



(21) 申请号 202311217076.5

(22) 申请日 2023.09.20

(65) 同一申请的已公布的文献号

申请公布号 CN 117238018 A

(43) 申请公布日 2023.12.15

(73) 专利权人 华南理工大学

地址 510640 广东省广州市天河区五山路  
381号

专利权人 人工智能与数字经济广东省实验  
室(广州)

(72) 发明人 张通 郭继凤 陈业林 陈俊龙

(74) 专利代理机构 广州市华学知识产权代理有  
限公司 44245

专利代理师 霍健兰

(51) Int.Cl.

G06V 40/16 (2022.01)

G06T 7/55 (2017.01)

G06V 10/44 (2022.01)

G06V 10/82 (2022.01)

G06V 40/40 (2022.01)

(56) 对比文件

C. L. Philip Chen et al. Broad  
Learning System: An Effective and  
Efficient Incremental Learning System  
Without the Need for Deep  
Architecture. IEEE Transactions on Neural  
Networks and Learning Systems. 2017, 10 -  
24.

朱玉斌; 延向军; 申旭奇; 卢兆林. 基于级联  
宽度学习的疲劳驾驶检测. 计算机工程与设计  
. 2020, (第02期), 全文.

审查员 王齐强

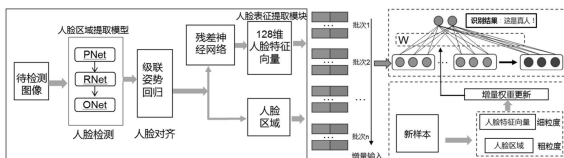
权利要求书3页 说明书8页 附图2页

(54) 发明名称

基于多粒度的可增量深宽网络活体检测方法、介质及设备

(57) 摘要

本发明提供了一种基于多粒度的可增量深宽网络活体检测方法、介质及设备; 其中方法为: 将待检测的图像依次进行人脸区域提取、人脸对齐、人脸特征向量提取处理, 形成多粒度信息; 将多粒度信息输入到稀疏宽度活体识别网络, 得到活体检测结果; 稀疏宽度活体识别网络的训练方法为: 将样本图像分成若干批次; 将样本图像的多粒度信息分批次输入到宽度活体识别网络进行学习, 更新权重; 当获取到新的样本图像时, 采用增量方式输入到宽度活体识别网络进行学习, 更新权重。该方法可解决学习时间消耗和资源占用问题, 检测精度高, 鲁棒性强。



1.一种基于多粒度的可增量深宽网络活体检测方法,其特征在于:将待检测的图像依次进行人脸区域提取、人脸对齐、人脸特征向量提取处理,得到人脸区域和 $C_2$ 维人脸特征;之后将包括人脸特征和人脸区域的多粒度信息融合,并输入到稀疏宽度活体识别网络,得到活体检测结果;

稀疏宽度活体识别网络是指经过训练的稀疏宽度活体识别网络;稀疏宽度活体识别网络的训练方法,包括如下步骤:

步骤Y1、获取当前的样本图像;根据样本图像对应的标签和机器容纳能力,将样本图像分成若干批次;

步骤Y2、将各个样本图像分别进行人脸区域提取、人脸对齐、人脸特征向量提取处理,得到各个样本图像的 $C_2$ 维人脸特征和人脸区域;对于各个批次,将批次对应样本图像的包括人脸特征和人脸区域的多粒度信息进行融合,得到数据 $B_a$ ;

步骤Y3、对各个批次数据 $B_a$ 依次进行学习:

步骤Y31、对于第1批次输入的数据 $B_1$ ,计算映射特征节点 $Z_1^n$ 和增强特征节点 $H_1^m$ ;合并映射特征节点 $Z_1^n$ 和增强特征节点 $H_1^m$ ,形成人脸活体识别特征层 $(A_n^m)_1 = [Z_1^n | H_1^m]$ ,并连接到稀疏宽度活体识别网络的输出层Y;

计算人脸活体识别特征层 $(A_n^m)_1$ 和输出层Y的连接权重 $(W_n^m)_1$ :

$$(W_n^m)_1 = (A_n^m)_1^{-1} Y ;$$

通过优化问题求解稀疏宽度学习的连接权重 $(W_n^m)_1$ :

$$\arg \min_{(W_n^m)_1} : \|(A_n^m)_1 (W_n^m)_1 - Y\|_v^{\sigma_1} + \lambda \|(W_n^m)_1\|_2^{\frac{1}{2}} ;$$

步骤Y32、对于第a批次输入的数据 $B_a$ ,  $a \geq 2$ , 计算增量映射特征节点 $Z_a^n$ 和增强特征节点 $H_a^m$ ;形成新的人脸活体识别特征层 $(A_n^m)_a$ 为:

$$(A_n^m)_a = \begin{bmatrix} (A_n^m)_{a-1} \\ [Z_a^n | H_a^m]^T \end{bmatrix};$$

输出层权重 $(W_n^m)_a$ 由如下公式更新:

$$D^T = (A_n^m)_a^T (A_n^m)_{a-1}^+;$$

$$C = (A_n^m)_a^T - D^T (A_n^m)_{a-1};$$

$$E^T = \begin{cases} C^+ & \text{if } C \neq 0 \\ ((1 + D^T D)^{-1} (A_n^m)_{a-1}^+ D & \text{if } C = 0 \end{cases};$$

$$(W_n^m)_a = (W_n^m)_{a-1} + [Y_{B_a}^T - (A_n^m)_a^T (W_n^m)_{a-1}] E;$$

其中,  $Y_{B_a}^T$ 为数据 $B_a$ 对应的标签;

当获取到新的样本图像时,根据新的样本图像对应的标签和机器容纳能力,将新的样本图像分成若干批次;之后执行步骤Y2和步骤Y32;

所述稀疏宽度活体识别网络的输出Y为:

$$Y = \text{Max}([Z_a^n | H_a^m](W_n^m)_a)。$$

2. 根据权利要求1所述的基于多粒度的可增量深宽网络活体检测方法,其特征在於:在稀疏宽度活体识别网络的训练方法中,对于第a批次输入的数据 $B_a$ ,  $a \geq 1$ , 计算增量映射特征节点 $Z_a^n$ 和增强特征节点 $H_a^m$ 是指:

$$Z_a^n = [\varphi(B_a W_{e1} + \beta_{e1}), \dots, \varphi(B_a W_{en} + \beta_{en})];$$

$$H_a^m = [\zeta([Z_a^n] W_{h1} + \beta_{h1}), \dots, \zeta([Z_a^n] W_{hm} + \beta_{hm})];$$

其中, $\varphi$ 为随机线性映射函数; $\zeta$ 为一个非线性激活函数;连接权重 $W_{ei}$ 和 $\beta_{ei}$ ,以及连接权重 $W_{hj}$ 和 $\beta_{hj}$ 是随机产生; $e=1,2,\dots,n$ ;  $j=1,2,\dots,m$ 。

3. 根据权利要求1所述的基于多粒度的可增量深宽网络活体检测方法,其特征在於:所述将待检测的图像依次进行人脸区域提取、人脸对齐、人脸特征向量提取处理,得到人脸区域和 $C_2$ 维人脸特征;之后将包括 $C_2$ 维人脸特征和人脸区域的多粒度信息融合,并输入到稀疏宽度活体识别网络,得到活体检测结果,是指:包括如下步骤:

步骤S1、将待检测的图像输入到人脸区域提取模型;人脸区域提取模型对图像中的人脸进行检测,得到人脸区域的坐标信息;

步骤S2、对人脸区域,通过关键点估计方法找到 $C_1$ 个初步的人脸关键点;通过级联姿势回归算法使各个人脸关键点对齐,从而将人脸区域中的脸部调整为设定的正脸姿态和大小,得到调整后的人脸图像;

步骤S3、使用人脸表征提取模块对人脸图像进行 $C_2$ 维人脸特征向量提取;

步骤S4、将包括人脸特征和人脸区域的多粒度信息融合,并输入到稀疏宽度活体识别网络;根据稀疏宽度活体识别网络的输出,得到活体检测结果。

4. 根据权利要求3所述的基于多粒度的可增量深宽网络活体检测方法,其特征在於:所述步骤S1中,人脸区域提取模型包括P-Net、R-Net和O-Net;

所述P-Net用于判别是否存在人脸,并使用边界框回归和特征点定位器初步提取人脸区域,输出多张人脸区域;

所述R-Net用于对P-Net输出的多张人脸区域进行再次判别是否存在人脸,删除错误输入,再次使用边界框回归和特征点定位器进行人脸区域的边界框回归和人脸特征点定位,输出可信度更高的人脸区域;

所述O-Net用于对R-Net输出的多张人脸区域进行第三次判别是否存在人脸,使用边界框回归和特征点定位器进行人脸区域的边界框回归和人脸特征点定位,最终输出人脸区域的坐标信息与人脸的五个特征点。

5. 根据权利要求4所述的基于多粒度的可增量深宽网络活体检测方法,其特征在於:所述人脸区域的坐标信息包括人脸区域的左上角坐标和右下角坐标;人脸的五个特征点包括左眼位置、右眼位置、鼻子位置以及嘴巴左、右位置;每个特征点用二维表示。

6. 根据权利要求3所述的基于多粒度的可增量深宽网络活体检测方法,其特征在於:所述步骤S2中, $C_1$ 个初步的人脸关键点包括下巴顶部的关键点、每只眼睛外部轮廓的关键点、每条眉毛内部轮廓的关键点、鼻子外部轮廓的关键点、嘴唇内部轮廓和外部轮廓的关键点。

7. 根据权利要求3所述的基于多粒度的可增量深宽网络活体检测方法,其特征在於:所述S3中,人脸表征提取模块是指在ResNet模型的基础上添加池化层,用于将ResNet模型的

输出转化为一个 $C_2$ 维人脸特征向量；

人脸表征提取模块是指经过训练的人脸表征提取模块；在人脸表征提取模块的训练过程中，每次随机选择一个三元组，并使用随机梯度下降优化器来最小化损失函数Loss；所述三元组是指在训练样本中抽出任三张人脸图像，其中第一张和第二张人脸图像为真人正样本，分别记为 $x_{1k}$ 和 $x_{2k}$ ，第三张人脸图像为攻击手段对应的负样本，记为 $x_{3k}$ ；

损失函数Loss为：

$$\text{Loss} = \sum_{k=1}^K [\|f(x_{1k}) - f(x_{2k})\|_2^2 - \|f(x_{1k}) - f(x_{3k})\|_2^2 + l];$$

其中，N表示三元组数量； $\|f(x_{1k}) - f(x_{2k})\|_2^2$ 表示第一张人脸图像与第二张人脸图像的人脸特征向量的欧氏距离或马氏距离； $\|f(x_{1k}) - f(x_{3k})\|_2^2$ 表示第一张人脸图像与第三张人脸图像的人脸特征向量的欧氏距离或马氏距离；1代表用于强制增加正样本和负样本间距离的参数。

8. 根据权利要求3所述的基于多粒度的可增量深宽网络活体检测方法，其特征在于：根据稀疏宽度活体识别网络的输出Y，计算置信度 $Cd_p$ ：

$$Cd_p = \frac{e^{Y_p}}{\sum_q e^{Y_q}} \quad p, q = 1, 2;$$

其中， $Y_p$ 代表稀疏宽度活体识别网络的输出； $Y_q$ 代表稀疏宽度活体识别网络判断图像为第q种类别对应的输出；

根据置信度 $Cd_p$ ，得到活体检测结果。

9. 一种可读存储介质，其特征在于，其中所述存储介质存储有计算机程序，所述计算机程序当被处理器执行时使所述处理器执行权利要求1-8中任一项所述的基于多粒度的可增量深宽网络活体检测方法。

10. 一种计算机设备，包括处理器以及用于存储处理器可执行程序存储器，其特征在于，所述处理器执行存储器存储的程序时，实现权利要求1-8中任一项所述的基于多粒度的可增量深宽网络活体检测方法。

## 基于多粒度的可增量深宽网络活体检测方法、介质及设备

### 技术领域

[0001] 本发明涉及活体检测技术领域,更具体地说,涉及一种基于多粒度的可增量深宽网络活体检测方法、介质及设备。

### 背景技术

[0002] 活体检测技术主要分为基于手工设计特征的方法和基于深度学习的方法。人工设计的特征针对图像采集时的信息损失和噪声引入,对比图像的纹理差异,如局部高光、阴影变化、模糊程度和低频分量信息损失等实现识别目的。随着高清摄像机和高清3D面具的应用,其不足之处日益凸显。基于运动信息的检测方法是常见识别率较高的人脸认证技术,但它需要认证人员的高度配合,检测过程不友好,且耗时较长。这类方法虽识别率较高,但严重依赖于特征表达(需要解决细节损失、颜色失真、阴影模糊和图像高光等问题)和硬件支持,在视频回放、3D面具等逼真的伪信息下,鲁棒性和泛化能力有限。

[0003] 相较于基于手工设计特征的方法,基于深度学习的活体检测方式适用于各种欺骗手段。如针对照片和视频攻击的双流CNN的人脸反欺骗方法;反3D面具欺骗方法;使用Inception和ResNet架构在不同的环境下的人脸欺骗检测等。然而该类方法成本高、体量大,轻量化部署难度高,不能满足在线实时处理需求。

[0004] 宽度学习作为一种单隐层神经网络,保留高度非线性逼近能力的同时,能在短时间内实现较高识别率。若能将宽度学习网络应用于活体检测领域,则可克服深度学习资源消耗高,训练时间长等缺点,满足工业中轻量化部署的需求。同时,考虑到随着攻击手段更新,神经网络模型也需要不断学习,因此,如何加快新样本学习,对活体检测技术也十分重要。

### 发明内容

[0005] 为克服现有技术中的缺点与不足,本发明的目的在于提供一种基于多粒度的可增量深宽网络活体检测方法、介质及设备;该方法可解决学习时间消耗和资源占用问题,检测精度高,鲁棒性强。

[0006] 为了达到上述目的,本发明通过下述技术方案予以实现:一种基于多粒度的可增量深宽网络活体检测方法,将待检测的图像依次进行人脸区域提取、人脸对齐、人脸特征向量提取处理,得到人脸区域和 $C_2$ 维人脸特征;之后将包括人脸特征和人脸区域的多粒度信息融合,并输入到稀疏宽度活体识别网络,得到活体检测结果;

[0007] 稀疏宽度活体识别网络是指经过训练的稀疏宽度活体识别网络;稀疏宽度活体识别网络的训练方法,包括如下步骤:

[0008] 步骤Y1、获取当前的样本图像;根据样本图像对应的标签和机器容纳能力,将样本图像分成若干批次;

[0009] 步骤Y2、将各个样本图像分别进行人脸区域提取、人脸对齐、人脸特征向量提取处理,得到各个样本图像的 $C_2$ 维人脸特征和人脸区域;对于各个批次,将批次对应样本图像的

包括人脸特征和人脸区域的多粒度信息进行融合,得到数据 $B_a$ ;

[0010] 步骤Y3、对各个批次数据 $B_a$ 依次进行学习;

[0011] 步骤Y31、对于第1批次输入的数据 $B_1$ ,计算映射特征节点 $Z_1^n$ 和增强特征节点 $H_1^m$ ;

合并映射特征节点 $Z_1^n$ 和增强特征节点 $H_1^m$ ,形成人脸活体识别特征层 $(A_n^m)_1 = [Z_1^n | H_1^m]$ ,并连接到稀疏宽度活体识别网络的输出层Y;

[0012] 计算人脸活体识别特征层 $(A_n^m)_1$ 和输出层Y的连接权重 $(W_n^m)_1$ ;

[0013]  $(W_n^m)_1 = (A_n^m)_1^{-1} Y$ ;

[0014] 通过优化问题求解稀疏宽度学习的连接权重 $(W_n^m)_1$ ;

[0015]  $\arg \min_{(W_n^m)_1} : \|(A_n^m)_1 (W_n^m)_1 - Y\|_v^{\sigma_1} + \lambda \|(W_n^m)_1\|_2^{\frac{1}{2}}$ ;

[0016] 步骤Y32、对于第a ( $a \geq 2$ ) 批次输入的数据 $B_a$ ,计算增量映射特征节点 $Z_a^n$ 和增强特征节点 $H_a^m$ ;形成新的人脸活体识别特征层 $(A_n^m)_a$ 为:

[0017]  $(A_n^m)_a = \begin{bmatrix} (A_n^m)_{a-1} \\ [Z_a^n | H_a^m]^T \end{bmatrix}$ ;

[0018] 输出层权重 $(W_n^m)_a$ 由如下公式更新:

[0019]  $D^T = (A_n^m)_a^T (A_n^m)_{a-1}^+$ ;

[0020]  $C = (A_n^m)_a^T - D^T (A_n^m)_{a-1}$ ;

[0021]  $E^T = \begin{cases} C^+ & \text{if } C \neq 0 \\ ((1 + D^T D)^{-1} (A_n^m)_{a-1}^+ D & \text{if } C = 0 \end{cases}$ ;

[0022]  $(W_n^m)_a = (W_n^m)_{a-1} + [Y_{B_a}^T - (A_n^m)_a^T (W_n^m)_{a-1}] E$ ;

[0023] 其中, $Y_{B_a}^T$ 为数据 $B_a$ 对应的标签;

[0024] 当获取到新的样本图像时,根据新的样本图像对应的标签和机器容纳能力,将新的样本图像分成若干批次;之后执行步骤Y2和步骤Y32;

[0025] 所述稀疏宽度活体识别网络的输出Y为:

[0026]  $Y = \text{Max}([Z_a^n | H_a^m] (W_n^m)_a)$ 。

[0027] 在稀疏宽度活体识别网络中,由于稀疏宽度学习输入的是人脸特征和脸部区域形成的高维数据,直接使用一次性样本内存和计算的压力较大,因此宽度活体识别网路采用增量学习的方式进行训练;这样既能缓解计算压力,又能在动态环境中不断更新数据和模型。

[0028] 此外,由于在实际场景中,随着新的攻击手段更新,会有新的样本图像;如果新的样本图像(包括原始攻击手段和新攻击手段)到来时,如果需要重新训练稀疏宽度活体识别网络,无疑会重复相同的工作并浪费大量的计算资源。本发明中的增量学习设计致力于解决这一尴尬问题。对于新的样本图像,直接使用相同的过程来提取人脸的多粒度信息;采用

与以上a ( $a \geq 2$ ) 批次计算过程,执行增量学习更新。

[0029] 优选地,在稀疏宽度活体识别网络的训练方法中,对于第a ( $a \geq 1$ ) 批次输入的数据  $B_a$ ,计算增量映射特征节点  $Z_a^n$  和增强特征节点  $H_a^m$  是指:

[0030]  $Z_a^n = [\varphi(B_a W_{e1} + \beta_{e1}), \dots, \varphi(B_a W_{en} + \beta_{en})];$

[0031]  $H_a^m = [\zeta([Z_a^n] W_{h1} + \beta_{h1}), \dots, \zeta([Z_a^n] W_{hm} + \beta_{hm})];$

[0032] 其中, $\varphi$ 为随机线性映射函数; $\zeta$ 为一个非线性激活函数;连接权重  $W_{ei}$  和  $\beta_{ei}$  ( $e=1, 2, \dots, n$ ),以及连接权重  $W_{hj}$  和  $\beta_{hj}$  ( $j=1, 2, \dots, m$ ) 是随机产生。

[0033] 优选地,所述将待检测的图像依次进行人脸区域提取、人脸对齐、人脸特征向量提取处理,得到人脸区域和  $C_2$  维人脸特征;之后将包括  $C_2$  维人脸特征和人脸区域的多粒度信息融合,并输入到稀疏宽度活体识别网络,得到活体检测结果,是指:包括如下步骤:

[0034] 步骤S1、将待检测的图像输入到人脸区域提取模型;人脸区域提取模型对图像中的人脸进行检测,得到人脸区域的坐标信息;

[0035] 步骤S2、对人脸区域,通过关键点估计方法找到  $C_1$  个初步的人脸关键点;通过级联姿势回归算法使各个人脸关键点对齐,从而将人脸区域中的脸部调整为设定的正脸姿态和大小,得到调整后的人脸图像;

[0036] 步骤S3、使用人脸表征提取模块对人脸图像进行  $C_2$  维人脸特征向量提取;

[0037] 步骤S4、将包括人脸特征和人脸区域的多粒度信息融合,并输入到稀疏宽度活体识别网络;根据稀疏宽度活体识别网络的输出,得到活体检测结果。

[0038] 优选地,所述步骤S1中,人脸区域提取模型包括P-Net、R-Net和O-Net;

[0039] 所述P-Net用于判别是否存在人脸,并使用边界框回归和特征点定位器初步提取人脸区域,输出多张人脸区域;

[0040] 所述R-Net用于对P-Net输出的多张人脸区域进行再次判别是否存在人脸,删除错误输入,再次使用边界框回归和特征点定位器进行人脸区域的边界框回归和人脸特征点定位,输出可信度更高的人脸区域;

[0041] 所述O-Net用于对R-Net输出的多张人脸区域进行第三次判别是否存在人脸,使用边界框回归和特征点定位器进行人脸区域的边界框回归和人脸特征点定位,最终输出人脸区域的坐标信息与人脸的五个特征点。

[0042] 优选地,所述人脸区域的坐标信息包括人脸区域的左上角坐标和右下角坐标;人脸的五个特征点包括左眼位置、右眼位置、鼻子位置以及嘴巴左、右位置;每个特征点用二维表示。

[0043] 优选地,所述步骤S2中,  $C_1$  个初步的人脸关键点包括下巴顶部的关键点、每只眼睛外部轮廓的关键点、每条眉毛内部轮廓的关键点、鼻子外部轮廓的关键点、嘴唇内部轮廓和外部轮廓的关键点。

[0044] 优选地,所述S3中,人脸表征提取模块是指在ResNet模型的基础上添加池化层,用于将ResNet模型的输出转化为一个  $C_2$  维人脸特征向量;

[0045] 人脸表征提取模块是指经过训练的人脸表征提取模块;在人脸表征提取模块的训练过程中,每次随机选择一个三元组,并使用随机梯度下降优化器来最小化损失函数Loss;所述三元组是指在训练样本中抽出任三张人脸图像,其中第一张和第二张人脸图像为真人

正样本,分别记为 $x_{1k}$ 和 $x_{2k}$ ,第三张人脸图像为攻击手段对应的负样本,记为 $x_{3k}$ ;

[0046] 损失函数Loss为:

$$[0047] \quad \text{Loss} = \sum_{k=1}^K [\|f(x_{1k}) - f(x_{2k})\|_2^2 - \|f(x_{1k}) - f(x_{3k})\|_2^2 + l];$$

[0048] 其中,N表示三元组数量; $\|f(x_{1k}) - f(x_{2k})\|_2^2$ 表示第一张人脸图像与第二张人脸图像的人脸特征向量的欧氏距离或马氏距离; $\|f(x_{1k}) - f(x_{3k})\|_2^2$ 表示第一张人脸图像与第三张人脸图像的人脸特征向量的欧氏距离或马氏距离;l代表用于强制增加正样本和负样本间距离的参数。

[0049] 优选地,根据稀疏宽度活体识别网络的输出Y,计算置信度 $Cd_p$ :

$$[0050] \quad Cd_p = \frac{e^{Y_p}}{\sum_q e^{Y_q}} \quad p, q = 1, 2;$$

[0051] 其中, $Y_p$ 代表稀疏宽度活体识别网络的输出; $Y_q$ 代表稀疏宽度活体识别网络判断图像为第q种类别对应的输出;

[0052] 根据置信度 $Cd_p$ ,得到活体检测结果。

[0053] 一种可读存储介质,其中所述存储介质存储有计算机程序,所述计算机程序当被处理器执行时使所述处理器执行上述基于多粒度的可增量深宽网络活体检测方法。

[0054] 一种计算机设备,包括处理器以及用于存储处理器可执行程序存储器,所述处理器执行存储器存储的程序时,实现上述基于多粒度的可增量深宽网络活体检测方法。

[0055] 与现有技术相比,本发明具有如下优点与有益效果:

[0056] 1、本发明采用人脸区域和多维人脸特征的多粒度信息输入到稀疏宽度活体识别网络中;人脸区域能够涵盖更多信息,但是没有针对性的细节信息,只利用人脸区域进一步提取特征可能会丢弃后期宽度无法获取的细节信息;而人脸特征则在一定程度上弥补了这一缺点,能提供初步的脸部相关的细节信息,实现互补;

[0057] 2、为避免高维多粒度信息中存在的冗余和无用信息干扰,本发明设计基于L1/2正则化的稀疏宽度学习,使其具有特征选择和产生稀疏权重的能力,从而提高模型识别效果;

[0058] 3、为解决同时学习全部多粒度信息对运算资源需求太大的问题,本发明中宽度学习采用增量输入方式,形成数据流模式,宽度学习一次只学习部分数据,并在前期知识上根据新批次的数据不断更新权重;另一方面,本发明采用增量学习方式能够方便利用新增加的各种类型攻击手段数据,无需模型重建和重新训练,使用最低成本提高模型鲁棒性;

[0059] 4、本发明具有检测精度高、耗时短、鲁棒性强、实时响应等优点。

## 附图说明

[0060] 图1是本发明基于多粒度的可增量深宽网络活体检测方法的流程示意图;

[0061] 图2是本发明基于多粒度的可增量深宽网络活体检测方法中,人脸区域提取模型的结构示意图;

[0062] 图3是本发明基于多粒度的可增量深宽网络活体检测方法中,稀疏宽度活体识别网络的训练流程示意图。

## 具体实施方式

[0063] 下面结合附图与具体实施方式对本发明作进一步详细的描述。

[0064] 实施例一

[0065] 本实施例一种基于多粒度的可增量深宽网络活体检测方法,如图1所示,包括如下步骤:

[0066] 步骤S1、将待检测的图像输入到人脸区域提取模型;人脸区域提取模型对图像中的人脸进行检测,得到人脸区域的坐标信息。

[0067] 人脸区域提取模型,如图2所示,包括P-Net、R-Net和O-Net;

[0068] 所述P-Net用于判别是否存在人脸,并使用边界框回归和特征点定位器初步提取人脸区域,输出多张人脸区域;

[0069] 所述R-Net用于对P-Net输出的多张人脸区域进行再次判别是否存在人脸,删除错误输入,再次使用边界框回归和特征点定位器进行人脸区域的边界框回归和人脸特征点定位,输出可信度更高的人脸区域;

[0070] 所述O-Net用于对R-Net输出的多张人脸区域进行第三次判别是否存在人脸,使用边界框回归和特征点定位器进行人脸区域的边界框回归和人脸特征点定位,最终输出人脸区域的坐标信息与人脸的五个特征点。

[0071] 所述P-Net输入为 $12 \times 12 \times 3$ 的图像,经过三个 $3 \times 3$ 的卷积层;P-Net的输出包括三个部分:P-Net的第一部分输出大小为 $1 \times 1 \times 2$ 的向量,表示是否存在人脸;P-Net的第二部分输出大小为 $1 \times 1 \times 4$ 的向量,表示人脸边界框位置相对完美人脸框位置的偏移信息;偏移信息包括:人脸边界框左上角横坐标的相对偏移、人脸边界框左上角纵坐标的相对偏移、人脸边界框的宽度误差以及人脸边界框的高度误差;P-Net的第三部分输出大小为 $1 \times 1 \times 10$ 的向量,表示人脸的五个特征点位置;

[0072] 所述R-Net输入为 $24 \times 24 \times 3$ 的图像,经过大小分别为 $3 \times 3$ 、 $3 \times 3$ 和 $2 \times 2$ 的三个卷积层后接一个128全连接层;R-Net的输出包括三个部分:R-Net的第一部分输出大小为 $1 \times 1 \times 2$ 的向量,用于进行多张人脸区域进行细化选择,删除部分错误输入;R-Net的第二部分输出大小为 $1 \times 1 \times 4$ 的向量,表示人脸边界框位置相对完美人脸框位置的偏移;R-Net的第三部分输出大小为 $1 \times 1 \times 10$ 的向量,表示人脸的五个特征点位置;

[0073] 所述O-Net输入为 $48 \times 48 \times 3$ 的图像,经过大小分别为大小分别为 $3 \times 3$ 、 $3 \times 3$ 、 $3 \times 3$ 和 $2 \times 2$ 的三个卷积层后接一个256全连接层,输出人脸区域的坐标信息和人脸的五个特征点位置。

[0074] 人脸区域的坐标信息包括人脸区域的左上角坐标和右下角坐标;人脸的五个特征点包括左眼位置、右眼位置、鼻子位置以及嘴巴左、右位置;每个特征点用二维表示。

[0075] 步骤S2、对人脸区域,通过关键点估计方法找到 $C_1$ 个(例如:68个)初步的人脸关键点;68个初步的人脸关键点包括下巴顶部的关键点、每只眼睛外部轮廓的关键点、每条眉毛内部轮廓的关键点、鼻子外部轮廓的关键点、嘴唇内部轮廓和外部轮廓的关键点。通过级联姿势回归算法使各个人脸关键点对齐,从而将人脸区域中的脸部调整为设定的正脸姿态和大小,得到调整后的人脸图像。

[0076] 步骤S3、使用人脸表征提取模块对人脸图像进行 $C_2$ 维(例如:128维)人脸特征向量提取。

[0077] 人脸表征提取模块是指在ResNet模型的基础上添加池化层,用于将ResNet模型的输出转化为一个 $C_2$ 维人脸特征向量;

[0078] 人脸表征提取模块是指经过训练的人脸表征提取模块;在人脸表征提取模块的训练过程中,每次随机选择一个三元组,并使用随机梯度下降优化器来最小化损失函数Loss;所述三元组是指在训练样本中抽出任三张人脸图像,其中第一张和第二张人脸图像为真人正样本,分别记为 $x_{1k}$ 和 $x_{2k}$ ,第三张人脸图像为攻击手段对应的负样本,记为 $x_{3k}$ ;

[0079] 损失函数Loss为:

$$[0080] \quad \text{Loss} = \sum_{k=1}^K [\|f(x_{1k}) - f(x_{2k})\|_2^2 - \|f(x_{1k}) - f(x_{3k})\|_2^2 + l];$$

[0081] 其中,N表示三元组数量; $\|f(x_{1k}) - f(x_{2k})\|_2^2$ 表示第一张人脸图像与第二张人脸图像的人脸特征向量的欧氏距离或马氏距离; $\|f(x_{1k}) - f(x_{3k})\|_2^2$ 表示第一张人脸图像与第三张人脸图像的人脸特征向量的欧氏距离或马氏距离;l代表用于强制增加正样本和负样本间距离的参数。

[0082] 因此,距离 $\|f(x_{ai}) - f(x_{pi})\|^2$ 应该较小,而距离 $\|f(x_{ai}) - f(x_{ni})\|^2$ 应该较大。

[0083] 步骤S4、将包括 $C_2$ 维人脸特征和人脸区域的多粒度信息融合为特征向量F;输入到稀疏宽度活体识别网络。

[0084] 稀疏宽度活体识别网络是指经过训练的稀疏宽度活体识别网络;稀疏宽度活体识别网络的训练方法如图3所示,包括如下步骤:

[0085] 步骤Y1、获取当前的采集样本图像;根据样本图像对应的标签和机器容纳能力,将样本图像分成若干批次;

[0086] 步骤Y2、将各个样本图像分别进行步骤S1~S3处理,得到各个样本图像的 $C_2$ 维人脸特征和人脸区域( $160 \times 160 \times 3$ );对于各个批次,将批次对应样本图像的包括 $C_2$ 维人脸特征和人脸区域的多粒度信息进行融合,得到数据 $B_a$ ;

[0087] 步骤Y3、对各个批次数据 $B_a$ 依次进行学习:

[0088] 步骤Y31、对于第1批次输入的数据 $B_1$ ,计算映射特征节点 $Z_1^n$ 和增强特征节点 $H_1^m$ :

$$[0089] \quad Z_1^n = [\varphi(B_1 W_{e1} + \beta_{e1}), \dots, \varphi(B_1 W_{en} + \beta_{en})];$$

$$[0090] \quad H_1^m = [\zeta([Z_1^n] W_{h1} + \beta_{h1}), \dots, \zeta([Z_1^n] W_{hm} + \beta_{hm})];$$

[0091] 其中, $\varphi$ 为随机线性映射函数; $\zeta$ 为一个非线性激活函数;连接权重 $W_{ei}$ 和 $\beta_{ei}$  ( $e=1, 2, \dots, n$ ),以及连接权重 $W_{hj}$ 和 $\beta_{hj}$  ( $j=1, 2, \dots, m$ )是随机产生。

[0092] 合并映射特征节点 $Z_1^n$ 和增强特征节点 $H_1^m$ ,形成人脸活体识别特征层 $(A_n^m)_1 = [Z_1^n | H_1^m]$ ,并连接到稀疏宽度活体识别网络的输出层Y;

[0093] 计算人脸活体识别特征层 $(A_n^m)_1$ 和输出层Y的连接权重 $(W_n^m)_1$ :

$$[0094] \quad (W_n^m)_1 = (A_n^m)^{-1} Y;$$

[0095] 通过优化问题求解稀疏宽度学习的连接权重 $(W_n^m)_1$ :

$$[0096] \quad \arg \min_{(W_n^m)_1} : \|(A_n^m)_1 (W_n^m)_1 - Y\|_v^{\sigma_1} + \lambda \|(W_n^m)_1\|_2^{\frac{1}{2}};$$

[0097] 步骤Y32、对于第a ( $a \geq 2$ ) 批次输入的数据 $B_a$ ,计算增量映射特征节点 $Z_a^n$ 和增强特

征节点  $H_a^m$ :

$$[0098] \quad Z_a^n = [\varphi(B_a W_{e1} + \beta_{e1}), \dots, \varphi(B_a W_{en} + \beta_{en})];$$

$$[0099] \quad H_a^m = [\zeta([Z_a^n] W_{h1} + \beta_{h1}), \dots, \zeta([Z_a^n] W_{hm} + \beta_{hm})];$$

[0100] 形成新的人脸活体识别特征层  $(A_n^m)_a$  为:

$$[0101] \quad (A_n^m)_a = \begin{bmatrix} (A_n^m)_{a-1} \\ [Z_a^n | H_a^m]^T \end{bmatrix};$$

[0102] 输出层权重  $(W_n^m)_a$  由如下公式更新:

$$[0103] \quad D^T = (A_n^m)_a^T (A_n^m)_{a-1}^+;$$

$$[0104] \quad C = (A_n^m)_a^T - D^T (A_n^m)_{a-1};$$

$$[0105] \quad E^T = \begin{cases} C^+ & \text{if } C \neq 0 \\ ((1 + D^T D)^{-1} (A_n^m)_{a-1}^+ D & \text{if } C = 0 \end{cases};$$

$$[0106] \quad (W_n^m)_a = (W_n^m)_{a-1} + [Y_{B_a}^T - (A_n^m)_a^T (W_n^m)_{a-1}] E;$$

[0107] 其中,  $Y_{B_a}^T$  为数据  $B_a$  对应的标签;

[0108] 在完成所有批次数据学习后,测试稀疏宽度活体识别网络的精度;当精度不满足要求时,重新进行各个批次数据学习,直至精度满足要求;

[0109] 当获取到新的样本图像时,根据新的样本图像对应的标签和机器容纳能力,将新的样本图像分成若干批次;之后执行步骤Y2和步骤Y32;

[0110] 所述稀疏宽度活体识别网络的输出Y为:

$$[0111] \quad Y = \text{Max}([Z_a^n | H_a^m] (W_n^m)_a)。$$

[0112] 根据稀疏宽度活体识别网络的输出,得到活体检测结果。

[0113] 具体地说,根据稀疏宽度活体识别网络的输出Y,计算置信度  $Cd_p$ :

$$[0114] \quad Cd_p = \frac{e^{Y_p}}{\sum_q e^{Y_q}} \quad p, q = 1, 2;$$

[0115] 其中,  $Y_p$  代表稀疏宽度活体识别网络的输出;  $Y_q$  代表稀疏宽度活体识别网络判断图像为第q种类别对应的输出;

[0116] 根据置信度  $Cd_p$ ,得到活体检测结果。

[0117] 实施例二

[0118] 本实施例一种可读存储介质,其中所述可读存储介质存储有计算机程序,所述计算机程序当被处理器执行时使所述处理器执行实施例一所述的基于多粒度的可增量深宽网络活体检测方法。

[0119] 实施例三

[0120] 本实施例一种计算机设备,包括处理器以及用于存储处理器可执行程序的存储器,所述处理器执行存储器存储的程序时,实现实施例一所述的基于多粒度的可增量深宽网络活体检测方法。

[0121] 上述实施例为本发明较佳的实施方式,但本发明的实施方式并不受上述实施例的限制,其他的任何未背离本发明的精神实质与原理下所作的改变、修饰、替代、组合、简化,均应为等效的置换方式,都包含在本发明的保护范围之内。

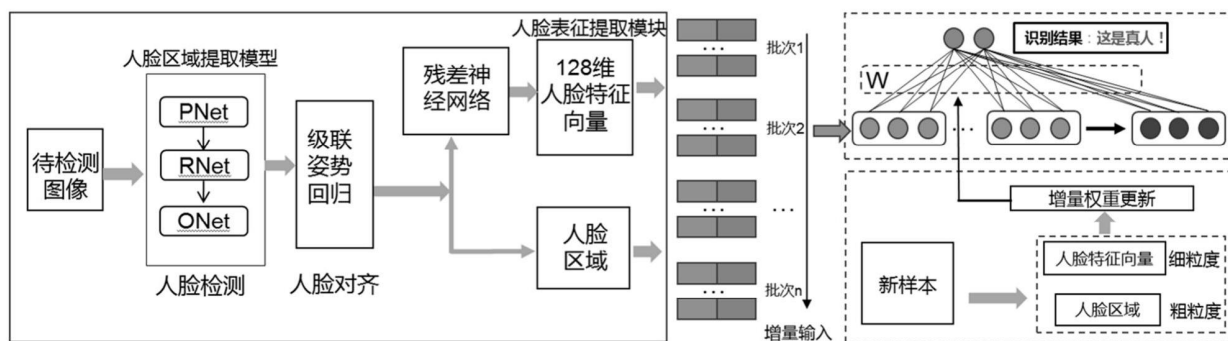


图1

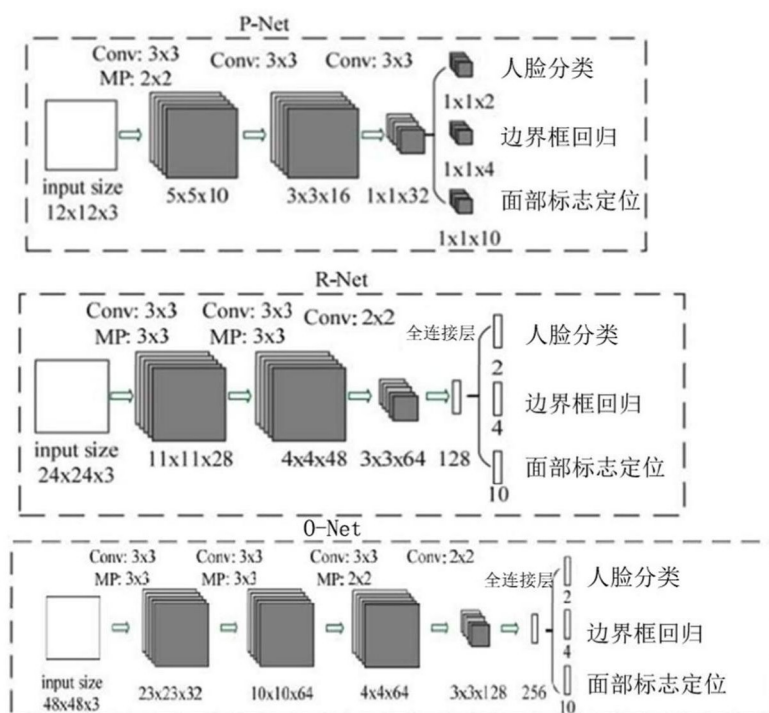


图2

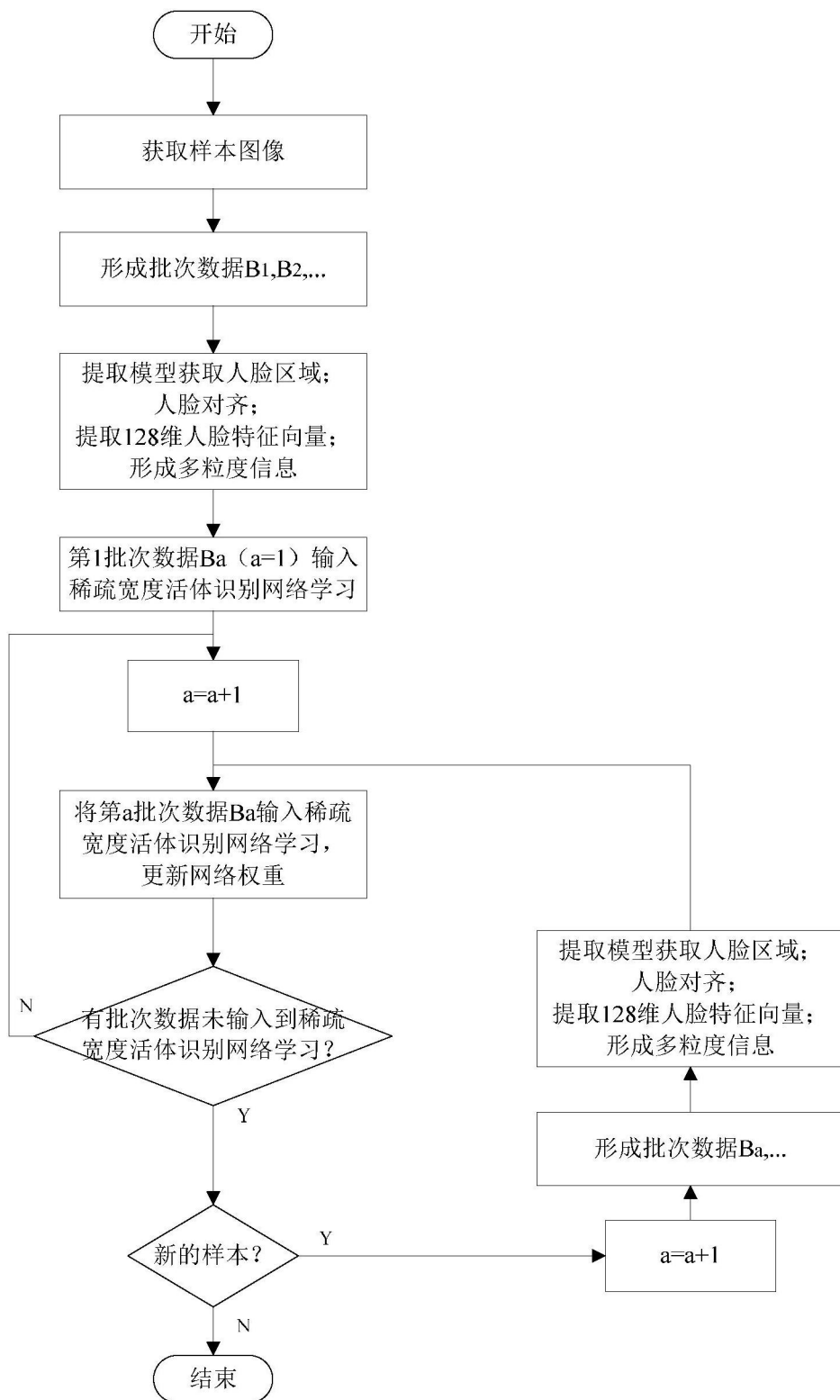


图3