的差分矩阵。

可选地，所述构造所述显著特征点的描述子的步骤包括：

5 旋转所述显著特征点的坐标，使坐标轴旋转为所述显著特征点的主方向上，确保旋转不变性；

在旋转后的图像上，以所述显著特征点为中心取 8\*8 大小的窗口，设所述显著特征点位于该窗口的中央位置，所述显著特征点周围的方格代表所述显著特征点的邻域范围，每个小方格代表一个像素，在每个像素上计算梯度，  
10 箭头的方向和长度分别表示像素的方向和模值；

将所述窗口平均分为 4\*4 的子窗口，再利用高斯模糊方法增加与所述显著特征点邻近域的权重值、降低所述显著特征点较远的权重值，最后计算每个区域的 0°、45°、90°、135°、180°、225°、270°、315°方向的梯度累加值，以得到所述显著特征点的描述子的特征向量。

15 模型训练模块 20 用于：构建卷积神经网络模型，将所有所述图像样本和上述得到的图像样本的图像标注分为测试集图像和训练集图像，利用所述训练集图像训练所述构建的卷积神经网络模型，并利用所述测试集图像对所述卷积神经网络模型进行测试。

分类模块 30 用于：将待分类的图像输入到训练好的卷积神经网络模型中，  
20 提取所述待分类的图像的显著特征点，并根据所述显著特征点进行图像分类。

上述特征点提取模块 10、模型训练模块 20、分类模块 30 等程序模块被执行时所实现的功能或操作步骤与上述实施例大体相同，在此不再赘述。

此外，本申请实施例还提出一种计算机可读存储介质，所述计算机可读存储介质上存储有图像分类程序，所述图像分类程序可被一个或多个处理器  
25 执行，以实现如下操作：

输入预设数量的图像样本，采用预设类型特征点提取算法分别提取各个所述图像样本中的显著特征点，并分别将各个所述图像样本对应的显著特征点保存为对应的图像样本的图像标注；

构建卷积神经网络模型，将所有所述图像样本和上述得到的图像样本的  
30 图像标注分为测试集图像和训练集图像，利用所述训练集图像训练所述构建的卷积神经网络模型，并利用所述测试集图像对所述卷积神经网络模型进行测试；

将待分类的图像输入到训练好的卷积神经网络模型中，提取所述待分类的图像的显著特征点，并根据所述显著特征点进行图像分类。

35 本申请计算机可读存储介质具体实施方式与上述图像分类装置和方法各实施例基本相同，在此不作累述。

需要说明的是，上述本申请实施例序号仅仅为了描述，不代表实施例的优劣。并且本文中的术语“包括”、“包含”或者其任何其他变体意在涵盖非排他

性的包含，从而使得包括一系列要素的过程、装置、物品或者方法不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种过程、装置、物品或者方法所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括该要素的过程、装置、物品或者方法中还存在另外的相同要素。

通过以上的实施方式的描述，本领域的技术人员可以清楚地了解到上述实施例方法可借助软件加必需的通用硬件平台的方式来实现，当然也可以通过硬件，但很多情况下前者是更佳的实施方式。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来，该计算机软件产品存储在如上所述的一个存储介质(如 ROM/RAM、磁碟、光盘)中，包括若干指令用以使得一台终端设备(可以是手机，计算机，服务器，或者网络设备等)执行本申请各个实施例所述的方法。

以上仅为本申请的优选实施例，并非因此限制本申请的专利范围，凡是利用本申请说明书及附图内容所作的等效结构或等效流程变换，或直接或间接运用在其他相关的技术领域，均同理包括在本申请的专利保护范围内。

# 权 利 要 求 书

1、一种图像分类方法，其特征在于，所述方法包括：

输入预设数量的图像样本，采用预设类型特征点提取算法分别提取各个所述图像样本中的显著特征点，并分别将各个所述图像样本对应的显著特征点保存为对应的图像样本的图像标注；

构建卷积神经网络模型，将所有所述图像样本和上述得到的图像样本的图像标注分为测试集图像和训练集图像，利用所述训练集图像训练所述卷积神经网络模型，并利用所述测试集图像对所述卷积神经网络模型进行测试；

将待分类的图像输入到训练好的卷积神经网络模型中，提取所述待分类的图像的显著特征点，并根据所述显著特征点进行图像分类。

2、如权利要求1所述的图像分类方法，其特征在于，所述采用预设类型特征点提取算法分别提取各个所述图像样本中的显著特征点，并分别将各个所述图像样本对应的显著特征点保存为对应的图像样本的图像标注，包括：

逐一选择所述图像样本或者批量选择所述图像样本，在一个图像样本被选择后，针对该图像样本执行如下步骤：

利用该图像样本制作高斯金字塔：将该图像样本进行下采样得到N个不同尺寸的图像，再对每张图像使用k组不同大小的高斯核进行高斯模糊，得到N\*K个不同尺寸和模糊程度的图像，其中，N和K均为正整数；

将高斯金字塔转化为差分金字塔：在高斯金字塔中，每层的K张照片按照顺序两两相减，得到差分图像，最后得到N\*(K-1)张差分图像，组成所述差分金字塔；

对所述差分金字塔进行多尺度空间极值点检测，得到该图像样本的候选特征点；

对得到的该图像样本的候选特征点进行定位，并去其中的低对比度的候选特征点和不稳定的边缘候选特征点，得到该图像样本的显著特征点；

计算所述显著特征点的主方向，及构造所述显著特征点的描述子，并根据所述显著特征点的描述子，将所述显著特征点保存为该图像样本的图像标注。

3、如权利要求2所述的图像分类方法，其特征在于，所述对得到的该图像样本的候选特征点进行定位，并去掉其中的低对比度的候选特征点和不稳定的边缘候选特征点，得到该图像样本的显著特征点，包括：

将该图像样本 $I(x,y)$ 的尺度空间定义为： $L(x,y,\sigma) = G(x,y,\sigma) * I(x,y)$ ，其中 $L(x,y,\sigma)$ 代表该图像样本的尺度空间， $G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$ ， $\sigma$ 为像素的尺度参数，当 $\sigma$ 连续变化， $G(x,y,\sigma)$ 构成可变尺度空间的2维高斯函数；

对尺度空间 $G(x,y,\sigma)$ 进行二次泰勒展开： $D(x) = D + \frac{\partial D}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial x^2} x$ ，并进行求导，令 $D(x) = 0$ ，得到该图像样本的候选特征点的位置：

$$\hat{x} = -\frac{\partial^2 D^{-1}}{\partial x^2} \frac{\partial D}{\partial x};$$

将所述候选特征点带入到所述D(x)函数中，利用二次泰勒展开式进行最小二乘拟合，得到 $D(\hat{x}) = D + \frac{\partial D^T}{\partial x} \hat{x}$ ；

如果 $D(\hat{x}) \geq$  预设阈值，则保留该候选特征点，否则将该候选特征点作为低对比度的候选特征点丢弃；

将候选特征点的D(x)函数的主曲率与 $2 \times 2$  H的特征值成正比：H为候选特征点邻域对应位置的差分矩阵， $H = \begin{bmatrix} D_{xx} & D_{yx} \\ D_{xy} & D_{yy} \end{bmatrix}$ ，其中， $D_{xx}$ ， $D_{xy}$ ， $D_{yx}$ ， $D_{yy}$ 是候选特征点邻域对应位置的差分求得；

设 $\alpha = \lambda_{max}$ 为H的最大特征值， $\beta = \lambda_{min}$ 为H的最小特征值，则：  
 $Th(H) = D_{xx} + D_{yy} = \alpha + \beta$ ， $Det(H) = D_{xx} + D_{yy} - D_{xy}^2 = \alpha \cdot \beta$ ，其中， $Th(H)$ 为矩阵H的迹， $Det(H)$ 为矩阵H的行列式；

设 $\gamma = \frac{\alpha}{\beta}$ 表示最大特征值和最小特征值的比值，则 $\frac{Th(H)^2}{Det(H)} = \frac{(\alpha + \beta)^2}{\alpha \beta} = \frac{(\gamma \beta + \beta)^2}{\gamma \beta^2} = \frac{(\gamma + 1)^2}{\gamma}$ ；

当 $\frac{Th(H)^2}{Det(H)} > \frac{(T_\gamma + 1)^2}{T_\gamma}$ 时，剔除该候选特征点，否则保留该候选特征点，其中， $T_\gamma$ 为预设阈值。

4、如权利要求3所述的图像分类方法，其特征在于，所述构造所述显著特征点的描述子的步骤包括：

旋转所述显著特征点的坐标，使坐标轴旋转为所述显著特征点的主方向上，确保旋转不变性；

在旋转后的图像上，以所述显著特征点为中心取 $8 \times 8$ 大小的窗口，设所述显著特征点位于该窗口的中央位置，所述显著特征点周围的方格代表所述显著特征点的邻域范围，每个小方格代表一个像素，在每个像素上计算梯度，箭头的方向和长度分别表示像素的方向和模值；

将所述窗口平均分为 $4 \times 4$ 的子窗口，再利用高斯模糊方法增加与所述显著特征点邻近域的权重值、降低所述显著特征点较远的权重值，最后计算每个区域的 $0^\circ$ 、 $45^\circ$ 、 $90^\circ$ 、 $135^\circ$ 、 $180^\circ$ 、 $225^\circ$ 、 $270^\circ$ 、 $315^\circ$ 方向的梯度累加值，以得到所述显著特征点的描述子的特征向量。

5、如权利要求2至4中任意一项所述的图像分类方法，其特征在于，所述显著特征点的主方向为：

$$\theta(x, y) = \alpha \tan 2((L(x + 1, y) - L(x - 1, y)) / (L(x, y + 1) - L(x, y - 1)))$$

其中，L为每个显著特征点所在的尺度， $\alpha = \lambda_{max}$ 为H的最大特征值，H为显著特征点邻域对应位置的差分矩阵。

6、如权利要求1所述的图像分类方法，其特征在于，所述卷积神经网络模型为ResNet50模型，利用所述图像样本和上述得到的图像样本的图像标注

训练所述构建的 ResNet50 模型的步骤包括：

对训练集图像进行批量标准化；

将所述训练集图像输入至由输入层、卷积层、激活函数和输出层构成的 ResNet50 结构的卷积神经网络模型；

5 利用图像集合，构建标准 ResNet50 网络结构，计算显著特征点的损失函数与分类器的损失，并将其用于进行反向传播计算误差值，更新权重训练卷积神经网络模型，根据损失函数调整各层参数，不断迭代直至收敛，最终得到训练好的 ResNet50 模型；及

10 将待分类的图像输入到训练好的 ResNet50 模型中，提取所述待分类的图像的显著特征点，并根据所述显著特征点进行图像分类。

7、如权利要求 6 所述的图像分类方法，其特征在于，所述损失函数为：

$$\tilde{\ell}(y, z) = -\log\left(\frac{e^{z_y}}{\sum_{j=1}^m e^{z_j}}\right) = \log\left(\sum_{j=1}^m e^{z_j}\right) - z_y$$

y 为图像所属的类别编号，是真实的标签值；z 为图像所对应的数据，这个数据也是输入 softmax 层的数据，这个数据是 ResNet 输出的数据，即特征映射和分类器传出。

15 8、一种图像分类装置，其特征在于，所述装置包括存储器和处理器，所述存储器上存储有可在所述处理器上运行的图像分类程序，所述图像分类程序被所述处理器执行时实现如下步骤：

20 输入预设数量的图像样本，采用预设类型特征点提取算法分别提取各个所述图像样本中的显著特征点，并分别将各个所述图像样本对应的显著特征点保存为对应的图像样本的图像标注；

构建卷积神经网络模型，将所有所述图像样本和上述得到的图像样本的图像标注分为测试集图像和训练集图像，利用所述训练集图像训练所述卷积神经网络模型，并利用所述测试集图像对所述卷积神经网络模型进行测试；

25 将待分类的图像输入到训练好的卷积神经网络模型中，提取所述待分类的图像的显著特征点，并根据所述显著特征点进行图像分类。

9、如权利要求 8 所述的图像分类装置，其特征在于，采用预设类型特征点提取算法分别提取各个所述图像样本中的显著特征点，并分别将各个所述图像样本对应的显著特征点保存为对应的图像样本的图像标注，包括：

30 逐一选择所述图像样本或者批量选择所述图像样本，在一个图像样本被选择后，针对该图像样本执行如下步骤：

利用该图像样本制作高斯金字塔：将该图像样本进行下采样得到 N 个不同尺寸的图像，再对每张图像使用 k 组不同大小的高斯核进行高斯模糊，得到 N \* K 个不同尺寸和模糊程度的图像，其中，N 和 K 均为正整数；

35 将高斯金字塔转化为差分金字塔：在高斯金字塔中，每层的 K 张照片按照顺序两两相减，得到差分图像，最后得到 N \* (K-1) 张差分图像，组成所述差分金字塔；

对所述差分金字塔进行多尺度空间极值点检测，得到该图像样本的候选特征点；

对得到的该图像样本的候选特征点进行定位，并去掉其中的低对比度的候选特征点和不稳定的边缘候选特征点，得到该图像样本的显著特征点；

5 计算所述显著特征点的主方向，及构造所述显著特征点的描述子，并根据所述显著特征点的描述子，将所述显著特征点保存为该图像样本的图像标注。

10、如权利要求 9 所述的图像分类装置，其特征在于，所述对得到的该图像样本的候选特征点进行定位，并去掉其中的低对比度的候选特征点和不稳定的边缘候选特征点，得到该图像样本的显著特征点，包括：

将该图像样本  $I(x,y)$  的尺度空间定义为： $L(x,y,\sigma) = G(x,y,\sigma) * I(x,y)$ ，其中  $L(x,y,\sigma)$  代表着该图像样本的尺度空间， $G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$ ， $\sigma$  为像素的尺度参数，当  $\sigma$  连续变化， $G(x,y,\sigma)$  构成可变尺度空间的 2 维高斯函数；

15 对尺度空间  $G(x,y,\sigma)$  进行二次泰勒展开： $D(x) = D + \frac{\partial D^T}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial x^2} x$ ，并进行求导，令  $D(x) = 0$ ，得到该图像样本的候选特征点的位置： $\hat{x} = -\frac{\partial^2 D^{-1}}{\partial x^2} \frac{\partial D}{\partial x}$ ；

将所述候选特征点带入到所述  $D(x)$  函数中，利用二次泰勒展开式进行最小二乘拟合，得到  $D(\hat{x}) = D + \frac{\partial D^T}{\partial x} \hat{x}$ ；

20 如果  $D(\hat{x}) \geq$  预设阈值，则保留该候选特征点，否则将该候选特征点作为低对比度特征点丢弃；

将候选特征点的  $D(x)$  函数的主曲率与  $2 \times 2$   $H$  的特征值成正比： $H$  为候选特征点邻域对应位置的差分矩阵， $H = \begin{bmatrix} D_{xx} & D_{yx} \\ D_{xy} & D_{yy} \end{bmatrix}$ ，其中  $D_{xx}$ ， $D_{xy}$ ， $D_{yx}$ ， $D_{yy}$  是候选特征点邻域对应位置的差分求得；

25 设  $\alpha = \lambda_{max}$  为  $H$  的最大特征值， $\beta = \lambda_{min}$  为  $H$  的最小特征值，则： $Th(H) = D_{xx} + D_{yy} = \alpha + \beta$ ， $Det(H) = D_{xx} + D_{yy} - D_{xy}^2 = \alpha \cdot \beta$ ，其中  $Th(H)$  为矩阵  $H$  的迹， $Det(H)$  为矩阵  $H$  的行列式；

设  $\gamma = \frac{\alpha}{\beta}$  表示最大特征值和最小特征值的比值，则  $\frac{Th(H)^2}{Det(H)} = \frac{(\alpha+\beta)^2}{\alpha\beta} = \frac{(\gamma\beta+\beta)^2}{\gamma\beta^2} = \frac{(\gamma+1)^2}{\gamma}$ ；

30 当  $\frac{Th(H)^2}{Det(H)} > \frac{(T_\gamma+1)^2}{T_\gamma}$  时，剔除该候选特征点，否则保留该候选特征点，其中， $T_\gamma$  为预设阈值。

11、如权利要求 9 所述的图像分类装置，其特征在于，所述构造所述显著特征点的描述子的步骤包括：

旋转所述显著特征点的坐标，使坐标轴旋转为所述显著特征点的主方向

上，确保旋转不变性；

在旋转后的图像上，以所述显著特征点为中心取 8\*8 大小的窗口，设所述显著特征点位于该窗口的中央位置，所述显著特征点周围的方格代表所述显著特征点的邻域范围，每个小方格代表一个像素，在每个像素上计算梯度，箭头的方向和长度分别表示像素的方向和模值；

将所述窗口平均分为 4\*4 的子窗口，再利用高斯模糊方法增加与所述显著特征点邻近域的权重值、降低所述显著特征点较远的权重值，最后计算每个区域的 0°、45°、90°、135°、180°、225°、270°、315°方向的梯度累加值，以得到所述显著特征点的描述子的特征向量。

12、如权利要求 8 至 10 中任意一项所述的图像分类装置，其特征在于，所述显著特征点的主方向为：

$$\theta(x, y) = \alpha \tan 2((L(x+1, y) - L(x-1, y)) / (L(x, y+1) - L(x, y-1)))$$

其中，L 为每个显著特征点所在的尺度， $\alpha = \lambda_{max}$  为 H 的最大特征值，H 为显著特征点邻域对应位置的差分矩阵。

13、如权利要求 8 所述的图像分类装置，其特征在于，所述卷积神经网络模型为 ResNet50 模型，利用所述图像样本和上述得到的图像样本的图像标注训练所述构建的 ResNet50 模型的步骤包括：

对训练集图像进行批量标准化；

将所述训练集图像输入至由输入层、卷积层、激活函数和输出层构成的 ResNet50 结构的卷积神经网络模型；

利用图像集合，构建标准 ResNet50 网络结构，计算显著特征点的损失函数与分类器的损失，并将其用于进行反向传播计算误差值，更新权重训练卷积神经网络模型，根据损失函数调整各层参数，不断迭代直至收敛，最终得到训练好的 ResNet50 模型；及

将待分类的图像输入到训练好的 ResNet50 模型中，提取所述待分类的图像的显著特征点，并根据所述显著特征点进行图像分类。

14、如权利要求 13 所述的图像分类装置，其特征在于，所述损失函数为：

$$\tilde{\ell}(y, z) = -\log\left(\frac{e^{z_y}}{\sum_{j=1}^m e^{z_j}}\right) = \log\left(\sum_{j=1}^m e^{z_j}\right) - z_y$$

y 为图像所属于的类别编号，是真实的标签值；z 为图像所对应的数据，这个数据也是输入 softmax 层的数据，这个数据是 ResNet 输出的数据，即特征映射和分类器传出。

15、一种计算机可读存储介质，其特征在于，所述计算机可读存储介质上存储有图像分类程序，所述图像分类程序可被一个或者多个处理器执行，以实现如权利要求 1 至 5 中任一项所述的图像分类方法的步骤：

输入预设数量的图像样本，采用预设类型特征点提取算法分别提取各个所述图像样本中的显著特征点，并分别将各个所述图像样本对应的显著特征点保存为对应的图像样本的图像标注；

构建卷积神经网络模型，将所有所述图像样本和上述得到的图像样本的图像标注分为测试集图像和训练集图像，利用所述训练集图像训练所述卷积神经网络模型，并利用所述测试集图像对所述卷积神经网络模型进行测试；

将待分类的图像输入到训练好的卷积神经网络模型中，提取所述待分类的图像的显著特征点，并根据所述显著特征点进行图像分类。

16、如权利要求 15 所述的计算机可读存储介质，其特征在于，采用预设类型特征点提取算法分别提取各个所述图像样本中的显著特征点，并分别将各个所述图像样本对应的显著特征点保存为对应的图像样本的图像标注，包括：

逐一选择所述图像样本或者批量选择所述图像样本，在一个图像样本被选择后，针对该图像样本执行如下步骤：

利用该图像样本制作高斯金字塔：将该图像样本进行下采样得到  $N$  个不同尺寸的图像，再对每张图像使用  $k$  组不同大小的高斯核进行高斯模糊，得到  $N * K$  个不同尺寸和模糊程度的图像，其中， $N$  和  $K$  均为正整数；

将高斯金字塔转化为差分金字塔：在高斯金字塔中，每层的  $K$  张照片按照顺序两两相减，得到差分图像，最后得到  $N * (K-1)$  张差分图像，组成所述差分金字塔；

对所述差分金字塔进行多尺度空间极值点检测，得到该图像样本的候选特征点；

对得到的该图像样本的候选特征点进行定位，并去掉其中的低对比度的候选特征点和不稳定的边缘候选特征点，得到该图像样本的显著特征点；

计算所述显著特征点的主方向，及构造所述显著特征点的描述子，并根据所述显著特征点的描述子，将所述显著特征点保存为该图像样本的图像标注。

17、如权利要求 16 所述的计算机可读存储介质，其特征在于，所述对得到的该图像样本的候选特征点进行定位，并去掉其中的低对比度的候选特征点和不稳定的边缘候选特征点，得到该图像样本的显著特征点，包括：

将该图像样本  $I(x,y)$  的尺度空间定义为： $L(x,y,\sigma) = G(x,y,\sigma) * I(x,y)$ ，其中  $L(x,y,\sigma)$  代表着该图像样本的尺度空间， $G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$ ， $\sigma$  为像素的尺度参数，当  $\sigma$  连续变化， $G(x,y,\sigma)$  构成可变尺度空间的 2 维高斯函数；

对尺度空间  $G(x,y,\sigma)$  进行二次泰勒展开： $D(x) = D + \frac{\partial D^T}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial x^2} x$ ，并进行求导，令  $D(x) = 0$ ，得到该图像样本的候选特征点的位置： $\hat{x} = -\frac{\partial^2 D^{-1}}{\partial x^2} \frac{\partial D}{\partial x}$ ；

将所述候选特征点带入到所述  $D(x)$  函数中，利用二次泰勒展开式进行最小二乘拟合，得到  $D(\hat{x}) = D + \frac{\partial D^T}{\partial x} \hat{x}$ ；

如果  $D(\hat{x}) \geq$  预设阈值，则保留该候选特征点，否则将该候选特征点作为

低对比度特征点丢弃；

将候选特征点的  $D(x)$  函数的主曲率与  $2 \times 2$   $H$  的特征值成正比： $H$  为候选特征点邻域对应位置的差分矩阵， $H = \begin{bmatrix} D_{xx} & D_{yx} \\ D_{xy} & D_{yy} \end{bmatrix}$ ，其中  $D_{xx}$ 、 $D_{xy}$ 、 $D_{yx}$ 、 $D_{yy}$  是候选特征点邻域对应位置的差分求得；

5 设  $\alpha = \lambda_{max}$  为  $H$  的最大特征值， $\beta = \lambda_{min}$  为  $H$  的最小特征值，则： $Th(H) = D_{xx} + D_{yy} = \alpha + \beta$ ， $Det(H) = D_{xx} + D_{yy} - D_{xy}^2 = \alpha \cdot \beta$ ，其中  $Th(H)$  为矩阵  $H$  的迹， $Det(H)$  为矩阵  $H$  的行列式；

设  $\gamma = \frac{\alpha}{\beta}$  表示最大特征值和最小特征值的比值，则  $\frac{Th(H)^2}{Det(H)} = \frac{(\alpha + \beta)^2}{\alpha \beta} = \frac{(\gamma \beta + \beta)^2}{\gamma \beta^2} = \frac{(\gamma + 1)^2}{\gamma}$ ；

10 当  $\frac{Th(H)^2}{Det(H)} > \frac{(T_\gamma + 1)^2}{T_\gamma}$  时，剔除该候选特征点，否则保留该候选特征点，其中， $T_\gamma$  为预设阈值。

18、如权利要求 16 所述的计算机可读存储介质，其特征在于，所述构造所述显著特征点的描述子的步骤包括：

15 旋转所述显著特征点的坐标，使坐标轴旋转为所述显著特征点的主方向上，确保旋转不变性；

在旋转后的图像上，以所述显著特征点为中心取  $8 \times 8$  大小的窗口，设所述显著特征点位于该窗口的中央位置，所述显著特征点周围的方格代表所述显著特征点的邻域范围，每个小方格代表一个像素，在每个像素上计算梯度，箭头的方向和长度分别表示像素的方向和模值；

20 将所述窗口平均分为  $4 \times 4$  的子窗口，再利用高斯模糊方法增加与所述显著特征点邻近域的权重值、降低所述显著特征点较远的权重值，最后计算每个区域的  $0^\circ$ 、 $45^\circ$ 、 $90^\circ$ 、 $135^\circ$ 、 $180^\circ$ 、 $225^\circ$ 、 $270^\circ$ 、 $315^\circ$  方向的梯度累加值，以得到所述显著特征点的描述子的特征向量。

25 19、如权利要求 15 所述的计算机可读存储介质，其特征在于，所述显著特征点的主方向为：

$$\theta(x, y) = \alpha \tan 2((L(x + 1, y) - L(x - 1, y)) / (L(x, y + 1) - L(x, y - 1)))$$

其中， $L$  为每个显著特征点所在的尺度， $\alpha = \lambda_{max}$  为  $H$  的最大特征值， $H$  为显著特征点邻域对应位置的差分矩阵。

30 20、如权利要求 15 所述的计算机可读存储介质，其特征在于，所述卷积神经网络模型为 ResNet50 模型，利用所述图像样本和上述得到的图像样本的图像标注训练所述构建的 ResNet50 模型的步骤包括：

对训练集图像进行批量标准化；

将所述训练集图像输入至由输入层、卷积层、激活函数和输出层构成的 ResNet50 结构的卷积神经网络模型；

利用图像集合，构建标准 ResNet50 网络结构，计算显著特征点的损失函

数与分类器的损失，并将其用于进行反向传播计算误差值，更新权重训练卷积神经网络模型，根据损失函数调整各层参数，不断迭代直至收敛，最终得到训练好的 ResNet50 模型；及

- 5 将待分类的图像输入到训练好的 ResNet50 模型中，提取所述待分类的图像的显著特征点，并根据所述显著特征点进行图像分类。

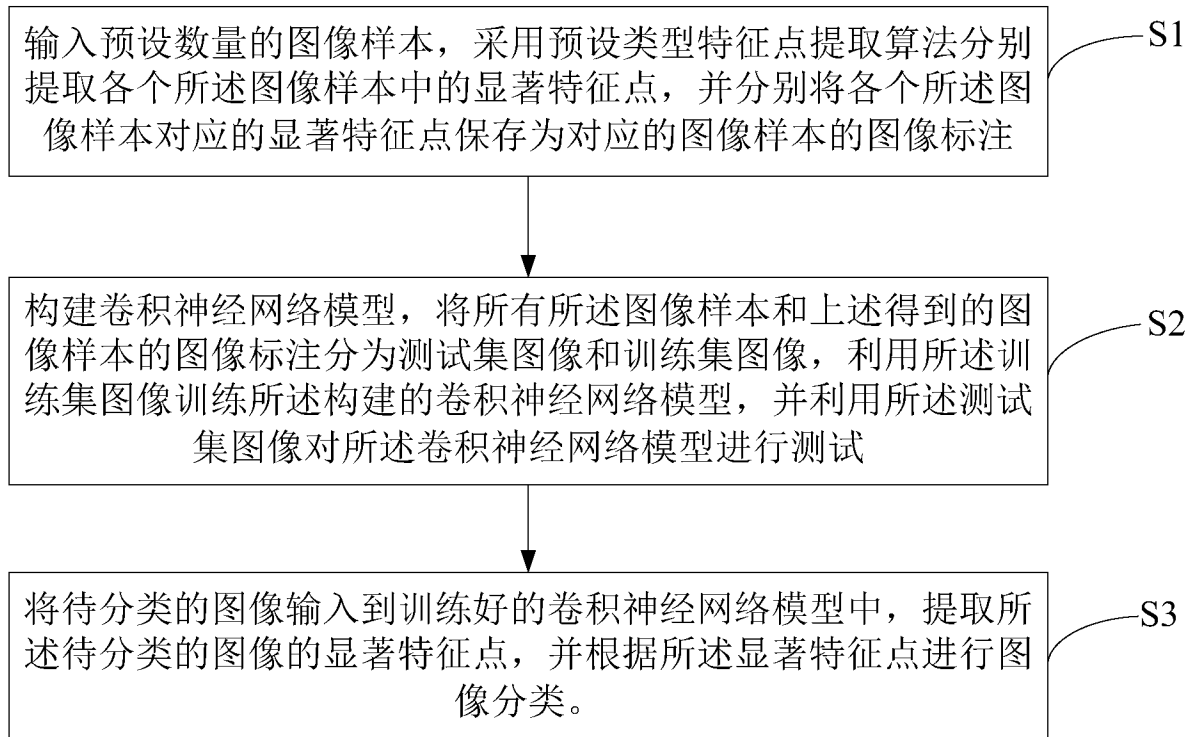


图 1

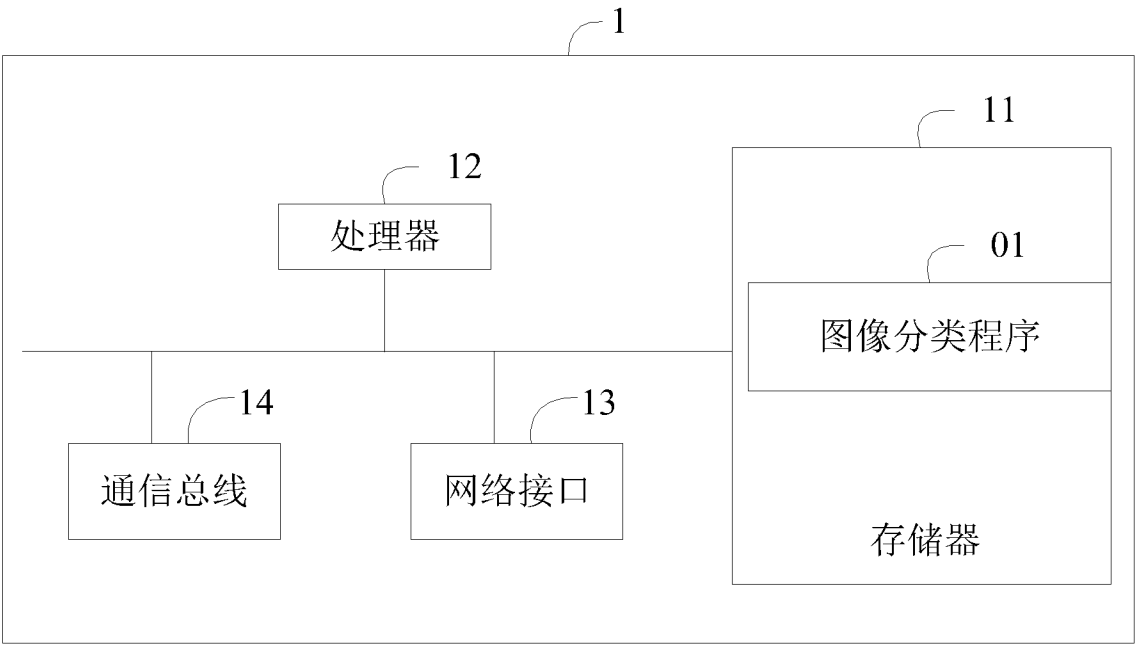


图 2

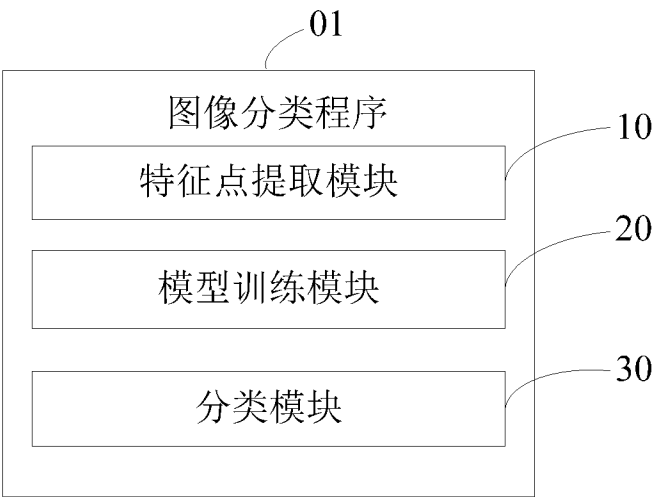


图 3

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/102186

**A. CLASSIFICATION OF SUBJECT MATTER**

G06K 9/32(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

G06K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT, CNKI, WPI, EPODOC: 中国科学技术大学, 吴朝, X射线, 相位衬度, 相衬, 成像, 背景, 光栅, 位移, 移动, 步进, 光强曲线, 背景位移曲线, 光强, 可见度, 对比度, 衬比度, 初始相位, 傅里叶, 傅立叶, 灰度, 正, 反, 图像, x-ray, phase-contrast, phase contrast, imaging, XPCI, stepping, grating, forward image, reverse image

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

| Category* | Citation of document, with indication, where appropriate, of the relevant passages                                                            | Relevant to claim No. |
|-----------|-----------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| PX        | CN 110097051 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 06 August 2019 (2019-08-06)<br>claims 1-10 and description, paragraphs [0064]-[0115] | 1-20                  |
| X         | CN 105894037 A (BEIHANG UNIVERSITY et al.) 24 August 2016 (2016-08-24)<br>description, paragraphs [0055]-[0088]                               | 1-3, 6-10, 13-17, 20  |
| Y         | CN 105894037 A (BEIHANG UNIVERSITY et al.) 24 August 2016 (2016-08-24)<br>description, paragraphs [0055]-[0088]                               | 4, 5, 11, 12, 18, 19  |
| Y         | CN 106446965 A (NORTHWESTERN POLYTECHNICAL UNIVERSITY) 22 February 2017 (2017-02-22)<br>claim 1                                               | 4, 5, 11, 12, 18, 19  |
| A         | CN 107871098 A (BEIJING EYECOOOL TECHNOLOGY CO., LTD.) 03 April 2018 (2018-04-03)<br>entire document                                          | 1-20                  |
| A         | CN 104700078 A (WUHAN INSTITUTE OF TECHNOLOGY) 10 June 2015 (2015-06-10)<br>entire document                                                   | 1-20                  |

☐ Further documents are listed in the continuation of Box C.☒ See patent family annex.

\* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance  
 “E” earlier application or patent but published on or after the international filing date  
 “L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)  
 “O” document referring to an oral disclosure, use, exhibition or other means  
 “P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention  
 “X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone  
 “Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art  
 “&” document member of the same patent family

Date of the actual completion of the international search

20 December 2019

Date of mailing of the international search report

30 December 2019

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/  
CN)  
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing  
100088  
China

Authorized officer

Facsimile No. (86-10)62019451

Telephone No.

**INTERNATIONAL SEARCH REPORT**  
**Information on patent family members**

International application No.

**PCT/CN2019/102186**

| Patent document<br>cited in search report |           |   | Publication date<br>(day/month/year) | Patent family member(s) | Publication date<br>(day/month/year) |
|-------------------------------------------|-----------|---|--------------------------------------|-------------------------|--------------------------------------|
| CN                                        | 110097051 | A | 06 August 2019                       | None                    |                                      |
| CN                                        | 105894037 | A | 24 August 2016                       | None                    |                                      |
| CN                                        | 106446965 | A | 22 February 2017                     | None                    |                                      |
| CN                                        | 107871098 | A | 03 April 2018                        | None                    |                                      |
| CN                                        | 104700078 | A | 10 June 2015                         | None                    |                                      |

## 国际检索报告

国际申请号

PCT/CN2019/102186

## A. 主题的分类

G06K 9/32 (2006.01) i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

## B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G06K

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNPAT;CNKI;WPI;EPDOC:图像, 分类, 识别, 特征, 显著, 关键, 神经网络, 训练, 测试, 模型, image, recognition, classify, characteristic, prominent, key, neural network, train, test, model

## C. 相关文件

| 类 型* | 引用文件, 必要时, 指明相关段落                                                                        | 相关的权利要求              |
|------|------------------------------------------------------------------------------------------|----------------------|
| PX   | CN 110097051 A (平安科技深圳有限公司) 2019年 8月 6日 (2019 - 08 - 06)<br>权利要求1-10以及说明书第[0064]-[0115]段 | 1-20                 |
| X    | CN 105894037 A (北京航空航天大学 等) 2016年 8月 24日 (2016 - 08 - 24)<br>说明书第[0055]-[0088]段          | 1-3, 6-10, 13-17, 20 |
| Y    | CN 105894037 A (北京航空航天大学 等) 2016年 8月 24日 (2016 - 08 - 24)<br>说明书第[0055]-[0088]段          | 4, 5, 11, 12, 18, 19 |
| Y    | CN 106446965 A (西北工业大学) 2017年 2月 22日 (2017 - 02 - 22)<br>权利要求1                           | 4, 5, 11, 12, 18, 19 |
| A    | CN 107871098 A (北京眼神科技有限公司) 2018年 4月 3日 (2018 - 04 - 03)<br>全文                           | 1-20                 |
| A    | CN 104700078 A (武汉工程大学) 2015年 6月 10日 (2015 - 06 - 10)<br>全文                              | 1-20                 |

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

\* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&amp;” 同族专利的文件

国际检索实际完成的日期

2019年 12月 20日

国际检索报告邮寄日期

2019年 12月 30日

ISA/CN的名称和邮寄地址

中国国家知识产权局(ISA/CN)

中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

受权官员

焦月

电话号码 86-(10)-53961306

国际检索报告  
关于同族专利的信息

国际申请号

PCT/CN2019/102186

| 检索报告引用的专利文件 |           |   | 公布日<br>(年/月/日) | 同族专利 | 公布日<br>(年/月/日) |
|-------------|-----------|---|----------------|------|----------------|
| CN          | 110097051 | A | 2019年 8月 6日    | 无    |                |
| CN          | 105894037 | A | 2016年 8月 24日   | 无    |                |
| CN          | 106446965 | A | 2017年 2月 22日   | 无    |                |
| CN          | 107871098 | A | 2018年 4月 3日    | 无    |                |
| CN          | 104700078 | A | 2015年 6月 10日   | 无    |                |