**深 圳 大 学 实 验 报 告**

**课程名称： 智能无人系统与边缘计算**

**实验项目名称： 视频流目标检测创新应用**

**学院： 电子与信息工程学院**

**专业： 电子信息工程（文华班）**

**指导教师： 蒙山**

**报告人： 余韦藩 学号： 2020285102 班级： 文华班**

**实验时间： 2022.12.21-2023.2.18**

**实验报告提交时间： 2023年2月18日**

**教务部制**

|  |
| --- |
| **实验目的与要求：**  **Basic requirements:**   1. Innovate image target detection application, and use YOLO v5 to complete image target training and detection. 2. Get video and image data through RTSP pull stream (you can build RTSP server by yourself), and display video stream content in the window. 3. Complete the self-defined image target detection reasoning frame by frame, display the target detection results in the window, and mark different detection targets through different color boxes.   **Additional requirements:**   1. Use stream processing framework, such as GStreamer, and YOLO v5 to build a complete image target detection application. 2. Get video stream data directly from USB camera or network IP camera to complete innovative application of image target detection. 3. Use the common NVIDIA Jetson Nano, or raspberry pie, or Horizon X3 pie, or Ruixin micro-series AI computing board to develop and realize the corresponding image target detection function. |
| **Experiment process and experiment result：**  **Basic requirements:**   1. **Innovate image target detection application, and use YOLO v5 to complete image target training and detection.**   **Experiment process:** Here we write a LoadImages class object to deal with the image input or video input as shown in the following figure. We have write some notes in the program so we don’t explain it in detail. The main program part of the LoadImages is judge whether the input is image or video. Then save all the frames in the class object, resize them and convert to the format of RGB. Combined with the yolo-v5 program, we achieve perform object detection for image or video.  微信截图_20230218093735  **Experiment result:**  The object detection for example image is shown in the following figure. Here we observe the inter time is tiny. What’s more, I achieve object detection for mp4 video. The original mp4 video(MP4原视频) and the process mp4 video(MP4视频检测) have been attached to this report.  微信图片_20230218001101微信图片_20230218001108   1. **Get video and image data through RTSP pull stream (you can build RTSP server by yourself), and display video stream content in the window.**   **Experiment process:** First, we should program to achieve rtsp pull stream. The core is to extract each frame in the video stream and display it in the screen. The program is shown in the following. Some notes are marked in it so we will not explain it in detail. It is worth to say that we use a yolo signal variance to control the output from original video stream to processed(object detection) video stream. What’s more, in order to display normal for the cv2 format image, we should convert it to the format of RGB before display it. The output of the video stream is equivalent to display each frame in the video stream. In order to play the video stream that we created, we should modify the corresponding rtsp address.  微信截图_20230218093528  微信截图_20230218093541  微信截图_20230218093558  In addition, we should build a rtsp server. Here we use the VLC software to generate our own rtsp server. All the steps are shown in the following figure.  微信截图_20230218092855微信截图_20230218092916  微信截图_20230218093006微信截图_20230218092932微信截图_20230218092946 微信截图_20230218093018  微信截图_20230218093029 微信截图_20230218093042  微信截图_20230218093052  After finish all the two process, we achieve get video and image data through RTSP pull stream.  **Experiment result:** The original mp4 video(rtsp原视频) and the pull video stream(rtsp拉流) have been attached to this report. Here is a screen shoot for the pull video stream.  **微信截图_20230218092711**   1. **Complete the self-defined image target detection reasoning frame by frame, display the target detection results in the window, and mark different detection targets through different color boxes.**   **Experiment process:** We find out the part of extracting each frame in the video stream program. If we want to perform object detection for the video stream, we do perform object detection for each frame. So before display the original frame, we use yolo-v5 model to perform object detection for each frame and output each processed frame in the screen. The following shows the corresponding program for each frame.  微信截图_20230218093945  Here we write a LoadStreams class object to deal with the video stream input as shown in the following figure. We have write some notes in the program so we don’t explain it in detail. The main program part of the LoadStreams is read each frame from video stream and save them in the class object.  微信截图_20230218093802  **Experiment result:** The original mp4 video(rtsp原视频) and the processed pull video stream(rtsp拉流检测) have been attached to this report. Here is a screen shoot for the processed pull video stream. From the figure, we observe that the frame has been marked by different color of rectangle which indicates our successfully performance. However, the video stream may unstable sometime and appears drop frame phenomenon. I support the reason of this problem may be related to the cv2 package.  **微信截图_20230218090445**  **微信截图_20230218090530**  **Additional requirements:**   1. **Get video stream data directly from USB camera or network IP camera to complete innovative application of image target detection.**   **Experiment process:** I purchase a USB camera from online to achieve this task. The following figure shows the camera.  **506fe2956de584d888b088717f9db47**  Here we write a Loadwebcam class object to deal with the external device input as shown in the following figure. We have write some notes in the program so we don’t explain it in detail. Combined with the yolo-v5 program, we achieve performing object detection for stream.  微信截图_20230218093749  **Experiment result:** The processed video stream(USB摄像头检测) have been attached to this report. Here is a screen shoot for the processed pull video stream. With the movement of the camera, the real-time object detection is performed smoothly which proves that our success.  **微信截图_20230218092750**   1. **Use stream processing framework, such as GStreamer, and YOLO v5 to build a complete image target detection application.**   **Experiment process:** I have been finished this task. However, last week my SSD fail to work and all the data in it including the former finished experiment report and all the result has been lost. If I finish this task again, I should reset the environment of QT and Gstreamer. It is too time consuming that I can’t finish it on time. So in the rest of the part, I will briefly introduce the idea to achieve this task.  In the former experiment six, we have the basic of achieving extracting each frame under Gstreamer framework(in this experiment, I successfully achieve all the task) as shown in the following figure.  微信截图_20230218105547微信截图_20230218105600  Let us look back to the idea. Each frame of the video stream is pass through pad, so we continuous extract each frame through pad and save them in a queue as shown in the following figure. Then after adding caption, we display them continuous by cv2 window. In order to achieve this task, we only replace caption operation to perform yolo-v5 detection for each frame. Save each processed frame in the queue and display it in original rate.  微信截图_20230218105723微信截图_20230218105629  Here we should search yolo-v5 C++ version and perform it according to the following figure.  微信截图_20230218111218  **Experiment result:** The result is similar to the experiment six. It is shame to unable to display the original experiment result.  微信截图_20230218105616 |
| **Experiment conclusion：**  Through this experiment, I deeply understand Yolov5 neural network from the following perspectives: network structure, run and application as shown in the following figure.  微信截图_20230217222522  First, we use the yolo-v5 to perform object detection for a given image. In addition, I understand the rtsp protocol which is often used to pull the video stream. Since the video consists of sequential images, the rest task require us to perform object detection for video stream. In this experiment, we extract each frame in the rtsp stream and perform object detection by yolo-v5. Moreover, we collect the external USB capture video stream and also perform object detection for each frame. The object recorded video is saved in convince to observe. Finally, we use the gstreamer framework to transmit the video stream and perform object detection for each frame. All the task is well finish and the corresponding result has been attached to the report. |
| **指导教师批阅意见：**    **成绩评定：**  **指导教师签字：**  **年 月 日** |
| **备注：** |

注：1、报告内的项目或内容设置，可根据实际情况加以调整和补充。

2、教师批改学生实验报告时间应在学生提交实验报告时间后10日内。