



(12) 发明专利申请

(10) 申请公布号 CN 104850845 A

(43) 申请公布日 2015. 08. 19

(21) 申请号 201510290648. 1

(22) 申请日 2015. 05. 30

(71) 申请人 大连理工大学

地址 116024 辽宁省大连市甘井子区凌工路  
2号

(72) 发明人 葛宏伟 何鹏程 孙亮 谭贞刚

(74) 专利代理机构 大连理工大学专利中心  
21200

代理人 关慧贞 梅洪玉

(51) Int. Cl.

G06K 9/00(2006. 01)

G06K 9/62(2006. 01)

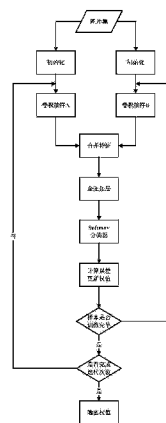
权利要求书2页 说明书5页 附图3页

(54) 发明名称

一种基于非对称卷积神经网络的交通标志识别方法

(57) 摘要

本发明属于智能交通标志识别技术领域, 涉及一种基于非对称卷积神经网络的交通标志识别方法, 用以解决交通标志识别问题中识别速度较慢, 且鲁棒性不强的问题。本发明的方法采用两个不同结构的卷积神经网络并行地进行特征映射和提取, 最后将特征进行合并, 再经过全连接层和最后的分类器, 完成整个分类过程。两个不同结构的卷积神经网络分别采用了随机池化操作和maxout单元, 确保了图像特征的多样性, 提高了识别精度并加快了网络运算速度。本发明对传统卷积神经网络进行了结构上的改进, 使用两个不同结构的卷积神经网络代替了传统的卷积神经网络结构, 该方法确保了图像特征的多样性, 提高了识别精度并加快了网络运算速度。



1. 一种基于非对称卷积神经网络的交通标志识别方法,包含图像预处理、非对称卷积神经网络的训练、准确率验证和图像识别四个部分;其特征在于以下步骤,

#### 一、图像预处理

取原始图像中的目标区域取出,将彩色图像转化为灰度图像,并缩放到同一的  $A \times A$  大小,为增强对比度,再将目标区域经过对比度增强处理,得到原始的训练集,测试集采用同样的方式处理;将原始训练集经过旋转  $[-10^\circ, 10^\circ]$ , 缩放  $[0.9, 1.1]$  后,加入原始数据集中,组成新的训练集;在新的数据集中随机取出与测试集数目相当的样本组成验证集,剩余的样本组成最终的训练集;

#### 二、非对称卷积神经网络的训练

模型训练方法采用 min batch 反向传播方法,每次从训练集中取固定数量的样本参加培训;

##### (1) 前向传播阶段

①从训练集中取出训练样本  $(x_t, y_t)$ , 将  $x_t$  作为两个独立的卷积神经网络的输入,  $y_t$  作为前向传播结果的比对标准;

②采用随机 pooling 的卷积神经网络卷积过程为,采用可训练的卷积核  $W$  与输入图像做卷积,得到的卷积加上一个可训练的偏置  $b$ , 然后经过 ReLU 激活函数得到卷积特征图  $C_{1s}$ ;而采用 maxout 的卷积神经网络是采用训练的方法求得激活函数,即在加上可训练的偏置  $b$  之后,须在  $k$  个卷积特征图中,按每个位置选取值最大的作为最后的神经元激活值,得到卷积特征图  $C_{1m}$ ;

③在随机 pooling 的卷积神经网络中,从 pooling 区域的多项式分布中选择出该 pooling 区域的响应值;多项式分布概率根据下面的公式计算得出,

$$p_i = \frac{a_i}{\sum_{k \in R_j} a_k} \quad (1)$$

$R_j$  是第  $j$  个 pooling 区域,  $a_i, a_k$  分别为 pooling 区域中元素  $i$  和  $k$  的值,  $p_i$  为 pooling 区域中元素  $i$  对应的概率;根据得到的概率,从  $a_i$  中采样得到 pooling 结果:

$$s_j = a_l \text{ where } l \sim P(p_1, \dots, p_{|R_j|}) \quad (2)$$

通过以上操作,得到 pooling 特征图  $S_{2s}$ ;

在 maxout 卷积神经网络中,采用最大值采样,即在 pooling 区域中取最大激活值,通过上述操作得到 pooling 特征图  $C_{2m}$ ;

④重复②、③,依次得到特征图  $C_{3s}$ 、 $C_{3m}$ 、 $S_{4s}$  和  $S_{4m}$ ;

⑤将  $S_{4s}$  和  $S_{4m}$  合并,形成图像的特征,经过全连接层之后送入 softmax 分类器,得到分类结果  $O_t$ ;

##### (2) 反向传播阶段

①计算实际输出  $O_t$  与相应的理想输出  $y_t$  的差;

②计算整体的代价函数为

$$J(W, b) = \frac{1}{m} \sum_{i=1}^m \left( \frac{1}{2} \left\| o_t^{(i)} - y_t^{(i)} \right\|^2 \right) \quad (3)$$

反向传播,对于第  $n_1$  层(输出层)的每个输出单元  $i$ ,我们根据以下公式计算残差:

$$\delta_i^{(n_l)} = -\left(y_i - a_i^{(n_l)}\right) \cdot f' \left(z_i^{(n_l)}\right) \quad (4)$$

对  $l = n_1-1, n_1-2, n_1-3, \dots, 2$  的各个层,第  $l$  层的第  $i$  个节点的残差计算方法如下:

$$\delta_i^{(l)} = \left(\sum_{j=1}^{s_{l+1}} W_{ji}^{(l)} \delta_j^{(l+1)}\right) \cdot f' \left(z_i^{(l)}\right) \quad (5)$$

计算偏导数,计算方法如下:

$$\nabla_{W^{(l)}} J(W, b; x, y) = \delta^{(l+1)} (a^{(l)})^T \quad (6)$$

$$\nabla_{b^{(l)}} J(W, b; x, y) = \delta^{(l+1)} \quad (7)$$

$$\Delta W^{(l)} := \Delta W^{(l)} + \nabla_{W^{(l)}} J(W, b; x, y) \quad (8)$$

$$\Delta b^{(l)} := \Delta b^{(l)} + \nabla_{b^{(l)}} J(W, b; x, y) \quad (9)$$

更新权重参数:

$$W^{(l)} = W^{(l)} - \alpha \left(\frac{1}{m} \Delta W^{(l)}\right) \quad (10)$$

$$b^{(l)} = b^{(l)} - \alpha \left[\frac{1}{m} \Delta b^{(l)}\right] \quad (11)$$

至此  $l$  次迭代完成,保存调整后的权值,重复第一阶段,直至达到设定的迭代次数后,训练结束,保存并把此权值输出给验证集以验证;

### 三、准确率验证

使用验证集验证输入的权值,将输入和映射的结果进行对比,输出错误率,如果错误率达到满意的要求,则可以进行测试识别,否则,调整网络规模或者增加迭代次数重新进行训练;

### 四、图像识别

将检测到的交通标志图像经过预处理之后,使用训练好权值的网络模型进行识别并输出其对应的类别。

## 一种基于非对称卷积神经网络的交通标志识别方法

### 技术领域

[0001] 本发明属于智能交通标志识别技术领域,涉及一种基于非对称卷积神经网络的交通标志识别方法,用以解决交通标志识别问题中识别速度较慢,且鲁棒性不强的问题。

### 背景技术

[0002] 近年来,智能交通系统得到了很大的发展,2010年,奥迪无人驾驶自动汽车行驶12.42公里,抵达落基山派克峰顶,标志着智能交通系统中的无人驾驶汽车技术逐渐走向成熟,而交通标志自动识别是无人驾驶汽车技术的重要组成部分。

[0003] 交通标志识别有着直接的现实应用,如安全驾驶,自动驾驶,场景理解和信号检测等。交通标志的识别是一个受限的分类问题,因为交通标志类别数目是一定的,并且每一类交通标志的形态是严格一致的。通过前面的介绍,交通标志识别似乎是比较容易的问题,其实不然。在实际的交通标志识别过程中,会存在视角变化、不同的光照条件、运动模糊、物理损坏、遮挡、涂鸦和褪色等多种问题,给交通标志的识别问题造成了困难。

[0004] 分类算法通常分为两类,一类是学习型算法,通过学习找到一个最优的分割平面,将两类或者多类区分开来,比如SVM(Support Vector Machine,支持向量机),Adaboost(Adaptive Boosting,自适应增强),神经网络等;另一类为最邻近算法,是在已有的训练样本中找到与测试样本最近的一个,进而判断测试样本类别,比如说模板匹配,k-d(k-dimensionality,k维)树等。

[0005] 针对道路交通标志的识别,已经有不少方法被提出。FatinZaklouta等采用k-d树和随机森林的方法来对交通标志进行分类(Zaklouta F, Stanciulescu B, Hamdoun O. Traffic sign classification using kd trees and random forests[C]. Neural Networks(IJCNN), The 2011 International Joint Conference on. IEEE, 2011:2151-2155.),具有快速建立、更新和搜索等优点,且更加适合非平衡数据。Yaxiang Fan等采用层次性稀疏表示方法(Fan Y, Sun H, Zhou S, et al. Hierarchical Sparse Representation for Traffic Sign Recognition[C]. Proceedings of 2013 Chinese Intelligent Automation Conference. Springer Berlin Heidelberg, 2013:653-660.),该方法是利用了稀疏表示对噪音和遮挡具有鲁棒性,且交通标志不同大类之间的区别明显,而同一类之间的区别相对来说不明显等特点。Suisui Tang等利用在计算机视觉中已被证明有效的多个特征,HOG(Histogram of Oriented Gradient,梯度方向直方图)、Garbor滤波器特征和LBP(Local Binary Pattern,局部二值模式)特征,将这三个特征组合起来,用SVM进行分类(Tang S, Huang L L. Traffic Sign Recognition Using Complementary Features[C]. Pattern Recognition(ACPR), 2013 2nd IAPR Asian Conference on. IEEE, 2013:210-214.),减小了计算量,加快了分类速度。Ciresan等提出的Multi-column deep neural network通过将多个卷积神经网络的结果通过线性组合得到最后的结果(Ciresan D, Meier U, Masci J, et al. Multi-column deep neural network for traffic sign classification[J]. Neural

Networks, 2012, 32:333-338.), 取得了非常好的效果; Pierre Sermanet 等提出的多尺度卷积神经网络 (Sermanet P, LeCun Y. Traffic sign recognition with multi-scale convolutional networks[C]. Neural Networks (IJCNN), The 2011 International Joint Conference on. IEEE, 2011:2809-2813.), 将第一层子采样层和第二层子采样层的结果组合后, 形成多尺度特征, 再经过分类器进行分类, 还有 Ciresan 等结合卷积神经网络与 HOG+SVM 分类等等, 在交通标志分类上面都取得了很好的效果。

[0006] 卷积神经网络因为其对比例缩放和平移等变形有着不变形, 在很多领域, 包括交通标志识别领域都取得了很优秀的结果。

[0007] 传统的卷积神经网络中, 在同一层通常对前一层得到的特征是采用的同样的处理方式, 因此最终得到的特征类型单一, 可能只适用于某一类的分类任务, 而且分类效果有待提升。在这样的背景下, 有不少文章都提出了改进措施, 例如 bagging[Breiman, 1994] 将多个不同模型在不同的子集上训练, 取得了不错的效果, Multi-column deep neural network[Ciresan] 将多个相同的模型在不同的数据变形中训练, 最后将分类结果进行线性组合, 最近提出的 dropout 在训练的时候将神经结点以 50% 的概率随机隐藏 (置为 0) 起来, 而在测试的时候不再隐藏结点激活值, 而是将激活值乘以 50%, 取得了不错的效果。Stochastic pooling 是专门针对卷积神经网络提出来的一种平均模型, 在训练时 pooling 阶段按权值随机选取 pooling 结果, 测试阶段采用按权值加权得到 pooling 结果。

[0008] 后面的两者可以看做是隐式的平均模型, 平均操作隐含在网络内部操作中, 而前面两种是显式的平均模型, 平均操作采用显而易见的组合方式实现。平均模型因为可以学习到多样的特征, 因此在分类任务中, 效果很好。

[0009] 其中, 显式的平均模型 Multi-column deep neural network[Ciresan] 取得了 2011 年, IJCNN 举办了 GTSRB (German Traffic Sign Recognition Benchmark) 的交通标志识别比赛的最好结果, 准确率达到 99.46%, 优于人类的表现。

[0010] 但是, 在该方法中因为采用了多个模型且各个模型的输入数据不同, 所以模型总的的数据量相对来说更大, 总的模型规模也更大, 不利于实际运用。

[0011] 为了提升效率, 总的卷积神经网络结构必然不会很大, 而且网络结构过大也会对学习过程造成更大的困难, 但是较小的网络结构产生的单一特征在分类时又不足以达到足够高的准确度。

## 发明内容

[0012] 本发明要解决的技术问题是卷积神经网络在进行交通标志识别时, 识别精度不高且识别速度较慢的问题, 提出了一种基于非对称性卷积神经网络的交通标志识别方法。

[0013] 本发明的技术方案包含图像预处理、非对称卷积神经网络的训练、准确率验证和图像识别四个部分。

[0014] 在非对称卷积神经网络中, 两个并行的卷积神经网络分别是采用随机池化的卷积神经网络和采用 maxout 单元的卷积神经网络, 卷积神经网络是由交替的卷积层和采样层组成的, 每个卷积神经网络共四层, 两个卷积神经网络的第四层得到所特征图合并成一个特征, 经过全连接层之后送入 softmax 分类器, 得到分类结果。如附图 2。

[0015] 1. 图像预处理部分

[0016] 取原始图像中的目标区域（交通标志区域）取出，为减少计算量，将彩色图像转化为灰度图像，并缩放到同一的  $A \times A$  大小，为增强对比度，再将目标区域经过对比度增强处理，得到原始的训练集，测试集采用同样的方式处理。为使所训练的模型有更好的鲁棒性，将原始训练集经过旋转  $[-10^\circ, 10^\circ]$ ，缩放  $[0.9, 1.1]$  后，加入原始数据集中，组成新的训练集，在新的数据集中随机取出与测试集数目相当的样本组成验证集，剩余的样本组成最终的训练集。处理后得到的部分图像如附图 1。

[0017] 2. 非对称卷积神经网络的训练阶段

[0018] 模型训练方法采用 min batch 反向传播方法，每次从训练集中取固定数量的样本参加训练。

[0019] (1) 前向传播阶段

[0020] ①从训练集中取出训练样本  $(x_t, y_t)$ ，将  $x_t$  作为两个独立的卷积神经网络的输入， $y_t$  作为前向传播结果的比对标准。

[0021] ②采用随机池化的卷积神经网络卷积过程为，采用可训练的卷积核  $W$  与输入图像做卷积，得到的卷积加上一个可训练的偏置  $b$ ，然后经过 ReLU 激活函数得到卷积特征图  $C_{1s}$ ；而采用 maxout 的卷积神经网络是采用训练的方法求得激活函数，即在加上可训练的偏置  $b$  之后，须在  $k$  个卷积特征图中，按每个位置选取最大值的作为最后的神经元激活值，得到卷积特征图  $C_{1m}$ 。

[0022] ③在随机池化的卷积神经网络中，从池化区域的多项式分布中随机选择出该池化区域的响应值。多项式分布概率根据下面的公式计算得出，

$$[0023] \quad p_i = \frac{a_i}{\sum_{k \in R_j} a_k} \quad (1)$$

[0024]  $R_j$  是第  $j$  个池化区域， $a_i$ 、 $a_k$  分别为池化区域中元素  $i$  和  $k$  的值， $p_i$  为池化区域中元素  $i$  对应的概率。根据得到的概率，从  $a_i$  中采样得到池化结果：

$$[0025] \quad s_j = a_l \text{ where } l \sim P(p_1, \dots, p_{|R_j|}) \quad (2)$$

[0026] 通过以上操作，得到池化特征图  $S_{2s}$ 。

[0027] 在 maxout 卷积神经网络中，采用最大值采样，即在池化区域中取最大激活值，通过上述操作得到池化特征图  $C_{2m}$ 。

[0028] ④重复②和③。依次得到特征图  $C_{3s}$ 、 $C_{3m}$ 、 $S_{4s}$  和  $S_{4m}$ 。

[0029] ⑤将  $S_{4s}$  和  $S_{4m}$  合并，形成图像的特征，经过全连接层之后送入 softmax 分类器，得到分类结果  $O_t$ 。

[0030] (2) 反向传播阶段

[0031] ①计算实际输出  $O_t$  与相应的理想输出  $y_t$  的差；

[0032] ②计算整体的代价函数为

$$[0033] \quad J(W, b) = \frac{1}{m} \sum_{i=1}^m \left( \frac{1}{2} \| O_t^{(i)} - y_t^{(i)} \|^2 \right) \quad (3)$$

[0034] 反向传播，对于第  $n_1$  层（输出层）的每个输出单元  $i$ ，我们根据以下公式计算残差：

$$[0035] \quad \delta_i^{(n_l)} = -\left(y_i - a_i^{(n_l)}\right) \cdot f' \left(z_i^{(n_l)}\right) \quad (4)$$

[0036] 对  $l = n_1-1, n_1-2, n_1-3, \dots, 2$  的各个层, 第  $l$  层的第  $i$  个节点的残差计算方法如下:

$$[0037] \quad \delta_i^{(l)} = \left(\sum_{j=1}^{s_{l+1}} W_{ji}^{(l)} \delta_j^{(l+1)}\right) \cdot f' \left(z_i^{(l)}\right) \quad (5)$$

[0038] 计算偏导数, 计算方法如下:

$$[0039] \quad \nabla_{W^{(l)}} J(W, b; x, y) = \delta^{(l+1)} \left(a^{(l)}\right)^T \quad (6)$$

$$[0040] \quad \nabla_{b^{(l)}} J(W, b; x, y) = \delta^{(l+1)} \quad (7)$$

$$[0041] \quad \Delta W^{(l)} := \Delta W^{(l)} + \nabla_{W^{(l)}} J(W, b; x, y) \quad (8)$$

$$[0042] \quad \Delta b^{(l)} := \Delta b^{(l)} + \nabla_{b^{(l)}} J(W, b; x, y) \quad (9)$$

[0043] 更新权重参数:

$$[0044] \quad W^{(l)} = W^{(l)} - \alpha \left(\frac{1}{m} \Delta W^{(l)}\right) \quad (10)$$

$$[0045] \quad b^{(l)} = b^{(l)} - \alpha \left[\frac{1}{m} \Delta b^{(l)}\right] \quad (11)$$

[0046] 至此 1 次迭代完成, 保存调整后的权值, 重复第一阶段, 直至达到设定的迭代次数后, 训练结束, 保存并把此权值输出给验证集以验证。

[0047] 3. 准确率验证阶段

[0048] 使用验证集验证输入的权值, 将输入和映射的结果进行对比, 输出错误率, 如果错误率达到满意的要求, 则可以进行测试识别, 否则, 调整网络规模或者增加迭代次数重新进行训练。

[0049] 4. 图像识别阶段

[0050] 将检测到的交通标志图像经过预处理之后, 使用训练好权值的网络模型进行识别并输出其对应的类别。

[0051] 本发明的方法采用两个不同结构的卷积神经网络并行地进行特征映射和提取, 最后将特征进行合并, 再经过全连接层和最后的分类器, 完成整个分类过程。两个不同结构的卷积神经网络分别采用了随机池化操作和 maxout 单元, 确保了图像特征的多样性, 提高了识别精度并加快了网络运算速度。

[0052] 本发明对传统卷积神经网络进行了结构上的改进, 使用两个不同结构的卷积神经网络代替了传统的卷积神经网络结构, 该方法确保了图像特征的多样性, 提高了识别精度并加快了网络运算速度。

## 附图说明

[0053] 图 1 是预处理后的图片。

[0054] 图 2 是非对称卷积神经网络示意图;

- [0055] 图 3 是非对称卷积神经网络的训练流程图；
- [0056] 图 4 是应用于交通标志中的非对称卷积神经网络示意图；
- [0057] 图 5 是错误率随循环次数变化曲线图。

### 具体实施方式

[0058] 1、确定训练集,本发明选用的是 GTSRB(德国交通标志识别基准, German traffic sign recognition benchmark) 中的训练集,包含训练图片 39,209 张,测试图片 12630 张。

[0059] 2、对训练集中的图片进行预处理,取原始图像中的目标区域(交通标志区域)取出,将彩色图像转化为灰度图像,并缩放到  $48 \times 48$  大小,再将目标区域经过 histeq 图像对比度增强处理,得到原始的训练集,测试集采用同样的方式处理。为使所训练的模型有更好的鲁棒性,将原始训练集经过旋转  $[-10^\circ, 10^\circ]$ , 缩放  $[0.9, 1.1]$  后,加入原始数据集中,组成新的训练集,在新的数据集中随机取出与测试集数目相当的样本组成验证集,剩余的样本组成最终的训练集。

[0060] 3、初始化参数,将各层的权值随机初始化为接近于 0 的随机值,并设定随机梯度下降法学习率为 0.01,循环次数为 100,确定卷积核大小为  $5 \times 5$ 。

[0061] 4、选定训练组;从训练集中每次分别随机地选取 50 个样本作为训练组。

[0062] 5、从训练集中取出训练样本  $(x_t, y_t)$ , 将  $x_t$  作为两个独立的卷积神经网络的输入,  $y_t$  作为前向传播结果的比对标准。

[0063] 6、第一层卷积层进行卷积操作,使用  $5 \times 5$  的卷积核对输入图像进行卷积,步长为 1,生成 6 幅卷积特征图,每个卷积神经网络 3 幅卷积特征图,如图 4 所示,两个网络的每幅特征图的大小均为  $44 \times 44$ 。

[0064] 7、采样层进行采样操作,采用  $2 \times 2$  采样尺寸,两个卷积神经网络特征图大小均变为  $22 \times 22$ 。

[0065] 8、再次进行卷积操作和采样操作,卷积特征图变为 12 幅,特征图大小为  $9 \times 9$ 。

[0066] 9、将两个卷积神经网络得到的特征合并为一个特征,为 972 维的向量,然后送入有 1000 个神经结点的全连接层,最后送入 43 个输出结点的 softmax 分类器,得到最终分类结果。

[0067] 10、经过反向传播调整参数,直到验证集正确率达到可接受的范围后,保存参数,用于测试测试集正确率。

[0068] 11、训练中,迭代次数往往会对最终的准确率产生很大影响,如附图 5,因此需要达到一定迭代训练次数才可获得较高的准确率。



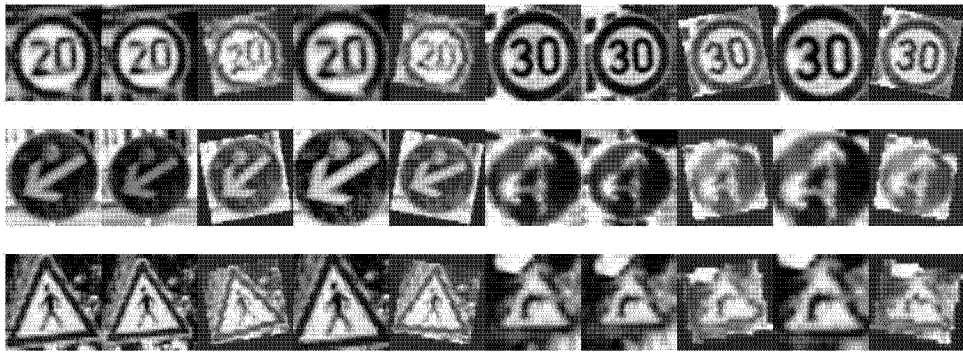


图 1

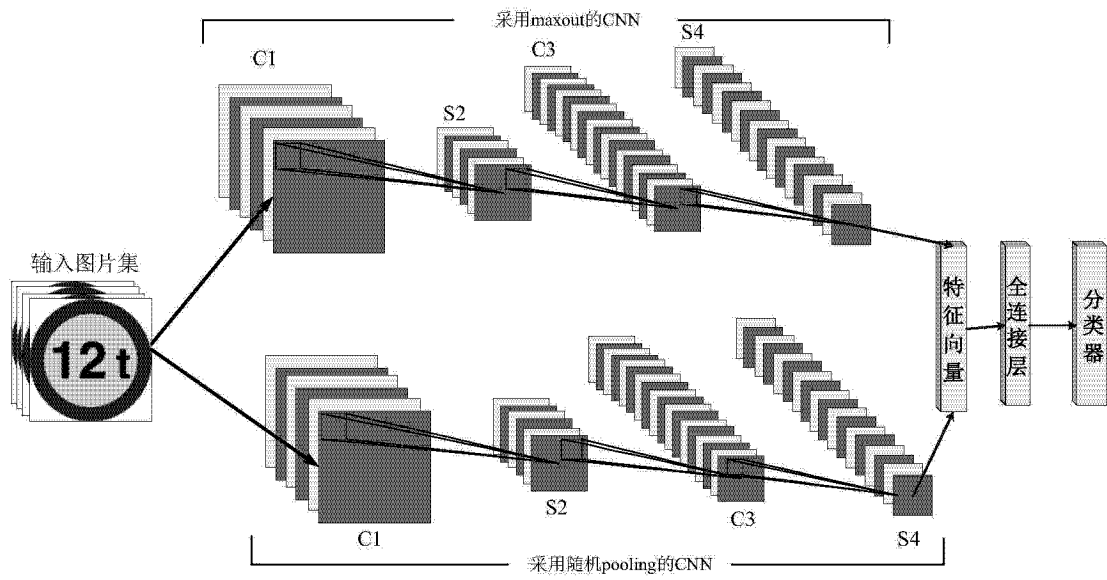


图 2

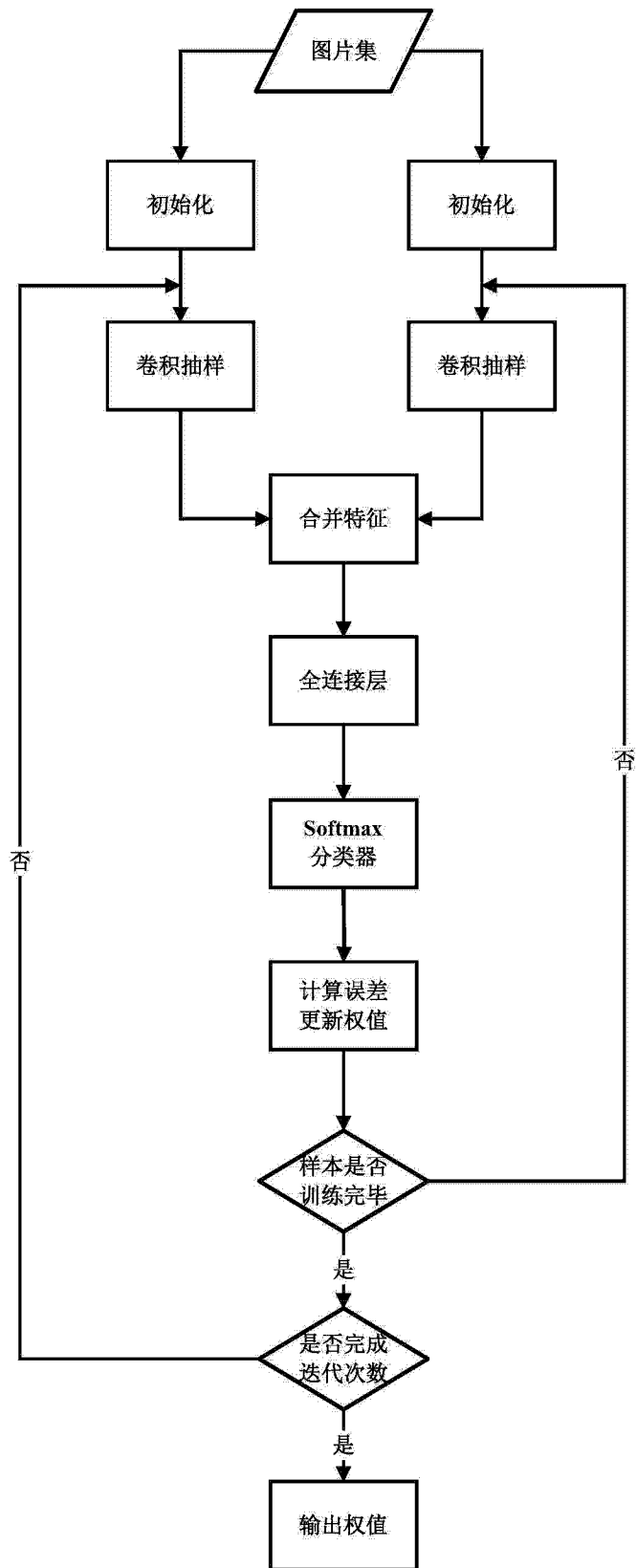


图 3

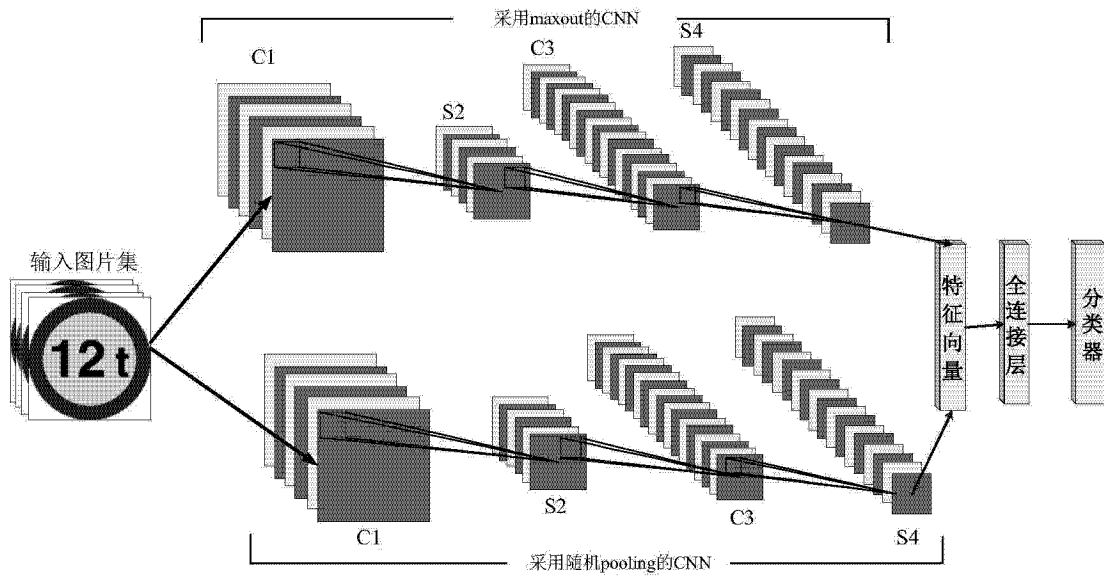


图 4

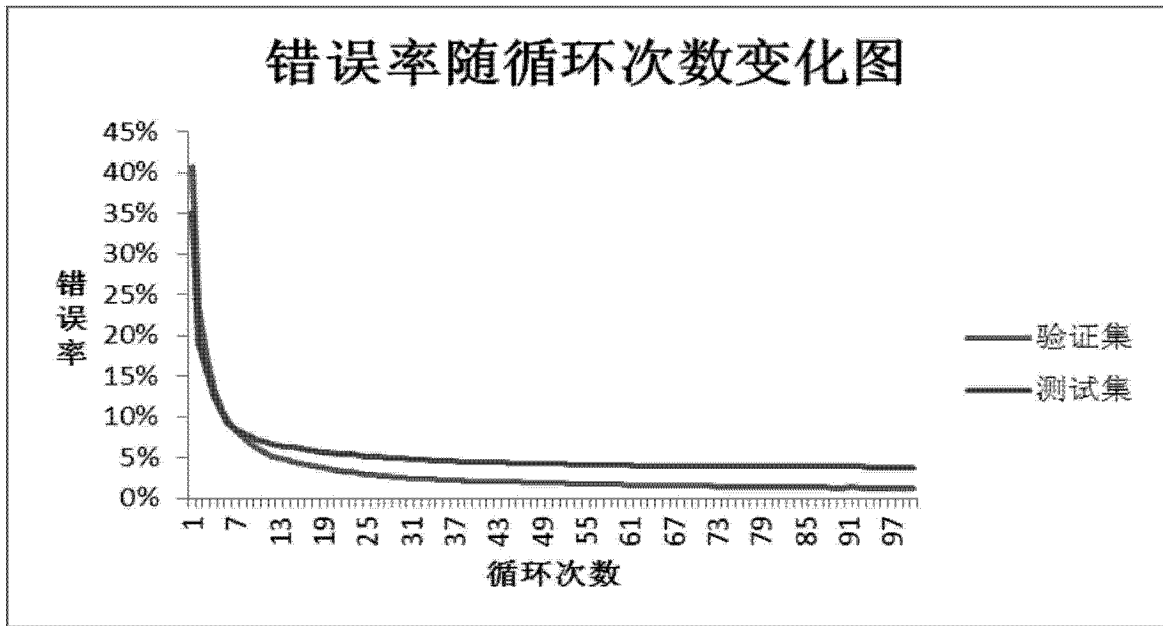


图 5