



(12) 发明专利申请

(10) 申请公布号 CN 113807164 A

(43) 申请公布日 2021. 12. 17

(21) 申请号 202110863986.5

(22) 申请日 2021.07.29

(71) 申请人 四川天翼网络服务有限公司  
地址 610041 四川省成都市高新区九兴大道10号三楼

(72) 发明人 刘栓

(74) 专利代理机构 成都金英专利代理事务所  
(普通合伙) 51218

代理人 袁英

(51) Int. Cl.

G06K 9/00 (2006.01)

G06K 9/62 (2006.01)

G06N 3/04 (2006.01)

G06N 3/08 (2006.01)

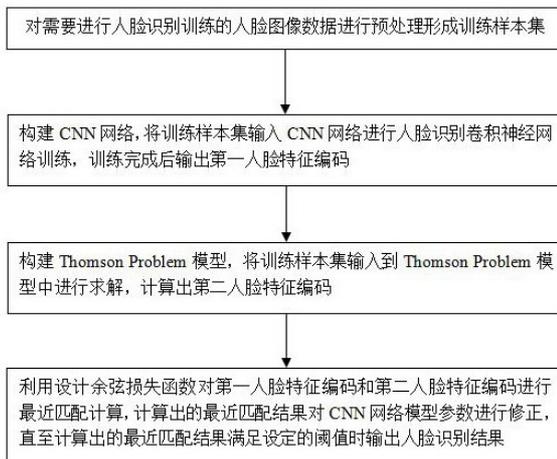
权利要求书1页 说明书5页 附图1页

(54) 发明名称

一种基于余弦损失函数的人脸识别方法

(57) 摘要

本发明公开了一种基于余弦损失函数的人脸识别方法,包括以下步骤:对需要进行人脸识别训练的人脸图像数据进行预处理形成训练样本集;构建CNN网络,将训练样本集输入CNN网络进行人脸识别卷积神经网络训练,训练完成后输出第一人臉特征编码;将训练样本集输入到Thomson Problem模型中进行求解,计算出第二人脸特征编码;对第一人臉特征编码和第二人脸特征编码进行最近匹配计算,根据最近匹配结果对CNN网络模型参数进行修正,直至最近匹配结果满足设定的阈值时输出人脸识别结果。本发明在人脸识别过程采用人脸特征的编码向量代替原来的FC,使得神经网络的训练不受训练数据集大小的影响,提高了人脸识别效率。



1. 一种基于余弦损失函数的人脸识别方法,其特征在於,包括以下步骤:

步骤一:对需要进行人脸识别训练的人脸图像数据进行预处理形成训练样本集;

步骤二:构建CNN网络,将训练样本集输入CNN网络进行人脸识别卷积神经网络训练,训练完成后输出第一人臉特征编码;

步骤三:构建Thomson Problem模型,将训练样本集输入到Thomson Problem模型中进行求解,计算出第二人脸特征编码;

步骤四:设计余弦损失函数,并利用余弦损失函数对第一人臉特征编码和第二人脸特征编码进行最近匹配计算,根据计算出的最近匹配结果对CNN网络模型参数进行修正,直至计算出的最近匹配结果满足设定的阈值时输出人脸识别结果。

2. 根据权利要求1所述的一种基于余弦损失函数的人脸识别方法,其特征在於,所述步骤三具体包括:设定Thomson Problem模型中的人脸特征编码维度 $m$ ,将训练样本集导入Thomson Problem模型进行人脸特征编码求解,将输出的所有人脸特征编码记为矩阵 $GT\_Maix$ ,即第二人脸特征编码。

3. 根据权利要求1所述的一种基于余弦损失函数的人脸识别方法,其特征在於,所述步骤四中设计的余弦损失函数如下式所示:

其中, $Lns$ 为交叉熵损失函数; $\theta_{y_i,i}$ 表示真实标签 $y$ 向量与预测标签 $i$ 向量的夹角。

4. 根据权利要求1所述的一种基于余弦损失函数的人脸识别方法,其特征在於,所述步骤四中利用余弦损失函数对第一人臉特征编码和第二人脸特征编码进行最近匹配计算过程包括:分别将第一人臉特征编码与第二人脸特征编码输入设计的余弦损失函数中,分别计算出第一人臉特征编码的夹角 $\theta_1$ 与第二人脸特征编码的夹角 $\theta_2$ ;根据最近匹配原则让夹角 $\theta_1$ 与夹角 $\theta_2$ 自动匹配,获得第一人臉特征编码和第二人脸特征编码之间的误差;根据余弦损失函数的导数,沿梯度最小方向将误差回传,修正CNN网络中的各层权重值,并重复进行人脸识别卷积神经网络训练,直至计算出的最近匹配结果满足设定的阈值时输出人脸识别结果。

5. 根据权利要求1所述的一种基于余弦损失函数的人脸识别方法,其特征在於,所述步骤二中CNN网络的网络结构包括:输入层、卷积层、池化层和输出层,输入层与卷积层连接,卷积层与池化层连接,池化层和输出层连接。

6. 根据权利要求4所述的一种基于余弦损失函数的人脸识别方法,其特征在於,所述步骤四中的夹角 $\theta_1$ 与夹角 $\theta_2$ 自动匹配过程还包括采用匈牙利算法进行夹角 $\theta_1$ 与夹角 $\theta_2$ 的自动匹配,获得第一人臉特征编码和第二人脸特征编码之间的误差。

## 一种基于余弦损失函数的人脸识别方法

### 技术领域

[0001] 本发明涉及人脸识别技术领域,尤其涉及一种基于余弦损失函数的人脸识别方法。

### 背景技术

[0002] 随着大数据的加持,人脸识别训练时可用的数据越来越多,从几万人到几十万人再到几百万人。在数据增多的同时给我们带来了新的问题,大数据集很难在当前主流的Loss函数上面进行训练(如Softmax系列),因为该类损失函数要求在训练时最后一层的全连接层(FC)神经元的个数要与数据集中的人数对应,当训练人数在几十万人时,最后一层的参数量比前面的参数量往往会高 10倍左右,最后一层可能就要占用将近1G的显存,这样一张12G显存的显卡训练时输入BatchSize也只能设置16左右(正常训练256是个不错的值),而且训练的大部分时间都是花在最后一层,然而我们实际做推理的时候并不使用最后一层,只使用前面的特征提取层,所以当数据量增多时特征提取层并不能得到很好的训练。

[0003] 如申请号为CN201810866133.5的专利公开了一种基于新型损失函数的人脸识别卷积神经网络训练方法,该方法将SoftMax loss、余弦相似度损失、Center Loss相结合,共同作为人脸识别卷积神经网络训练过程中的目标函数,在保证增大类间距、减小类内距离的同时,加入余弦相似度损失来克服由于训练与测试时人脸特征比对的度量方法不同造成的差异性。包括:将人脸识别数据行划分为训练样本集与测试样本集;对训练样本集中的人脸图像进行数据预处理;构建基于新型损失函数的人脸识别卷积神经网络结构;将训练样本集输入到人脸识别卷积神经网络中进行训练;保存人脸识别模型参数;利用经过数据预处理后的测试样本集对人脸识别模型进行测试。该技术方案虽然通过将SoftMax loss、余弦相似度损失、Center Loss相结合作为目标函数,克服了训练和测试的差异性,但是该方法并不能适用于样本数据量大的人脸识别训练,训练效率不高。

[0004] 如申请号为CN202010188585.X的专利申请公开了一种基于余弦损失非约束条件人脸识别方法,它包括以下步骤,S1、获取待识别图像,对待识别图像进行多尺度变换,得到图像金字塔;S2、将步骤S1得到的图像金字塔输入到MTCNN网络,MTCNN网络对图像进行处理后,得到面部特征点;S3、根据步骤S2的面部特征点,进行人脸校正;S4、利用步骤S3处理后的数据训练 Inception-ResnetV1卷积神经网络,虽然该技术方案采用余弦损失函数作为监督信号,进行分类器模型的训练,是该方法也不能适用于样本数据量大的人脸识别训练,训练效率不高。

### 发明内容

[0005] 本发明的目的在于克服现有技术的不足,提供一种基于余弦损失函数的人脸识别方法,同时设计了新的Loss函数,该Loss函数的输入基于Thomson problem的输出结果,该结果的维度可以任意设置,不受训练数据集人数的影响。

[0006] 本发明的目的是通过以下技术方案来实现的:

[0007] 一种基于余弦损失函数的人脸识别方法,包括以下步骤:

[0008] 步骤一:对需要进行人脸识别训练的人脸图像数据进行预处理形成训练样本集;

[0009] 步骤二:构建CNN网络,将训练样本集输入CNN网络进行人脸识别卷积神经网络训练,训练完成后输出第一人臉特征编码;

[0010] 步骤三:构建Thomson Problem模型,将训练样本集输入到Thomson Problem 模型中进行求解,计算出第二人脸特征编码;

[0011] 步骤四:设计余弦损失函数,并利用余弦损失函数对第一人臉特征编码和第二人脸特征编码进行最近匹配计算,计算出的最近匹配结果对CNN网络模型参数进行修正,直至计算出的最近匹配结果满足设定的阈值时输出人脸识别结果。

[0012] 具体的,所述步骤三具体包括:设定Thomson Problem模型中的人脸特征编码维度m,将训练样本集导入Thomson Problem模型进行人脸特征编码求解,将输出的所有人臉特征编码记为矩阵GT\_Maix,即第二人脸特征编码。

[0013] 具体的,所述步骤四中设计的余弦损失函数如下式所示:

$$[0014] \quad Lns = \frac{1}{N} \sum_i -\log \frac{e^{s \cdot \cos(\theta_{y_i, i})}}{\sum_j e^{s \cdot \cos(\theta_{j_i})}}$$

[0015] Lns为交叉熵损失函数,其中 $\theta_{y_i, i}$ 表示真实标签y向量与预测标签i向量的夹角;

[0016] cos计算两个向量的夹角距离;s为缩放因子,默认取值64。

[0017] 具体的,所述步骤四中利用余弦损失函数对第一人臉特征编码和第二人脸特征编码进行最近匹配计算过程包括:分别将第一人臉特征编码与第二人脸特征编码输入设计的余弦损失函数中,分别计算出第一人臉特征编码的夹角 $\theta_1$ 与第二人脸特征编码的夹角 $\theta_2$ ;根据最近匹配原则让夹角 $\theta_1$ 与夹角 $\theta_2$ 自动匹配,获得第一人臉特征编码和第二人脸特征编码之间的误差;根据余弦损失函数的导数,沿梯度最小方向将误差回传,修正CNN网络中的各层权重值,并重复进行人脸识别卷积神经网络训练,直至计算出的最近匹配结果满足设定的阈值时输出人脸识别结果。

[0018] 具体的,所述步骤二中CNN网络的网络结构包括:输入层、卷积层、池化层和输出层,输入层与卷积层连接,卷积层与池化层连接,池化层和输出层连接。

[0019] 具体的,步骤四中的夹角 $\theta_1$ 与夹角 $\theta_2$ 自动匹配过程还包括采用匈牙利算法进行夹角 $\theta_1$ 与夹角 $\theta_2$ 的自动匹配,获得第一人臉特征编码和第二人脸特征编码之间的误差。

[0020] 本发明的有益效果:

[0021] 1.本发明设计的Loss函数的输入基于Thomsonproblem的输出结果,该结果的维度可以任意设置,通过在人脸识别过程采用人脸特征的编码向量代替原来的FC,使得神经网络的训练不受训练数据集大小的影响,提高了人脸识别效率。

[0022] 2.本发明在训练过程中让CNN网络生成的人脸特征与超球面求解的人脸特征尽可能匹配,这样相似不相同人脸在超球面上的表现也是挨着的特征向量,那么当我们在实际推理的时候,模型也会表现出对于相同的人脸特征向量夹角会很接近,对于相似不相同的人脸特征向量夹角是比较小的,对于不相同的人脸特征向量夹角是比较大的,通过两个人脸特征的匹配提高了人脸识别时的精度和准确性。

## 附图说明

- [0023] 图1是本发明的方法流程图。  
 [0024] 图2是本发明的CNN网络训练结构图。  
 [0025] 图3是本发明实施例的超球面模型示意图。

## 具体实施方式

[0026] 为了对本发明的技术特征、目的和效果有更加清楚的理解,下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅是本发明的一部分实施例,而不是全部的实施例。

[0027] 基于本发明中的实施例,本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0028] 本实施例中,如图1和图2所示,一种基于余弦损失函数的人脸识别方法,包括以下步骤:

[0029] 步骤一:对需要进行人脸识别训练的人脸图像数据进行预处理形成训练样本集;

[0030] 步骤二:构建CNN网络,将训练样本集输入CNN网络进行人脸识别卷积神经网络训练,训练完成后输出第一人臉特征编码;

[0031] 步骤三:构建Thomson Problem模型,将训练样本集输入到Thomson Problem 模型中进行求解,计算出第二人脸特征编码;

[0032] 步骤四:设计余弦损失函数,并利用余弦损失函数对第一人臉特征编码和第二人脸特征编码进行最近匹配计算,计算出的最近匹配结果对CNN网络模型参数进行修正,直至计算出的最近匹配结果满足设定的阈值时输出人脸识别结果。

[0033] 本实施例,由于现有的人脸识别方法在面对大数据集时的人脸识别效果不好,即当数据量增多时人脸识别过程的特征提取层并不能得到很好的训练,会影响到最终的人脸识别效率。

[0034] 本实施例中,为了解决上述问题,通过增加第二人脸特征编码来提高人脸识别效率。步骤三的过程具体包括:设定Thomson Problem模型中的人脸特征编码维度m,将训练样本集导入Thomson Problem模型进行人脸特征编码求解,将输出的所有人脸特征编码记为矩阵GT\_Maix,即第二人脸特征编码。

[0035] 具体的,步骤四中设计的余弦损失函数如下式所示:

$$[0036] \quad Lns = \frac{1}{N} \sum_i -\log \frac{e^{s \cdot \cos(\theta_{y_i, i})}}{\sum_j e^{s \cdot \cos(\theta_{j, i})}}$$

[0037] 其中,Lns为交叉熵损失函数; $\theta_{y_i, i}$ 表示真实标签y向量与预测标签i向量之间的夹角,cos函数用于计算两个y向量与i向量之间的夹角距离;s为缩放因子,默认取值64。

[0038] 本实施例中可以达到以下技术效果:

[0039] 本实施例设计的Loss函数的输入基于Thomsonproblem的输出结果,该结果的维度可以任意设置,通过在人脸识别过程采用人脸特征的编码向量代替原来的FC,使得神经网络的训练不受训练数据集大小的影响,提高了人脸识别效率。

[0040] 实施例二:

[0041] 本实施例中,在实施例一提供的方法基础上做进一步改进,实施例一设计的余弦损失函数可以让同一个人对应的人脸向量编码的夹角尽可能的小,让不同人脸向量编码的夹角尽可能的大。

[0042] 因此,本实施在训练过程中让CNN输出的Embedding自动匹配Thomson Problem求解的特征向量编码,本实施例采用最近匹配方案来进行Embedding自动匹配ThomsonProblem求解的特征向量编码过程。

[0043] 具体的,步骤四中利用余弦损失函数对第一人脸特征编码和第二人脸特征编码进行最近匹配计算过程包括:分别将第一人脸特征编码与第二人脸特征编码输入设计的余弦损失函数中,分别计算出第一人脸特征编码的夹角 $\theta_1$ 与第二人脸特征编码的夹角 $\theta_2$ ;根据最近匹配原则让夹角 $\theta_1$ 与夹角 $\theta_2$ 自动匹配,获得第一人脸特征编码和第二人脸特征编码之间的误差;根据余弦损失函数的导数,沿梯度最小方向将误差回传,修正CNN网络中的各层权重值,并重复进行人脸识别卷积神经网络训练,直至计算出的最近匹配结果满足设定的阈值时输出人脸识别结果。

[0044] 具体的,步骤二中CNN网络的网络结构包括:输入层、卷积层、池化层和输出层,输入层与卷积层连接,卷积层与池化层连接,池化层和输出层连接。

[0045] 本发明通过用人脸特征的编码向量代替原来的FC,由于人脸特征的向量编码是基于超球面的,超球面的维度我们可以做任意设置,所以我们训练模型的最后一层可以任意设置而不用根据人脸数据集中人的多少来更改。当然为了提高人脸特征的表达能力,我们同常会把超球面的维度设置为256维或者512维度,这样的维度是可以表示足够多的不同人脸的。同时我们在训练过程中让CNN生成的人脸特征与我们超球面求解的人脸特征尽可能匹配,这样相似不相同人脸在超球面上的表现也是挨着的特征向量,那么当我们在实际推理的时候,模型也会表现出对于相同的人脸特征向量夹角会很接近,对于相似不相同的人脸特征向量夹角是比较小的,对于不相同的人脸特征向量夹角是比较大的,这样在识别时也会得到比较好的人脸识别效果。

[0046] 本实施例可以达到以下技术效果:

[0047] 本实施例通过最近匹配原则使得CNN网络生成的人脸特征与超球面求解的人脸特征尽可能匹配,这样相似不相同人脸在超球面上的表现也是挨着的特征向量,那么当我们在实际推理的时候,模型也会表现出对于相同的人脸特征向量夹角会很接近,对于相似不相同的人脸特征向量夹角是比较小的,对于不相同的人脸特征向量夹角是比较大的,通过两个人脸特征的匹配提高了人脸识别时的精度和准确性。

[0048] 实施例三:

[0049] 本实施例中,一种基于余弦损失函数的人脸识别方法,包括以下步骤:

[0050] 步骤一:对需要进行人脸识别训练的人脸图像数据进行预处理形成训练样本集;

[0051] 步骤二:构建CNN网络,将训练样本集输入CNN网络进行人脸识别卷积神经网络训练,训练完成后输出第一人脸特征编码;

[0052] 步骤三:构建Thomson Problem模型,将训练样本集输入到Thomson Problem模型中进行求解,计算出第二人脸特征编码;

[0053] 步骤四:设计余弦损失函数,并利用余弦损失函数对第一人脸特征编码和第二人脸特征编码进行最近匹配计算,计算出的最近匹配结果对CNN网络模型参数进行修正,直至

计算出的最近匹配结果满足设定的阈值时输出人脸识别结果。

[0054] 本实施例中,在实施例二的基础上还可以采用匈牙利匹配算法来进行第一人臉特征编码和第二人脸特征编码的匹配计算。

[0055] 匈牙利算法是一种在多项式时间内求解任务分配问题的组合优化算法,并推动了后来的原始对偶方法。

[0056] 匈牙利算法的核心思想:初始时最大匹配为控,然后不断寻找增广路,并扩展增广路。不断重复这一过程直到找不到增广路为止。

[0057] 如果二分图左半边U集合中共有n个点,那么最多找n条增广路径。如果图中共有m条边,那么每找一条增广路径(DFS或BFS)时最多把所有边遍历一遍,所花时间也就是m。所以总时间大概就是 $O(nm)$ 。

[0058] 本实施例中,人脸识别过程主要包括以下过程:

[0059] 1. Thomson problem求解。在一个超球面上找到N个点,并让这N个点尽可能的均匀分布。如图3所示,假如超球面是二维,既是一个圆。白色圆点表示表示N的个数这里 $N=8$ ,黑色圆点表示超球面的球心即圆心。Thomson Problem 算法就是让这N个点均匀的分布在超球面上,这里超球面是二维的圆,所以就是让N的点均匀的分布在圆上,可利用现有的Thomson Problem模型求解方法求解人脸特征编码。

[0060] 2. 求解第一人臉特征编码(Embedding)。假设超球面上的N个白色圆点表示N个人对应的人脸特征编码。这样N个人脸的编码在球面就会尽可能的分开(均匀分布),那么在做人脸识别时就可以把不同的人脸轻易的分开。把超球心到超球面上的点组成的向量(上图3中的8条有方向的线),作为人脸的特征编码,假如训练集有50万人,即 $N=50$ 万。人脸特征编码的维度即可以理解成超球面的维度,图3中假设的是2维,超球面的维度是256维度或者512维度,在数据集增大时可以适当提高维度,来提高模型的识别能力。假如我们以256维作为超球面的维度,训练集有50万人,根据Thomson Problem就可以得到一个由所有球面上的点组成的特征编码,维度为 $[500000, 256]$ ,记为矩阵  $GT\_Maix$ ,每个256就是对一个人脸的编码。在获得人脸特征编码后训练人脸识别模型,让训练集的50万人自动找到自身对应的编码。

[0061] 以上显示和描述了本发明的基本原理、主要特征和本发明的优点。本行业的技术人员应该了解,本发明不受上述实施例的限制,上述实施例和说明书中描述的只是说明本发明的原理,在不脱离本发明精神和范围的前提下,本发明还会有各种变化和改进,这些变化和改进都落入要求保护的本发明范围内。本发明要求保护的范围由所附的权利要求书及其等效物界定。

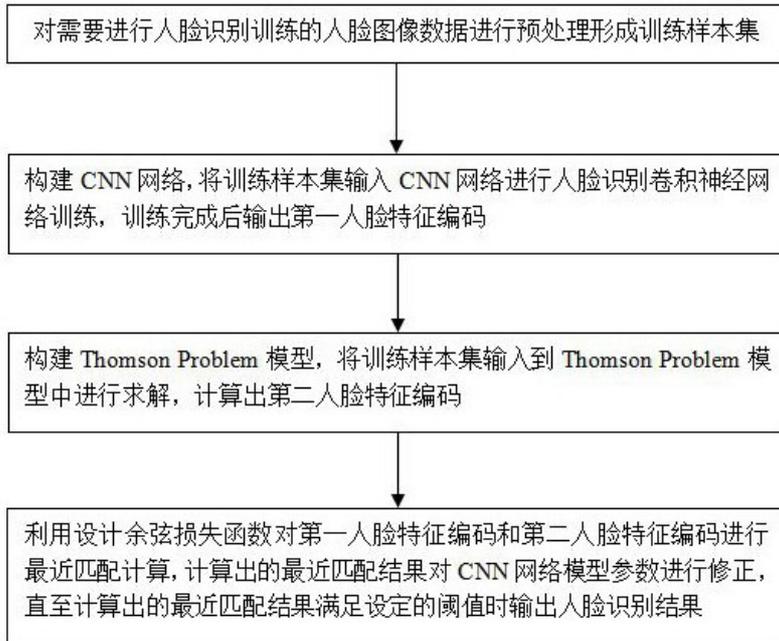


图1

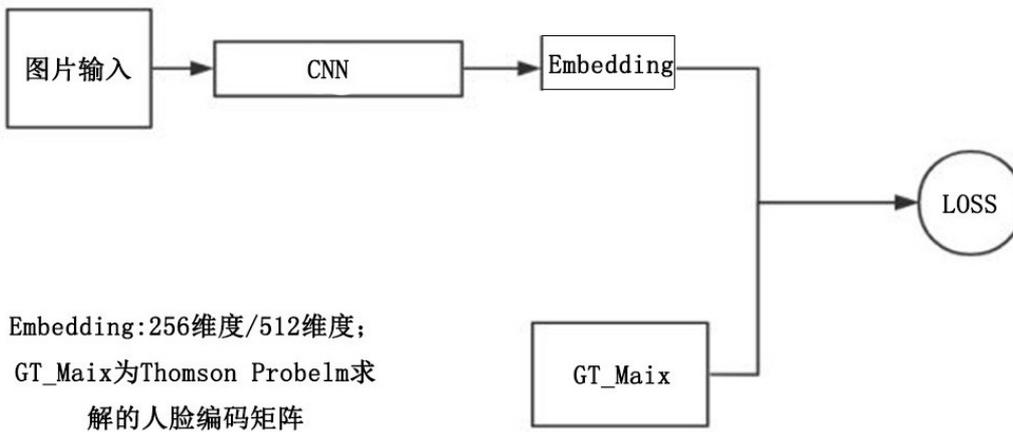


图2

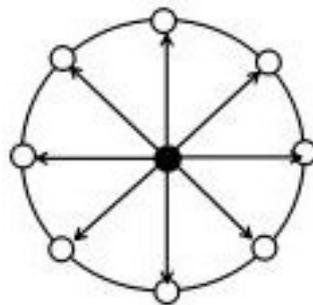


图3