

权利要求书

- 1、一种目标跟踪方法，包括：
获取视频图像；
针对除所述视频图像中的参考帧图像之后的待跟踪图像，生成所述待跟踪图像中的
5 搜索区域与所述参考帧图像中的目标图像区域之间的图像相似性特征图；其中，所述目
标图像区域内包含待跟踪对象；
根据所述图像相似性特征图，确定所述搜索区域中的待定位区域的定位位置信息；
响应于在所述搜索区域中确定出所述待定位区域的定位位置信息，根据确定的待定
10 位区域的定位位置信息，确定所述待跟踪对象在包含所述搜索区域的待跟踪图像中的检
测框。
- 2、根据权利要求1所述的目标跟踪方法，其中，根据所述图像相似性特征图，确定
所述搜索区域中的待定位区域的定位位置信息，包括：
根据所述图像相似性特征图，预测所述待定位区域的尺寸信息；
根据所述图像相似性特征图，预测所述搜索区域的特征图中的每个特征像素点的概
15 率值，一个特征像素点的概率值表征所述搜索区域中与该特征像素点对应的像素点位于
所述待定位区域内的几率；
根据所述图像相似性特征图，预测所述搜索区域中与每个所述特征像素点对应的像
素点与所述待定位区域的位置关系信息；
从预测的概率值中选取所述概率值最大的特征像素点对应的所述搜索区域中的
20 像素点作为目标像素点；
基于所述目标像素点、所述目标像素点与所述待定位区域的位置关系信息、以及所
述待定位区域的尺寸信息，确定所述待定位区域的定位位置信息。
- 3、根据权利要求1或2所述的目标跟踪方法，其中，根据以下步骤从所述参考帧图
像中提取所述目标图像区域：
25 确定所述待跟踪对象在所述参考帧图像中的检测框；
基于所述参考帧图像中的所述检测框的尺寸信息，确定所述参考帧图像中的所述检
测框对应的第一延伸尺寸信息；
基于所述第一延伸尺寸信息，以所述参考帧图像中的所述检测框为起始位置向周围
延伸，得到所述目标图像区域。
- 4、根据权利要求1或2所述的目标跟踪方法，其中，根据以下步骤从待跟踪图像中
30 提取搜索区域：
获取在所述视频图像中当前帧待跟踪图像的前一帧待跟踪图像中，所述待跟踪对象
的检测框；
基于所述待跟踪对象的检测框的尺寸信息，确定所述待跟踪对象的检测框对应的第
35 二延伸尺寸信息；
基于所述第二延伸尺寸信息和所述待跟踪对象的检测框的尺寸信息，确定当前帧待
跟踪图像中的搜索区域的尺寸信息；
以所述待跟踪对象的检测框的中心点为当前帧待跟踪图像中的搜索区域的中心，根
据当前帧待跟踪图像中的搜索区域的尺寸信息确定所述搜索区域。
- 5、根据权利要求1至4任一项所述的目标跟踪方法，其中，所述生成所述待跟踪图
40 像中的搜索区域与所述参考帧图像中的目标图像区域之间的图像相似性特征图，包括：
将所述搜索区域缩放至第一预设尺寸，以及，将所述目标图像区域缩放至第二预设

尺寸;

生成所述搜索区域中的第一图像特征图,以及所述目标图像区域中的第二图像特征图;所述第二图像特征图的尺寸小于所述第一图像特征图的尺寸;

确定所述第二图像特征图与所述第一图像特征图中的每个子图像特征图之间的相关性特征;所述子图像特征图与所述第二图像特征图的尺寸相同;

基于确定的多个相关性特征,生成所述图像相似性特征图。

6、根据权利要求1至5任一项所述的目标跟踪方法,其中,

所述目标跟踪方法由跟踪定位神经网络执行;其中所述跟踪定位神经网络由标注有目标对象的检测框的样本图像训练得到。

7、根据权利要求6所述的目标跟踪方法,其中,所述方法还包括训练所述跟踪定位神经网络的步骤:

获取样本图像,所述样本图像包括参考帧样本图像和待跟踪的样本图像;

将所述样本图像输入待训练的跟踪定位神经网络,经过所述待训练的跟踪定位神经网络对输入的样本图像进行处理,预测所述目标对象在所述待跟踪的样本图像中的检测框;

基于所述待跟踪的样本图像中标注的检测框,和所述待跟踪的样本图像中预测的检测框,调整所述待训练的跟踪定位神经网络的网络参数。

8、根据权利要求7所述的目标跟踪方法,其中,将所述待跟踪的样本图像中的待定位区域的定位位置信息,作为所述待跟踪的样本图像中预测的检测框的位置信息,

所述基于所述待跟踪的样本图像中标注的检测框,和所述待跟踪的样本图像中预测的检测框,调整所述待训练的跟踪定位神经网络的网络参数,包括:

基于所述预测的检测框的尺寸信息、

所述待跟踪的样本图像中搜索区域中每个像素点位于所述预测的检测框内的预测概率值、

所述待跟踪的样本图像中搜索区域中每个像素点与所述预测的检测框的预测位置关系信息、

所述标注的检测框的标准尺寸信息、

所述待跟踪的样本图像中标准搜索区域中每个像素点是否位于标注的检测框中的信息和

所述标准搜索区域中每个像素点与所述标注的检测框的标准位置关系信息,调整所述待训练的跟踪定位神经网络的网络参数。

9、一种目标跟踪装置,包括:

图像获取模块,配置为获取视频图像;

相似性特征提取模块,配置为针对除所述视频图像中的参考帧图像之后的待跟踪图像,生成所述待跟踪图像中的搜索区域与所述参考帧图像中的目标图像区域之间的图像相似性特征图;其中,所述目标图像区域内包含待跟踪对象;

定位模块,配置为根据所述图像相似性特征图,确定所述搜索区域中的待定位区域的定位位置信息;

跟踪模块,配置为响应于在所述搜索区域中确定出所述待定位区域的定位位置信息,根据确定的待定位区域的定位位置信息,确定所述待跟踪对象在包含所述搜索区域的待跟踪图像中的检测框。

10、根据权利要求9所述的目标跟踪装置,其中,所述定位模块配置为:

根据所述图像相似性特征图,预测所述待定位区域的尺寸信息;

根据所述图像相似性特征图,预测所述搜索区域的特征图中的每个特征像素点的概

率值,一个特征像素点的概率值表征所述搜索区域中与该特征像素点对应的像素点位于所述待定位区域内的几率;

根据所述图像相似性特征图,预测所述搜索区域中与每个所述特征像素点对应的像素点与所述待定位区域的位置关系信息;

- 5 从预测的概率值中选取所述概率值最大的特征像素点所对应的所述搜索区域中的像素点作为目标像素点;

基于所述目标像素点、所述目标像素点与所述待定位区域的位置关系信息、以及所述待定位区域的尺寸信息,确定所述待定位区域的定位位置信息。

- 10 11、根据权利要求9或10所述的目标跟踪装置,其中,所述相似性特征提取模块配置为利用以下步骤从所述参考帧图像中提取所述目标图像区域:

确定所述待跟踪对象在所述参考帧图像中的检测框;

基于所述参考帧图像中的所述检测框的尺寸信息,确定所述参考帧图像中的所述检测框对应的第一延伸尺寸信息;

- 15 基于所述第一延伸尺寸信息,以所述参考帧图像中的所述检测框为起始位置向周围延伸,得到所述目标图像区域。

12、根据权利要求9或10所述的目标跟踪装置,其中,所述相似性特征提取模块配置为利用以下步骤从待跟踪图像中提取搜索区域:

获取在所述视频图像中当前帧待跟踪图像的前一帧待跟踪图像中,所述待跟踪对象的检测框;

- 20 基于所述待跟踪对象的检测框的尺寸信息,确定所述待跟踪对象的检测框对应的第二延伸尺寸信息;

基于所述第二延伸尺寸信息和所述待跟踪对象的检测框的尺寸信息,确定当前帧待跟踪图像中的搜索区域的尺寸信息;

- 25 以所述待跟踪对象的检测框的中心点为当前帧待跟踪图像中的搜索区域的中心,根据当前帧待跟踪图像中的搜索区域的尺寸信息确定所述搜索区域。

13、根据权利要求9至12任一项所述的目标跟踪装置,其中,所述相似性特征提取模块配置为:

将所述搜索区域缩放至第一预设尺寸,以及,将所述目标图像区域缩放至第二预设尺寸;

- 30 生成所述搜索区域中的第一图像特征图,以及所述目标图像区域中的第二图像特征图;所述第二图像特征图的尺寸小于所述第一图像特征图的尺寸;

确定所述第二图像特征图与所述第一图像特征图中的每个子图像特征图之间的相关性特征;所述子图像特征图与所述第二图像特征图的尺寸相同;

基于确定的多个相关性特征,生成所述图像相似性特征图。

- 35 14、根据权利要求9至13任一项所述的目标跟踪装置,其中,所述目标跟踪装置利用跟踪定位神经网络确定所述待跟踪对象在包含所述搜索区域的待跟踪图像中的检测框;其中所述跟踪定位神经网络由标注有目标对象的检测框的样本图像训练得到。

15、根据权利要求14所述的目标跟踪装置,其中,所述目标跟踪装置还包括模型训练模块,配置为:

- 40 获取样本图像,所述样本图像包括参考帧样本图像和待跟踪的样本图像

将所述样本图像输入待训练的跟踪定位神经网络,经过所述待训练的跟踪定位神经网络对输入的样本图像进行处理,预测所述目标对象在所述待跟踪的样本图像中的检测框;

基于所述待跟踪的样本图像中标注的检测框和所述待跟踪的样本图像中预测的检

测框，调整所述待训练的跟踪定位神经网络的网络参数。

- 16、根据权利要求15所述的目标跟踪装置，其中，将所述待跟踪的样本图像中的待定位区域的定位位置信息作为所述待跟踪的样本图像中预测的检测框的位置信息，所述模型训练模块在基于所述待跟踪的样本图像中标注的检测框和所述待跟踪的样本图像中预测的检测框，调整所述待训练的跟踪定位神经网络的网络参数的情况下，配置为：
- 5 基于所述待跟踪的样本图像中预测的检测框的尺寸信息、所述待跟踪的样本图像中搜索区域中每个像素点位于所述待跟踪的样本图像中预测的检测框内的预测概率值、所述待跟踪的样本图像中搜索区域中每个像素点与所述待跟踪的样本图像中预测的检测框的预测位置关系信息、所述待跟踪的样本图像中标注的检测框的标准尺寸信息、所述待跟踪的样本图像中标准搜索区域中每个像素点是否位于标注的检测框中的信息、所述待跟踪的样本图像中标准搜索区域中每个像素点与所述待跟踪的样本图像中标注的检测框的标准位置关系信息，调整所述待训练的跟踪定位神经网络的网络参数。
- 10

- 17、一种电子设备，包括：处理器、存储介质和总线，所述存储介质存储有所述处理器可执行的机器可读指令，当电子设备运行时，所述处理器与所述存储介质之间通过总线通信，所述处理器执行所述机器可读指令，以执行如权利要求1至8任一所述的目标跟踪方法。
- 15

18、一种计算机可读存储介质，所述计算机可读存储介质上存储有计算机程序，所述计算机程序被处理器运行时执行如权利要求1至8任一所述的目标跟踪方法。