

权 利 要 求 书

- 1.一种人脸和人手关联检测方法，其特征在于，包括：
 - 获取第一图像，所述第一图像为人物对象的图像；
 - 对所述第一图像执行特征提取，得到多个尺度的第一特征图；
 - 对所述多个尺度的第一特征图执行特征融合处理，得到多个尺度的第二特征图，所述第二特征图
- 5 的尺度与所述第一特征图的尺度一一对应；
 - 基于得到的所述多个尺度的第二特征图检测所述第一图像中针对同一人物对象的关联的人脸位置
 - 和人手位置。
- 2.根据权利要求1所述的方法，其特征在于，所述获取第一图像，包括：
 - 获取第二图像，所述第二图像为包括至少一个物体对象的图像；
 - 10 对所述第二图像执行人体目标检测，得到所述第一图像中所述至少一个物体对象中任一物体对象的
 - 检测框；
 - 将所述任一物体对象的所述检测框在所述第二图像中对应的图像区域，确定为所述任一物体对象的
 - 第一图像。
- 3.根据权利要求1或2所述的方法，其特征在于，所述对所述第一图像执行特征提取，得到多个尺
- 15 度的第一特征图，包括：
 - 将所述第一图像调整为预设尺度的第三图像；
 - 将所述第三图像输入至残差网络，得到所述多个尺度的第一特征图。
- 4.根据权利要求1-3中任意一项所述的方法，其特征在于，所述对所述多个尺度的第一特征图执行
- 特征融合处理，得到多个尺度的第二特征图，包括：
 - 20 将所述多个尺度的第一特征图输入至特征金字塔网络，通过所述特征金字塔网络执行所述特征融
 - 合处理，得到所述多个尺度的第二特征图。
- 5.根据权利要求1-4中任意一项所述的方法，其特征在于，按照尺度从小到大的顺序，所述多个第
- 一特征图表示为 $\{C_1, \dots, C_n\}$ ，其中， n 表示第一特征图的数量， n 为大于1的整数；
 - 所述对所述多个尺度的第一特征图执行特征融合处理，得到多个尺度的第二特征图，包括：
 - 25 利用第一卷积核对第一特征图 C_n 执行卷积处理，获得与所述第一特征图 C_n 对应的第二特征图
 - F_n ，其中，所述第一特征图 C_n 的尺度与所述第二特征图 F_n 的尺度相同；
 - 对所述第二特征图 F_n 执行线性插值处理获得与所述第二特征图 F_n 对应的第一中间特征图 F'_n ，其
 - 中，所述第一中间特征图 F'_n 的尺度与第一特征图 C_{n-1} 的尺度相同；
 - 利用第二卷积核对所述第一特征图 C_n 以外的第一特征图 C_i 执行卷积处理，得到所述第一特征图
 - 30 C_i 对应的第二中间特征图 C'_i ，所述第二中间特征图 C'_i 的尺度与第一中间特征图 F'_{i+1} 的尺度相
 - 同，其中， i 为大于或者等于1且小于 n 的整数变量；
 - 利用所述第二中间特征图 C'_i 和对应的所述第一中间特征图 F'_{i+1} 得到所述第二特征图
 - F_n 以外的第二特征图 F_i ，其中，所述第一中间特征图 F'_{i+1} 由对应的所述第二特征图 F_{i+1} 经线性插值得
 - 到。
- 6.根据权利要求5所述的方法，其特征在于，所述利用所述第二中间特征图 C'_i 和对应的所

述第一中间特征图 $F_{i+1}^{'}$ 得到所述第二特征图 F_n 以外的第二特征图 F_i ，包括：

将所述第二中间特征图 $C_i^{'}$ 与对应的所述第一中间特征图 $F_{i+1}^{'}$ 进行加和处理，得到所述第二特征图 F_i 。

7. 根据权利要求1-6中任意一项所述的方法，其特征在于，所述基于得到的所述多个尺度的第二特征图检测所述第一图像中针对同一人物对象的关联的人脸位置和人手位置，包括：

对所述多个尺度的第二特征图中尺度最大的第二特征图执行卷积处理，分别得到表示所述人脸位置的掩码图，以及表示所述人手位置的掩码图；

基于所述人脸位置的掩码图以及所述人手位置的掩码图确定所述第一图像中关联的人手和人脸所在的位置区域。

8. 根据权利要求1-7中任意一项所述的方法，其特征在于，所述方法还包括以下方式中的至少一种：

在所述第一图像中突出显示所述关联的人手和人脸；

为所述第一图像中检测到的关联的人脸位置和人手位置分配相同的标签。

9. 根据权利要求1-8中任意一项所述的方法，其特征在于，所述方法通过神经网络实现，其中，训练所述神经网络的步骤包括：

获取训练图像，所述训练图像为包括人物对象的图像，所述训练图像具有真实关联的人脸位置和人手位置的标注信息；

将所述训练图像输入至所述神经网络，通过所述神经网络预测所述训练图像中针对同一人物对象的关联的人脸位置和人手位置；

- 基于预测出的关联的所述人脸位置以及人手位置以及所述标注信息确定网络损失，并根据所述网络损失调整所述神经网络的网络参数，直至满足训练要求。

10. 一种人脸和人手关联检测装置，其特征在于，包括：

获取模块，用于获取第一图像，所述第一图像为包括人物对象的图像；

特征提取模块，用于对所述第一图像执行特征提取，得到多个尺度的第一特征图；

- 融合模块，用于对所述多个尺度的第一特征图执行特征融合处理，得到多个尺度的第二特征图，所述第二特征图的尺度与所述第一特征图的尺度一一对应；

检测模块，用于基于得到的所述多个尺度的第二特征图检测所述第一图像中针对同一人物对象的关联的人脸位置和人手位置。

11. 根据权利要求10所述的装置，其特征在于，所述获取模块包括：

- 获取单元，用于获取第二图像，所述第二图像为包括至少一个物体对象的图像；

目标检测单元，用于对所述第二图像执行人体目标检测，得到所述第一图像中所述至少一个物体对象中任一物体对象的检测框；

确定单元，用于将所述任一物体对象的所述检测框在所述第二图像中对应的图像区域，确定为所述任一物体对象的第一图像。

12. 根据权利要求10或11所述的装置，其特征在于，所述特征提取模块还用于将所述第一图像调整为预设尺度的第三图像；

将所述第三图像输入至残差网络，得到所述多个尺度的第一特征图。

13. 根据权利要求10-12中任意一项所述的装置，其特征在于，所述融合单元还用于将所述多个尺度的第一特征图输入至特征金字塔网络，通过所述特征金字塔网络执行所述特征融合处理，得到所述多个尺度的第二特征图。

14. 根据权利要求10-13中任意一项所述的装置，其特征在于，按照尺度从小到大的顺序，所述多

个第一特征图表示为 $\{C_1, \dots, C_n\}$ ，其中， n 表示第一特征图的数量， n 为大于1的整数；

所述融合模块还用于利用第一卷积核对第一特征图 C_n 执行卷积处理，获得与所述第一特征图 C_n 对应的第二特征图 F_n ，其中，所述第一特征图 C_n 的尺度与所述第二特征图 F_n 的尺度相同；

对所述第二特征图 F_n 执行线性插值处理获得与所述第二特征图 F_n 对应的第一中间特征图 F'_n ，其中，
5 中，所述第一中间特征图 F'_n 的尺度与第一特征图 C_{n-1} 的尺度相同；

利用第二卷积核对所述第一特征图 C_n 以外的第一特征图 C_i 执行卷积处理，得到所述第一特征图 C_i 对应的第二中间特征图 C'_i ，所述第二中间特征图 C'_i 的尺度与第一中间特征图 F'_{i+1} 的尺度相同，其中， i 为大于或者等于1且小于 n 的整数变量；

利用所述第二中间特征图 C'_i 和对应的所述第一中间特征图 F'_{i+1} 得到所述第二特征图
10 F_n 以外的第二特征图 F_i ，其中，所述第一中间特征图 F'_{i+1} 由对应的所述第二特征图 F_{i+1} 经线性插值得到。

15. 根据权利要求14所述的装置，其特征在于，所述融合模块还用于将所述第二中间特征图 C'_i 与对应的所述第一中间特征图 F'_{i+1} 进行加和处理，得到所述第二特征图 F_i 。

16. 根据权利要求10-15中任意一项所述的装置，其特征在于，所述检测模块还用于对所述多个尺度的第二特征图中尺度最大的第二特征图执行卷积处理，分别得到表示所述人脸位置的掩码图，以及表示所述人手位置的掩码图；

基于所述人脸位置的掩码图以及所述人手位置的掩码图确定所述第一图像中关联的人手和人脸所在的位置区域。

17. 根据权利要求10-16中任意一项所述的装置，其特征在于，所述装置还包括显示模块和分配模块中的至少一种，其中

所述显示模块，用于在所述第一图像中突出显示所述关联的人手和人脸；

所述分配模块，用于为所述第一图像中检测到的人脸位置和人手位置分配相同的标签。

18. 根据权利要求10-17中任意一项所述的装置，其特征在于，所述装置包括神经网络，所述特征提取模块、所述融合模块和所述检测模块应用所述神经网络，

25 所述装置还包括训练模块，用于训练所述神经网络，其中，训练所述神经网络的步骤包括：

获取训练图像，所述训练图像为包括人物对象的图像，所述训练图像具有真实关联的人脸位置和人手位置的标注信息；

将所述训练图像输入至所述神经网络，通过所述神经网络预测所述训练图像中针对同一人物对象的关联的人脸位置和人手位置；

30 基于预测出的关联的所述人脸位置以及人手位置以及所述标注信息确定网络损失，并根据所述网络损失调整所述神经网络的网络参数，直至满足训练要求。

19. 一种电子设备，其特征在于，包括：

处理器；

用于存储处理器可执行指令的存储器；

其中，所述处理器被配置为调用所述存储器存储的指令，以执行权利要求1至9中任意一项所述的方法。

20. 一种计算机可读存储介质，其上存储有计算机程序指令，其特征在于，所述计算机程序指令被处理器执行时实现权利要求1至9中任意一项所述的方法。

- 5 21. 一种计算机程序，其特征在于，所述计算机程序包括计算机可读代码，当所述计算机可读代码在电子设备中运行时，所述电子设备中的处理器执行用于实现权利要求1至9中的任意一项所述的方法。