Detección de actividad en vídeos de tienda

Sistema funcional

Autores:

Jairo Carrillo Huélamo

Daniel Hernández Puerto

Águeda Sierra Carreto

Índice

1. Introducción

En esta fase introductoria se explicará el funcionamiento completo del sistema, cómo se podrá utilizar y las referencias a Github y DockerHub.

Según los requerimientos del cliente se debe desarrollar un sistema capaz de extraer información de una serie de vídeos diarios y proporcionar cuántos clientes entran, salen o pasan de largo de la tienda. Debido a esto, para la elaboración del sistema se ha utilizado una red YOLOv5 para la detección de las personas y un algoritmo propio para diferenciar el estado del cliente dentro del vídeo. Para finalizar, se le proporcionan al cliente los resultados del vídeo analizado.

1.1. Descripción de la instalación

Para ello es necesario tener instalada la aplicación de Docker en la máquina donde se vaya a realizar el despliegue del sistema. Una vez hecho esto, el siguiente paso es descargar la imagen docker redom69/aiva-2022-grupo-g que se encuentra alojada en DockerHub. Esto se puede hacer por línea de comandos o a través de la aplicación de Docker Desktop (MacOs y Windows) ejecutando el comando:

docker pull redom69/aiva-2022-grupo-g:aiva-grupog

El siguiente paso es lanzar esta imagen en un contenedor:

docker run --name ComVision redom69/aiva-2022-grupo-g:aiva-grupog

Por último, se puede comprobar que el contenedor se está ejecutando con:

docker ps

1.2. Enlaces a GitHub y DockerHub

La explicación del funcionamiento y todo lo referente al sistema se encuentra en el fichero readme.md del repositorio de GitHub: https://github.com/redom69/AIVA_2022-Grupo-G.

Por otro lado, la imagen está en: https://hub.docker.com/r/redom69/aiva-2022-grupo-g/tags

2. Despliegue

En la Figura 1 se muestra el diagrama UML que describe el sistema.

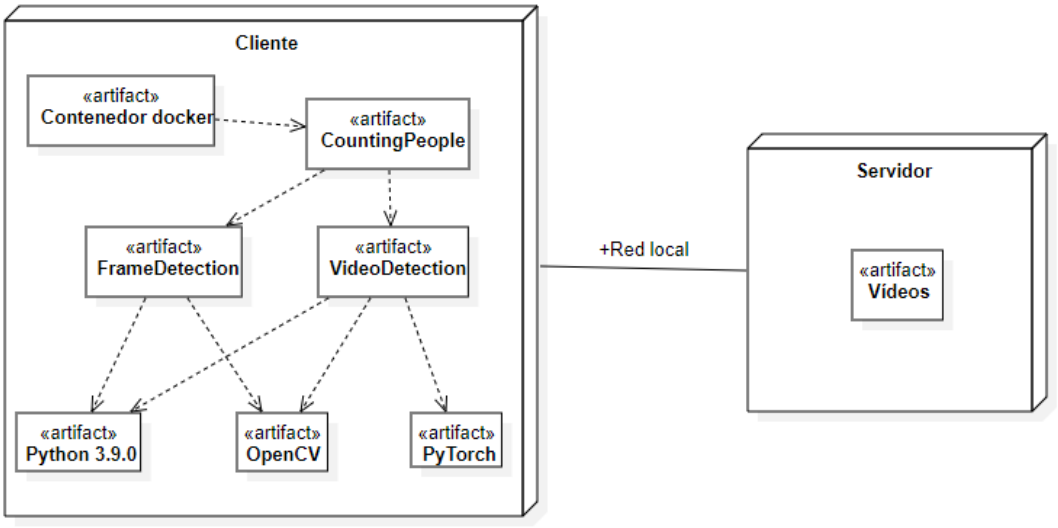

Figura 1: Diagrama de despliegue

Como se puede observar en el diagrama hay dos nodos. Uno de ellos es el servidor en el que se guardan los vídeos procedentes de la tienda. El otro nodo es el cliente que se conecta mediante una red local al servidor para obtener los vídeos a analizar.

En el nodo del cliente hay diversos artefactos. El principal es el contenedor en el que se encuentra el programa para hacer la detección (*CountingPeople*). Este programa cuenta con dos componentes: *FrameDetection* se encarga de hacer el conteo de las personas de cada grupo (entran, se paran o pasan de largo) en un *frame* y *VideoDetection* se encarga de juntar los resultados obtenidos con cada *frame* y devolver el resultado global. Ambos componentes utilizan la versión de Python 3.9.0 y OpenCV. *VideoDetection* además utiliza PyTorch para hacer la detección de las personas.

3. Funcionamiento del sistema

En la Figura 2 se representa un diagrama de actividad que explica cómo funciona el sistema.

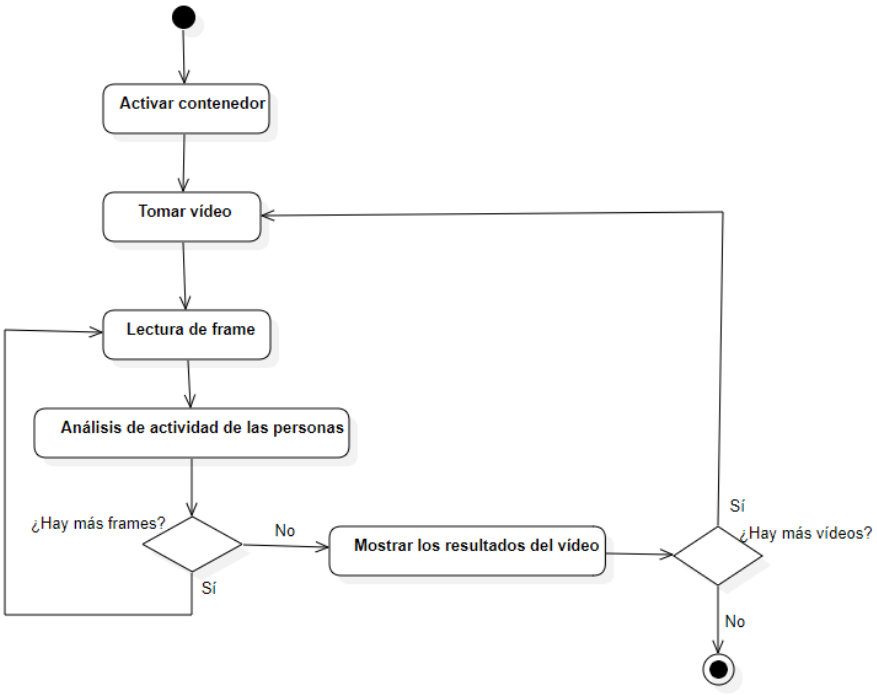

Figura 2: Diagrama de actividad del funcionamiento del sistema

El primer paso es que el directivo de la tienda active el contenedor para empezar a ejecutar el programa. Una vez activado, lo primero que hace el sistema es acceder al servidor en el que se almacenan los vídeos procedentes de las cámaras y que se encuentra en la misma red local. De esta manera obtiene el primer vídeo a analizar y, a continuación, el programa lee el primer *frame*. En él realiza la detección de las personas y luego hace el conteo del número de personas que pertenecen a cada grupo. Una vez terminada la detección y el conteo de este *frame*, pasa al siguiente con el que hace lo mismo. Esto lo va haciendo con todos los *frames* hasta que llega al último. Una vez terminado, ya se obtiene el resultado global del vídeo, se muestra por pantalla y pasa al siguiente vídeo.

Este proceso se va repitiendo hasta que termina con todos los vídeos obtenidos del servidor.

En la Figura 3 se muestra un ejemplo de la presentación de los resultados.

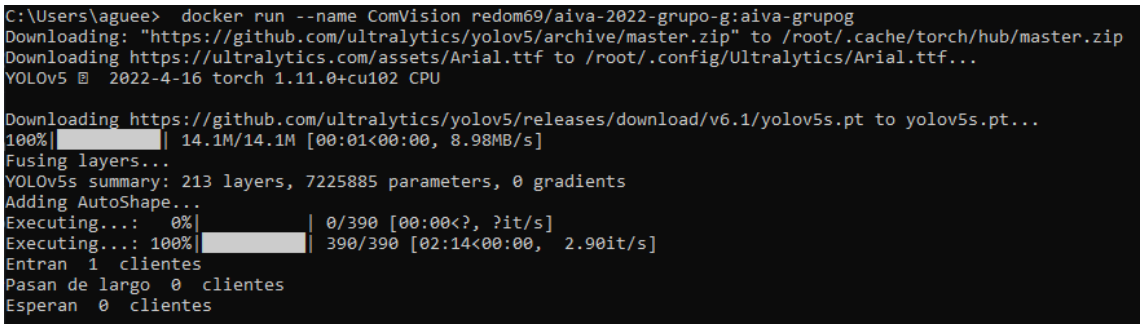

Figura 3: Resultados obtenidos con un vídeo

Lo primero que se ve es que se cargan los pesos de la red YOLOv5 para hacer la detección. Después de cargarlos se empieza a analizar el vídeo y mientras, se muestra el porcentaje del vídeo analizado mediante una barra de progreso. Una vez terminado, aparecen los contadores obtenidos para cada uno de los grupos.

Por último, en la Figura 3 aparece que se ha utilizado la CPU, sin embargo, existe la posibilidad de emplear la GPU para acelerar el procesado del vídeo.

Actualmente no se tiene acceso al servidor que almacena los vídeos de la tienda. Debido a esto, para comprobar el correcto funcionamiento del sistema, se ha incluido la carpeta de vídeos de prueba en la propia imagen de docker. Por tanto, para que el sistema funcione como se ha descrito anteriormente, solo habría que cambiar la ruta para acceder al servidor y en esa imagen de docker definitiva no habría ninguna carpeta con los vídeos.