(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 112613552 B
(45)授权公告日 2024.05.28

(21)申请号 202011506810.6

(22)申请日 2020.12.18

(65)同一申请的已公布的文献号
申请公布号 CN 112613552 A

(43)申请公布日 2021.04.06

(73)专利权人 北京工业大学
地址 100124 北京市朝阳区平乐园100号

(72)发明人 毋立芳 邓斯诺 张恒 石戈 简萌 相叶

(74)专利代理机构 北京思海天达知识产权代理有限公司 11203
专利代理师 刘萍

(51)Int.Cl.
G06V 10/764(2022.01)
G06V 10/82(2022.01)
G06N 3/0464(2023.01)
G06N 3/08(2023.01)

(56)对比文件
CN 107341506 A,2017.11.10
CN 107506722 A,2017.12.22
CN 110135461 A,2019.08.16
CN 111667559 A,2020.09.15
RU 2707147 C1,2019.11.22
傅博文;唐向宏;肖涛.Focal损失在图像情感分析上的应用研究.计算机工程与应用.(10),全文.

审查员 周瑜

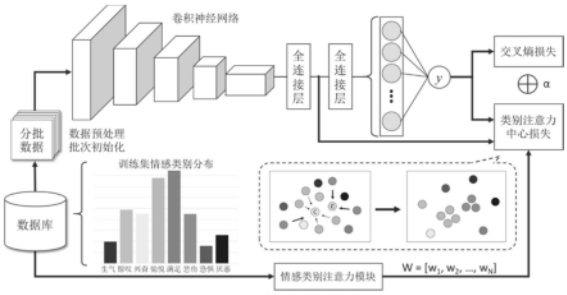
权利要求书2页 说明书5页 附图1页

(54)发明名称

一种结合情感类别注意力损失的卷积神经网络情感图像分类方法

(57)摘要

一种结合情感类别注意力损失的卷积神经网络情感图像分类方法,涉及智能媒体计算和计算机视觉技术领域;首先对训练样本进行类别权重计算,得到情感类别注意力权重向量;其次按情感类别数量和情感类别注意力损失修改卷积神经网络最后的分类层与损失函数;然后将训练样本进行预处理后传入网络中,使网络在损失函数和优化器的对参数的迭代更新后达到收敛,完成训练;最终将预处理后的测试图像送入网络中,计算得到的模型的情感图像分类准确率以及模型对测试情感图像的预测类别。本发明使得情感图像在通过卷积神经网络进行情感类别分类时,可以自适应的得到更符合数据集样本分布特点的分类结果,有助于情感分类算法在不同的实际应用场景中训练与使用。



CN 112613552 B

1.结合情感类别注意力损失的卷积神经网络情感图像分类方法,其特征在于,包含如下步骤:

步骤1,建立图像的情感类别权重向量:将数据集中每一类的图像视为一维,得到情感类别权重向量W;

步骤2,建立深度网络模型:选择深度网络模型ResNet-101,使用ResNet-101为基础骨干网络,并以待分类类别数作为输出向量维度替换原模型最后的分类层,以生成情感类别;

步骤3,损失函数设计:为了利用图像的情感类别权重信息,解决异类样本比例不均衡和类间类内距离收束的问题,损失函数包括对类间距离和类内距离的计算以及不同情感类别权重比例的考量;

步骤4,训练模型:将图像经过预处理后,输入到网络模型中,使用随机梯度下降方法进行优化,对模型参数进行学习;

步骤5,获得待测图像的情感类别:将数据库中的图像经过同步骤4一样的预处理步骤后,输入到步骤4训练好的模型中,得到对应的情感类别;

在所述步骤1中,建立图像的情感类别权重的具体方法为:将情感类别数为N的图像训练集中的同一类别图像视为一个维度C_i,i=1,2,…,N;计算统计不同维度中样本数量n_i,i=1,2,…,N;通过计算获得各类别权重,合并形成权重向量W=[w_1,w_2,…,w_N],各w_i计算方式如下:

$$C_i: \qquad w_i = \frac{n_i N}{\sum_{i=1}^{N} n_i};$$

在步骤3中,损失函数包含两项损失,情感类别注意力中心损失和交叉熵损失;

4.1情感类别注意力中心损失

情感类别注意力中心损失特别增加了权重部分来降低由于样本类别不均衡带来的影响;具体损失函数如下:

$$L_c = \frac{1}{2\mathrm{m}} \sum_{i=1}^{\mathrm{m}} (\mathrm{W} \| f_i - C_{y_i} \|_2)^2$$

其中m为训练时每批次的图像数量,W为得到的权重向量,f_i为图像经过基础骨干网络在分类层之前得到的特征,C_{y_i}为类别为y_i时的特征中心,特征中心即为当前批次下同类别图像得到的特征在各个维度中的均值组成的特征向量;

4.2交叉熵损失

使用交叉熵损失来进行基本的损失度量,其目的在于保持类间距离;构建的交叉熵损失用于让不同情感类别之间的图像更远;具体损失函数如下:

$$L_s = -\frac{1}{\mathrm{m}} \sum_{i=1}^{\mathrm{m}} \log \frac{e^{w_{y_i}^{\mathrm{T}} x_i + b_{y_i}}}{\sum_{j=1}^{\mathrm{N}} e^{w_j^{\mathrm{T}} x_i + b_j}}$$

其中m为训练时每批次的图像数量,N为数据集情感类别的数量,x_i为该批次中第i张图片从ResNet-101基础骨干网络在分类层之前得到的特征,w和b即为分类层中的权重和偏置

参数值,下标y_i和j表示为分类层之后类别,b_{y_i}表示该批次第i张图片判断为y_i类别时在分类层中的偏置参数的值;

总损失函数为:

$$L=(1-\alpha)L_c+\alpha L_s$$

$$=(1-\alpha)\frac{1}{2m}\sum_{i=1}^{m}\left(W\left\|f_i-C_{y_i}\right\|_2\right)^2-\alpha\frac{1}{m}\sum_{i=1}^{m}\log\frac{e^{w_{y_i}^T x_i+b_{y_i}}}{\sum_{j=1}^{N}e^{w_j^T x_i+b_j}}$$

其中,α为量化损失函数的超参数,设置为0.6;

将待训练图片进行随机中心裁剪和随机水平翻转的方式进行数据预处理,利用随机梯度下降的方式对模型进行训练,保存训练好的模型。

一种结合情感类别注意力损失的卷积神经网络情感图像分类方法

技术领域

[0001] 本发明属于计算机视觉技术领域,涉及一种结合情感类别注意力损失的卷积神经网络情感图像分类方法。

背景技术

[0002] 随着社交媒体的发展,人们越来越倾向于通过在网络上发布图像信息来记录和分享自己的心情。这些蕴含着情感信息的图片,往往也包含着发布者对某一类事物的情感趋向和态度。而通过从海量图像中获知人们的情感趋向来了解用户群体的态度,对商品推荐、舆情分析以及社交媒体管理都有着重要的作用。所以如何利用计算机算法进行高效自动化的对大量蕴含情感信息的图像进行识别和分析是现在亟需解决的问题。

[0003] 早前的情感分析方法采取使用底层手工特征用于情感分类,如颜色、线条、纹理等特征,或者利用名词形容词对探测器Sentibank提取图像中的中层表征。得益于深度学习强大的特征提取能力,使用基于卷积神经网络的算法在图像分类任务中取得的效果越来越好,情感图像分类任务也因此有了新的进展。如You等人在2016年设计了一种深度卷积神经网络来进行图像情感分类,在其反馈机制中加入了滤除功能,通过该功能滤除了训练集中错误的标注数据,从而有效地提高了图像情感分类的能力。She等人在2019年设计了一种基于深度卷积神经网络ResNet101的弱监督耦合卷积网络进行图像情感分类,通过类激活图捕捉引起情感的区域,并且利用误差反向传播机制进行反馈调整,从而进一步提升图像情感分类的准确性。

[0004] 但在通常的图像分类任务中,不同类别的图像区别相对明确,而不同情感图像在情感类别之间,却没有明确清晰的界限。这就需要通过对样本特征的相似度距离采取不同的度量方式计算损失函数,来提升不同类别的区分度。而在这方面Zhang等人于2017年提出了一种将深度卷积神经网络与中心损失相结合的人脸识别技术,通过将传统损失函数与中心损失函数两种损失函数共同作为迁移学习过程中的监督信号,使得提取到的特征在同类别时聚合在不同类时分散,以此来提高模型输出人脸特征的判别能力。Yang等人在2018年提出了一种使用三重约束来训练卷积神经网络的方法有效地将图像定位在情感级别,通过考虑不同情感极性的关系,计算特征间的相关性来实现情感图像检索和分类的多任务处理。尽管这些研究考虑了收束类内距离,但没有针对性的对样本数量不均衡的情感类别进行损失函数设计,这在真实的社交媒体环境中还是不够的。在真实的社交媒体环境中,情感图像的数量无法做到按类别均匀分布。这就会使得通常训练出的模型不能有针对性的考虑各个情感类别,而造成分类性能的损失。

[0005] 在这样的启发下,我们设计了一种结合了交叉熵损失与情感类别注意力中心损失的方法,扩大情感图像类间距离的同时也按不同类别有差异的收束类内距离,从而提升情感图像分类的准确率。

发明内容

[0006] 为了解决以上问题,我们发明了一种情感图像分类算法,将情感类别注意力中心损失与交叉熵损失相结合,使得不同类样本特征距离更远的同时,同类别样本间的距离更近。减少了部分样本由于缺少类内约束而分布在与类内中心距离更远的类间交界处的机会,从而减少了这些样本被错分的可能。同时针对不同样本类别采取不同的收束力度,使得当各类别的样本在数量不均衡时,可以得到相应的调整,从而使得通过本方法得到的网络模型具有更好的情感图像分类效果。

[0007] 本发明的具体步骤如下:

[0008] 步骤1,建立图像的情感类别权重向量:将有标注的情感图像数据集进行整理划分,将训练集部分中每一类的图像视为一维,得到情感类别权重向量W;

[0009] 步骤2,建立深度网络模型:选择深度网络模型,如ResNet-101,并以待分类类别数作为输出向量维度替换原模型最后的分类层;

[0010] 步骤3,损失函数设计:为了利用图像的情感类别权重信息,解决异类样本比例不均衡和类间类内距离收束的问题,损失函数包括对类间距离和类内距离的计算以及不同情感类别权重比例的考量;

[0011] 步骤4,训练模型:将步骤1划分的图像经过放缩随机翻转等方式进行预处理后,输入到步骤2的网络模型中,使用随机梯度下降方法进行优化,通过步骤3的损失函数计算损失对模型参数进行学习;

[0012] 步骤5,获得待测图像的情感类别:将数据集中的图像经过固定尺寸放缩和中心裁剪的预处理步骤后,输入到步骤4训练好的模型中,得到对应的情感类别。

[0013] 本发明与现有技术相比较,具有如下显而易见的突出实质性特点和显著技术进步:

[0014] 本发明提出了一种结合情感类别注意力损失的卷积神经网络情感图像分类方法。通过计算交叉熵损失和类别注意力中心损失相结合的方式反馈更新卷积神经网络的参数,使得不同类样本特征距离更远的同时,同类别样本间的距离更近。减少了部分样本由于缺少类内约束而分布在与类内中心距离更远的类间交界处的机会,从而减少了这些样本被错分的可能。同时针对不同样本类别采取不同的收束力度,使得当各类别的样本数量不均衡时,可以得到相应的调整,从而使得通过本方法得到的网络模型在类别不均衡的情感图像数据集的测试中,具有更好的情感分类效果。

附图说明

[0015] 下面结合附图和具体实施方式对本发明作进一步详细的说明:

[0016] 图1为训练基于本方法的图像情感分类卷积神经网络的架构图。

[0017] 图2为基于本方法进行情感图像分类的整体流程图。

具体实施方式

[0018] 本发明提出了一种结合情感类别注意力损失的卷积神经网络情感图像分类方法。本发明的整体结构如图1所示。本发明实例在win10和JupyterNotebook环境下进行仿真,使用FI数据集通过本发明中的方法进行训练,得到可实现高准确率的图像情感分类模型。得

到模型后,可将测试图像输入到模型中,得到图像的情感分类结果。本发明的具体实现流程如图2所示,具体实现步骤如下:

[0019] 步骤1:建立图像的情感类别权重向量:将有标注的情感图像数据集进行整理划分,将训练集部分中每一类的图像视为一维,得到情感类别权重向量W;

[0020] 步骤2:建立深度网络模型:选择ResNet-101作为深度网络模型的骨干网络,并以待分类类别数作为输出向量维度替换原模型最后的分类层,得到待使用的深度网络模型;

[0021] 步骤3:损失函数设计:为了利用图像的情感类别权重信息,解决异类样本比例不均衡和类间类内距离收束的问题,损失函数包括对类间距离和类内距离的计算以及不同情感类别权重比例的考量;

[0022] 步骤4:训练模型:将步骤1划分的图像经过放缩随机翻转等方式进行预处理后,输入到步骤2的网络模型中,使用随机梯度下降方法进行优化,通过步骤3的损失函数计算损失对模型参数进行学习;

[0023] 步骤5:获得待测图像的情感类别:将数据集中的图像经过固定尺寸放缩和中心裁剪的预处理步骤后,输入到步骤4训练好的模型中,得到对应的情感类别。

[0024] 在所述步骤1中,建立了图像的情感类别权重向量W:

[0025] 本方法可以用于大型真实社交网络中的图像进行情感分类,故在本实例中选取自Flickr和Instagram整理而成的通用公开情感数据集Flickr&Instagram(下文简称FI数据集),相较传统情感数据集,该数据集具有数据规模大,情感类别不均衡的特点,更符合真实网络环境。

[0026] 将FI数据集中的训练集按原8类标签整理,通过权重计算得到各类别的权重w_i,计算方式如下:

[0027] $$w_i = \frac{\sum_{i=1}^{N} n_i}{n_i N}$$

[0028] 其中,N为类别数,n_i为各类别的训练样本数,在本实例中按FI数据集的情感类别为8类,设定N=8。得到各类别的权重系数后,将每类的权重系数并接在一起,得到情感类别权重向量$W=[w_1, w_2, \cdots, w_N]$;

[0029] 在所述步骤2中,建立了深度网络模型:

[0030] 本实例采用ResNet-101,获取原先在ImageNet上预训练好的ResNet-101模型,加载后去除最后的输入维度为2048*1024,输出维度为1*1024的分类层,并以待分类类别数8,作为输出向量维度替换,最终的分类层输入维度为2048*8,输出维度为1*8,即各类别预测概率,取最大项所属位置的对应情感类别即为图像输出的情感类别;

[0031] 在所述步骤3中,进行了损失函数的设计:

[0032] 本实例为了利用图像的情感类别权重信息,解决异类样本比例不均衡和类间类内距离收束的问题,损失函数包括对类间距离和类内距离的计算以及不同情感类别权重比例的考量,设计的损失函数包含情感类别注意力中心损失和交叉熵损失两项:

[0033] 情感类别注意力中心损失:使用增加情感类别权重的中心损失来进行约束,以实现根据情感类别注意力控制的机制。构建加权中心损失用以有区分性的针对不同类别予以类内收束。为了解决数据的类别不均衡问题,加权中心损失特地的增加了权重部分来降低

由于样本类别不均衡带来的影响。具体损失函数如下：

[0034] $$L_c = \frac{1}{2m}\sum_{i=1}^{m}(W\|f_i - C_{y_i}\|_2)^2$$

[0035] 其中m为训练时每批次的图像数量，一般根据实验平台显存大小取值为2的整数次幂，如16或32等，本实例中设置为16，W为步骤1中得到的权重向量，f_i为图像经过基础步骤2中的基础骨干网络在分类层之前得到的特征，C_{y_i}为类别为y_i时的特征中心，特征中心即为当前批次下同类别图像得到的特征在各个维度中的均值组成的特征向量。

[0036] 交叉熵损失：使用交叉熵损失来进行基本的损失度量，其目的在于保持类间距离。构建的交叉熵损失用于让不同情感类别之间的图像更远。具体损失函数如下：

[0037] $$L_s = -\frac{1}{m}\sum_{i=1}^{m}\log\frac{e^{w_{y_i}^T x_i + b_{y_i}}}{\sum_{j=1}^{N} e^{w_j^T x_i + b_j}}$$

[0038] 其中m为训练时每批次的图像数量与上述L_c中的m相同，N为数据集情感类别的数量，本实例中根据FI数据集情况N＝8，x_i为该批次中第i张图片从步骤2中的基础骨干网络在分类层之前得到的特征，w和b即为分类层中的权重(weight)和偏置(bias)参数值，下标y_i和j表示为分类层之后类别，例如其中b_{y_i}表示该批次第i张图片判断为y_i类别时在分类层中的偏置参数的值。

[0039] 总损失函数为：

$$L = (1-\alpha)L_c + \alpha L_s$$

[0040] $$= (1-\alpha)\frac{1}{2m}\sum_{i=1}^{m}\left(W\|f_i - C_{y_i}\|_2\right)^2 - \alpha\frac{1}{m}\sum_{i=1}^{m}\log\frac{e^{w_{y_i}^T x_i + b_{y_i}}}{\sum_{j=1}^{N} e^{w_j^T x_i + b_j}}$$

[0041] 其中，α为量化损失函数的超参数，在本实例中设置为0.6。

[0042] 在所述步骤4中，进行模型的训练：

[0043] 将图像经过放缩随机翻转等方式对训练集数据进行预处理，在本实例中随机裁剪的参数设置为448*448，随机翻转的概率设置为0.5。按固定大小为一批次，输入到网络模型中，最后不足固定大小批次的样本作为一个批次。固定批次大小设置在一定程度上尽量的大会小幅度的提高模型的训练效果，但由于实验平台限制，推荐选取8、16或32，在本实例中固定批次的大小设置为16。通过最后的分类层，得到输出结果与输入的训练集标签进行自动比对，统计正确样本数量占整体训练样本的比例记为训练集在此轮的准确率。同时在得到输出向量时，按步骤3中计算的损失函数也取得了当前轮次的损失值，得到的损失值将反馈给优化器处理后进行反向传播更新模型中各个参数。

[0044] 出于对收敛速度和收敛效果的考虑，本方法中的优化器选取随机梯度下降法作为优化方法。优化器的参数设置主要有初始学习率(learning rate)和动量(momentum)两项，其中初始学习率一般为0.1、0.01、0.0001、0.00001等值中视模型收敛情况选取，本实例推荐取0.0001，在此初始值时收敛效果更加稳定。动量原则上在0到1之间取值，本实例中推荐

选取随机梯度下降法中的默认值0.9。由于固定学习率的设置不利于深度网络在训练后半程中寻找更好的参数,本方法在训练过程中增加固定轮次降低学习率的策略。其中降低的轮次推荐在14-20轮降低1-2次,总训练轮数建议50-80轮,直至损失函数值震荡,不再明显下降,即可认定模型已收敛,训练可以结束。在本实例中对优化器进行了每14轮和每20轮降低一半学习率的设置,并对模型参数进行60轮训练学习以保证训练效果的有效收敛,设置轮数过少可能还未收敛,轮数过多则会增加训练时间但不会提高效果。

[0045] 在每轮在训练样本训练完后,固定模型的参数并采用FI数据集中的验证集数据固定大小放缩裁剪传入到网络模型中,本实例中裁剪参数设置为448*448,将模型输出与样本的标签进行对比并统计正确样本所占比例即验证集准确率,如果当前轮数的验证集准确率比先前最高的验证集准确率要高就保存当前准确率为最高的验证集准确率,并保存当前轮数训练的模型。当所有轮次训练完毕后,最终保存的最高验证集准确率下的模型,即为训练好的最优模型;

[0046] 在所述步骤5中,获得待测图像的情感类别:

[0047] 将FI数据集中的测试集数据或任意图像按同步骤4中验证集图像一样的固定大小放缩中心裁剪的预处理步骤后可逐张或以固定数量按批次输入模型即可。本实例中固定大小放缩中心裁剪的参数设置为448*448,为在同等实验条件下提高处理效率,本实例中测试集数据推荐以16为批次大小,按批输出测试图像到模型中进行测试。经过模型处理,将分类层后的输出结果与样本的标签进行对比并统计正确样本所占比例即测试集准确率。输出结果对应的情感类别即为改模型判断的图像情感类别。

[0048] 在FI数据集中的测试集经过本实例中的模型测试,准确率结果为0.7087,高于目前同类方法研究内容中的最好效果:刊登在今年高水平杂志IEEE TRANSACTIONS ON MULTIMEDIA上的Adaptive Deep Metric Learning for Affective Image Retrieval and Classification的0.6837,同时也高于目前在此数据集同样仅以标签信息作为参考依据的多分类最好结果:发表在今年高水平杂志IEEE TRANSACTIONS ON MULTIMEDIA上的WSCNet: Weakly Supervised Coupled Networks for Visual Sentiment Classification and Detection的准确率0.7007。

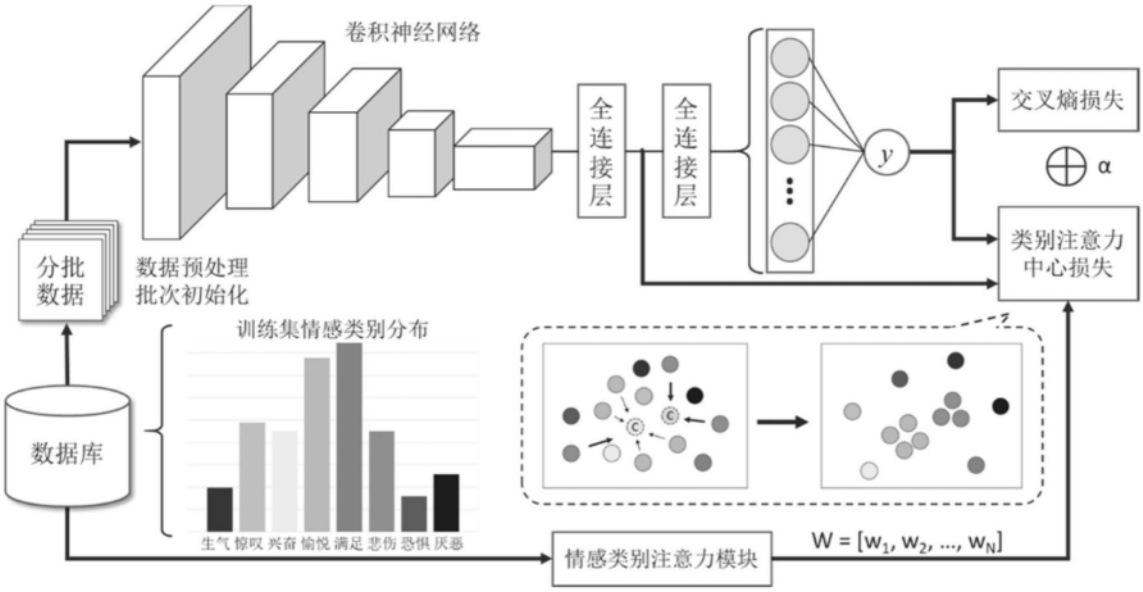

卷积神经网络
全连接层
全连接层
y
交叉熵损失
α
类别注意力中心损失
分批数据
数据预处理
批次初始化
训练集情感类别分布
数据库
生气 惊叹 兴奋 愉悦 满足 悲伤 恐惧 厌恶
情感类别注意力模块
W = [w1, w2, ..., wN]

图1

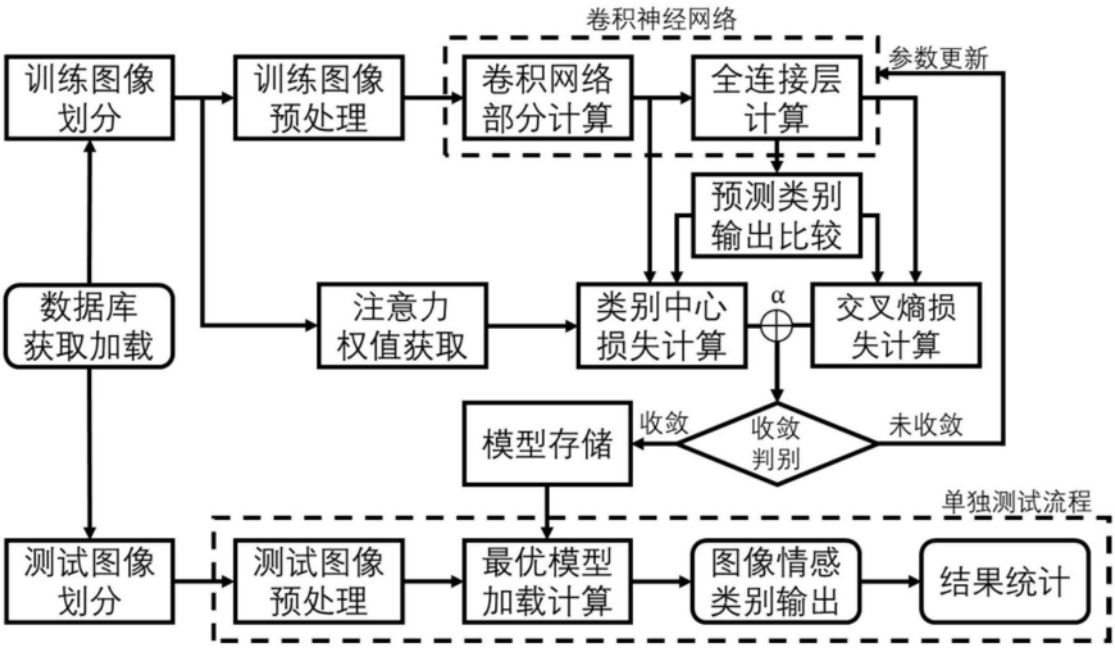

卷积神经网络
参数更新
训练图像划分
训练图像预处理
卷积网络部分计算
全连接层计算
预测类别输出比较
数据库获取加载
注意力权值获取
类别中心损失计算
α
交叉熵损失计算
模型存储
收敛
收敛判别
未收敛
单独测试流程
测试图像划分
测试图像预处理
最优模型加载计算
图像情感类别输出
结果统计

图2