

权利要求书

1、一种目标对象识别方法，包括：

对待检图像的对象进行目标对象检测，获得所述对象的目标对象
预测信息，所述目标对象预测信息为检测到的对象为目标对象的置信
5 信息；

对所述待检图像的所述对象进行关键点检测，获得所述对象的关键
点预测信息；所述关键点预测信息为检测到对象的关键点为目标对
象的关键点的置信信息；

将所述目标对象预测信息以及所述关键点预测信息进行融合，获
10 得所述对象的综合预测信息；

根据所述综合预测信息对所述目标对象进行识别。

2、根据权利要求 1 所述的方法，其中，所述对待检图像的对象进
行目标对象检测，和对所述待检图像的所述对象进行关键点检测之前，
包括：

15 获取所述待检图像的对象对应的图像区域；

所述对待检图像的对象进行目标对象检测，包括：

对待检图像的对象对应的图像区域进行目标对象检测；

对所述待检图像的对象进行关键点检测，包括：

对待检图像的对象对应的图像区域进行关键点检测。

20 3、根据权利要求 1 或 2 所述的方法，其中，所述将所述目标对象
预测信息以及所述关键点预测信息进行融合，获得所述对象的综合预
测信息，包括：

将所述目标对象预测信息以及所述关键点预测信息相乘，得到所
述对象的综合预测信息。

25 4、根据权利要求 1~3 中任一项所述的方法，其中，所述对所述待

检图像的所述对象进行关键点检测，获得所述对象的关键点预测信息，包括：

利用定位关键点的神经网络模型，对所述待检图像的对象进行关键点检测，获得所述对象的关键点预测信息。

- 5 5、根据权利要求 2 至 4 中任一项所述的方法，其中，所述获取所述待检图像的对象对应的图像区域之后，所述将所述目标对象预测信息以及所述关键点预测信息进行融合，获得所述对象的综合预测信息之前，还包括：

从所述图像区域，检测所述对象的偏转角度信息；

- 10 所述将所述目标对象预测信息以及所述关键点预测信息进行融合，获得所述对象的综合预测信息，包括：

根据所述目标对象预测信息、所述关键点预测信息和所述偏转角度信息进行融合，获得所述对象的综合预测信息。

- 15 6、根据权利要求 5 所述的方法，其中，所述从所述图像区域，检测所述对象的偏转角度信息，包括：

利用对象分类的神经网络模型，从所述图像区域检测所述对象的偏转角度信息。

7、根据权利要求 1 至 6 中任一项所述的方法，其中，

所述目标对象为：人脸；

- 20 所述将所述目标对象预测信息以及所述关键点预测信息进行融合，获得所述对象的综合预测信息之前，还包括：

从所述图像区域，检测人脸俯仰角度和/或人脸侧转角度；

所述将所述目标对象预测信息以及所述关键点预测信息进行融合，获得所述对象的综合预测信息，包括：

- 25 根据适用指数函数将所述人脸俯仰角度和/或人脸侧转角度进行归一化处理；将所述目标对象预测信息、所述关键点预测信息和归一

化的人脸俯仰角度相乘，获得所述对象的综合预测信息；

或，

将所述目标对象预测信息、所述关键点预测信息和归一化的人脸侧转角度相乘，获得所述对象的综合预测信息；

5 或，

将所述目标对象预测信息、所述关键点预测信息、归一化的人脸俯仰角度和归一化的人脸侧转角度相乘，获得所述对象的综合预测信息。

8、根据权利要求 1 至 7 中任一项所述的方法，其中，所述待检图
10 像为视频帧图像；

在根据所述综合预测信息对所述目标对象进行识别之后，还包括：

根据从多个所述视频帧图像对目标对象进行识别的结果，对所述目标对象进行跟踪；

或者，

15 根据为多个所述视频帧图像各自获得的综合预测信息，从多个所述视频帧图像选择综合预测质量最高的视频帧图像作为抓拍图像；

或者，

根据为多个所述视频帧图像各自获得的综合预测信息，从多个所述视频帧图像选择预定个数的视频帧图像，对选择的视频帧图像进行
20 特征融合。

9、一种目标对象识别装置，包括：

对象检测模块，配置为对待检图像的对象进行目标对象检测，获得所述对象的目标对象预测信息，所述目标对象预测信息为检测到的对象为目标对象的置信信息；

25 关键点检测模块，配置为对所述待检图像的所述对象进行关键点检测，获得所述对象的关键点预测信息；所述关键点预测信息为检测

到对象的关键点为目标对象的关键点的置信信息；

预测信息融合模块，配置为将所述对象检测模块获得的目标对象预测信息以及所述关键点检测模块获得的关键点预测信息进行融合，获得所述对象的综合预测信息；

- 5 对象识别模块，配置为根据所述预测信息融合模块获得的综合预测信息对所述目标对象进行识别。

10、根据权利要求 9 所述的装置，其中，所述装置还包括：

图像区域获取模块，配置为获取所述待检图像的对象对应的图像区域；

- 10 所述对象检测模块，配置为对所述图像区域获取模块获取的待检图像的对象对应的图像区域进行目标对象检测；

所述关键点检测模块，配置为对所述图像区域获取模块获取的待检图像的对象对应的图像区域进行关键点检测。

- 11、根据权利要求 9 或 10 所述的装置，其中，所述预测信息融合
15 模块，配置为将所述目标对象预测信息以及所述关键点预测信息相乘，得到所述对象的综合预测信息。

12、根据权利要求 9 至 11 中任一项所述的装置，其中，所述关键点检测模块，配置为利用定位关键点的神经网络模型，对所述待检图像的对象进行关键点检测，获得所述对象的关键点预测信息。

- 20 13、根据权利要求 10 至 12 中任一项所述的装置，其中，所述获取所述待检图像的对象对应的图像区域之后，所述将所述目标对象预测信息以及所述关键点预测信息进行融合，获得所述对象的综合预测信息之前，所述装置还包括：

25 偏转角度检测模块，配置为从所述图像区域获取模块获取的图像区域，检测所述对象的偏转角度信息；

所述预测信息融合模块，配置为根据所述目标对象预测信息、所

述关键点预测信息和所述偏转角度信息进行融合，获得所述对象的综合预测信息。

14、根据权利要求 13 所述的装置，其中，所述偏转角度检测模块配置为利用对象分类的神经网络模型，从所述图像区域检测所述对象的偏转角度信息。

15、根据权利要求 9 至 14 中任一项所述的装置，其中，

所述目标对象为：人脸；

所述将所述目标对象预测信息以及所述关键点预测信息进行融合，获得所述对象的综合预测信息之前，所述装置还包括：

10 人脸偏转角度检测模块，配置为从所述图像区域，检测人脸俯仰角度和/或人脸侧转角度；

所述预测信息融合模块，配置为根据适用指数函数将所述人脸俯仰角度和/或人脸侧转角度进行归一化处理；将所述目标对象预测信息、所述关键点预测信息和归一化的人脸俯仰角度相乘，获得所述对象的综合预测信息；

或，

将所述目标对象预测信息、所述关键点预测信息和归一化的人脸侧转角度相乘，获得所述对象的综合预测信息；

或，

20 将所述目标对象预测信息、所述关键点预测信息、归一化的人脸俯仰角度和归一化的人脸侧转角度相乘，获得所述对象的综合预测信息。

16、根据权利要求 9 至 15 中任一项所述的装置，其中，所述待检图像为视频帧图像；

25 在根据所述综合预测信息对所述目标对象进行识别之后，所述装置还包括：

对象跟踪模块，配置为根据从多个所述视频帧图像对目标对象进行识别的结果，对所述目标对象进行跟踪；

或者，

抓拍图像选取模块，配置为根据为多个所述视频帧图像各自获得的综合预测信息，从多个所述视频帧图像选择综合预测质量最高的视频帧图像作为抓拍图像；

或者，

特征融合模块，配置为根据为多个所述视频帧图像各自获得的综合预测信息，从多个所述视频帧图像选择预定个数的视频帧图像，对选择的视频帧图像进行特征融合。

17、一种电子设备，包括：处理器、存储器、通信元件和通信总线，所述处理器、所述存储器和所述通信元件通过所述通信总线完成相互间的通信；

所述存储器配置为存放至少一可执行指令，所述可执行指令使所述处理器执行如权利要求 1 至 8 中任一项所述的目标对象识别方法相应的操作。

18、一种计算机可读存储介质，其上存储有计算机程序指令，其中，所述程序指令被处理器执行时实现权利要求 1 至 8 中任一项所述的目标对象识别方法的步骤。

19、一种计算机程序，其包括有计算机程序指令，其中，所述程序指令被处理器执行时实现权利要求 1 至 8 中任一项所述的目标对象识别方法的步骤。