

Detección y control de aforo en ambientes del Centro de la Industria, la Empresa y los Servicios, implementando YOLO V5, Deep SORT.

AUTORES

Diego Camilo Celada Lozada, Ingeniero electrónico, Magister en control Industrial, Estudiante doctorado en modelación y computación científica Instructor SENA Huila, dcelada@sena.edu.co

Julián René Chaux, Ingeniero electrónico, Magister en control Industrial, Instructor SENA Huila. jchaux@sena.edu.co

Karol Johana Zambrano Cruz, Matemático, Magister en Educación, Estudiante doctorado en modelación y computación científica Instructora SENA Huila, kzambrano@sena.edu.co

Proyecto realizado con el objetivo de controlar el aforo de aprendices en los espacios comunes del Centro de la Industria, la Empresa y los Servicios de la Regional Huila, como medida de bioseguridad, implementando Machine Learning.

1. Resumen

El virus COVID-19 es una enfermedad que se propaga rápidamente amenazando a personas independientemente del lugar de procedencia. Frente a esta necesidad, el ser humano consciente de la problemática toma medidas de distanciamiento pues el contacto es el principal foco de contagio. El distanciamiento social es la medida adecuada para la protección. Según la OMS, para practicar actividades sociales las personas deben distanciarse 1m entre sí. Por lo tanto, se pretende reconocer la cantidad de aprendices que ingresan al Centro de la Industria, la Empresa y los Servicios y control de aforo de personas permitidas en los ambientes de aprendizaje, teniendo en cuenta el límite de personas por espacio de distanciamiento, implementando la detección de personas con YOLOv5 el sistema utiliza la detección de objetos para distinguir las personas en un video e imágenes continuamente, el modelo de detección distingue a los grupos de aprendices que utilizan datos de cuadros delimitadores y asignando un ID. Las distancias por pares del centroide de los cuadros delimitadores distinguidos de los individuos se resuelven utilizando el método euclidiano distancia. Para detectar la infracción de la distancia entre individuos, implementa estimación de la distancia real al píxel y configuración, borde. Se establece un margen de infracción para evaluar si la distancia la estima rompe el límite de distancia social básico. El marco asegura mayor velocidad de derivación y es apto para transmitir resultados en tiempo real sin perder precisión. Se implementa el algoritmo Deepsort para controlar las personas permitas por capacidad del ambiente.

Palabras clave: Detección imagen, control aforo, Yolo V5, Deepsort

2. Abstract

The COVID-19 virus is a rapidly spreading disease threatening people regardless of where they come from. Faced with this need, the human being aware of the problem takes distancing measures since contact is the main source of contagion. Social distancing is the appropriate measure for protection. According to the WHO, to practice social activities people must distance themselves 1m from each other. Therefore, it is intended to recognize the number of apprentices who enter the Center for Industry, Business and Services and control the capacity of people allowed in the learning environments, taking into account the limit of people per distance space, implementing YOLOv5 people detection system uses object detection to distinguish people in video and images continuously, the detection model distinguishes groups of learners using bounding box data and assigning an ID. Paired distances from the centroid of the distinguished bounding boxes of individuals are solved using the Euclidean distance method.

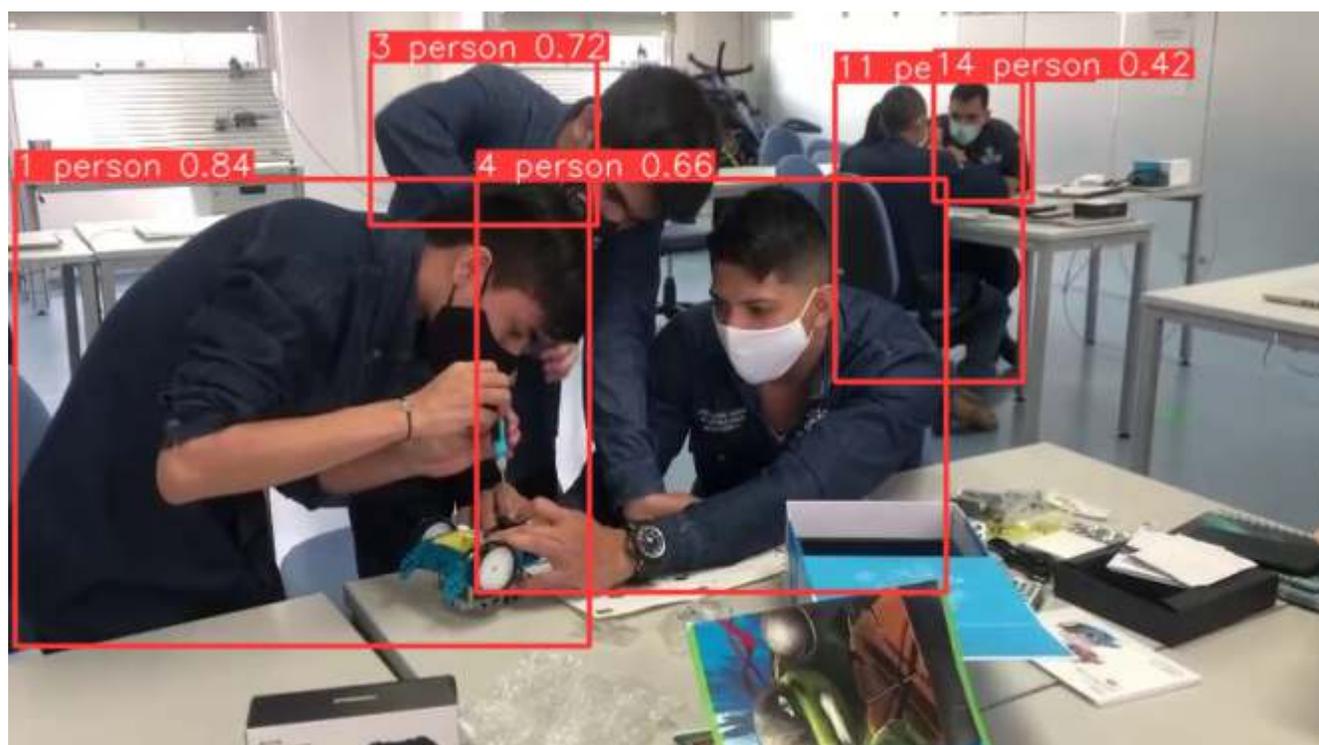
To detect the violation of the distance between individuals, it implements estimation

of the actual distance to the pixel and setting, edge. A margin of infringement is established to evaluate if the distance estimated breaks the basic social distance limit. The frame ensures higher derivation speed and is able to transmit results in real time without losing precision. The Deepsor algorithm is implemented to control the people you allow by environment capacity

Keywords: Image detection, capacity control, Yolo V5, Deepsor

3. Introducción

Actualmente, la humanidad enfrenta un momento de caos debido a una epidemia. La pandemia de COVID-19, popularmente conocida como el coronavirus, es una pandemia mundial en curso derivada de la enfermedad ocasionada por el virus SARS-CoV-2 (Gorbalenya et.all, 2020). Colombia es uno de los países afectados por este virus, al igual que el departamento del Huila en donde se han ocasionados problemas en la sociedad en diferentes aspectos, propiciando en las personas a la búsqueda de soluciones y a solventar situaciones en el diario vivir, para intentar reducir el impacto.



Uno de los aspectos que se obligó a cambiar sus condiciones son las condiciones de movilidad y distanciamiento, con el objetivo de mantener la seguridad entre los grupos de personas. Por este motivo el Centro de la Industria, la Empresa y los Servicios de la Regional Huila. Sin embargo, ha sido inevitable tener que realizar cambios en aspectos tan básicos como la movilidad en aras de la seguridad en distanciamiento frente al contacto con diferentes personas. De esta manera se encuentra la necesidad de realizar un control de aforo en el centro de formación, teniendo en cuenta que los ambientes de aprendizaje tienen una capacidad por metros cuadrados según las personas matriculas, y además, tiene zonas comunes que al ser compartidas, se deben tener en cuenta la distribución de horarios o seguimiento a los educandos para que no excedan los aforos permitidos, de esta manera se procede a plantear un proyecto que permita identificar a los aprendices en el momento de ingreso al ambiente de aprendizaje, reconocerlos e identificarlos y de esta forma poder encontrar el lugar de ubicación dentro del centro de formación, que permita al instructor tener conocimiento de su movilidad dentro del centro de formación y de esta forma tener conocimiento de los espacios que se encuentran con el aforo permitido sin entrar a romper el esquema de bioseguridad y controlar las medidas de aislamientos.

Los modelos más usados suelen ser costosos, o consisten en un contador humano en los accesos al recinto contabilizando las personas que entran y salen. Por todo esto, se planteó crear un sistema en tiempo real de control de personas en espacios cerrados que fuese asequible y automático.

El objetivo general es detectar y controlar el aforo en ambientes del CIES Neiva Implementando Yolo V5 y DeepSort.

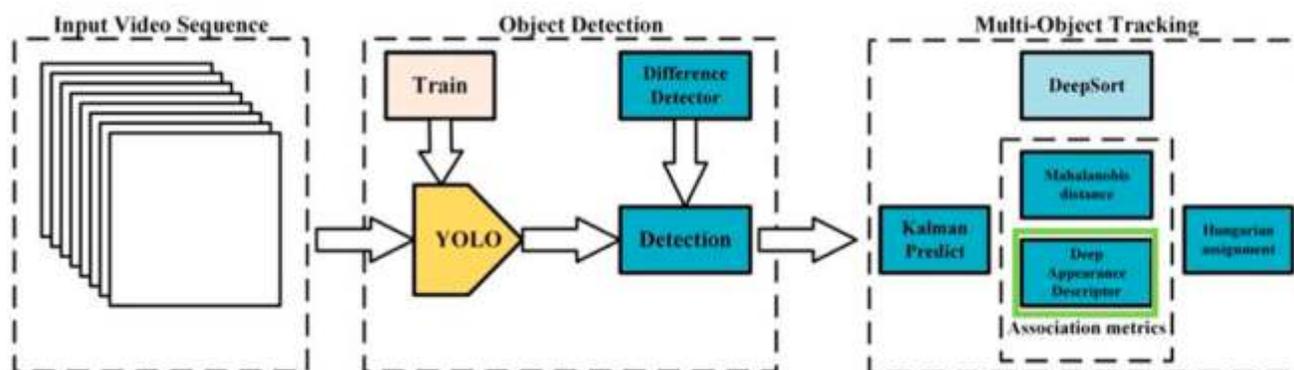
4. Materiales y métodos

Arquitectura

La arquitectura básica para el desarrollo del proyecto se muestra en la siguiente imagen, y se basa en tres bloques que se describen a continuación.

- Secuencia de video de entrada: Es la fuente del dispositivo de adquisición, que puede ser una fuente de video en tiempo real o un video guardado en el computador.
- Detección de objetos: Utilizando YOLO V5, que es una familia de arquitecturas y modelos de detección de objetos preentrenados en el conjunto de datos COCO, que utiliza el framework Pytorch. (yolov5, 2021).
- Seguimiento Multi-Objetos: Utilizando DeepSORT, que es el framework de seguimiento de objetos más popular y uno de los más utilizados y elegantes (Maiya, 2019).

Figura 6: Script



Fuente: Kanjee, 2020

Para la validación de modelo YOLOv5, se utiliza un conjunto de datos COCO val o testdev. Los modelos se descargan automáticamente desde la última versión de YOLOv5. Tenga en cuenta que las métricas de pycocotools pueden ser ~ 1% mejores que las métricas de repositorios equivalentes, debido a ligeras diferencias en el cálculo de mAP. Para el entrenamiento del modelo se puede obtener modelos previamente entrenados, que se descargan automáticamente desde la última versión de YOLOv5, los conjuntos de datos disponibles para la descarga automática incluyen:

COCO, COCO128, VOC, Argoverse, VisDrone, GlobalWheat, xView, Objects365, SKU-110K.

La otra opción es entrenar el modelo con datos personalizados con Roboflow, el cual permite organizar, etiquetar y preparar fácilmente un conjunto de datos de alta calidad con sus propios datos personalizados. Roboflow también facilita el establecimiento de un canal de aprendizaje activo, colabora con tu equipo en la mejora del conjunto de datos e integra directamente en tu flujo de trabajo de construcción de modelos con el paquete roboflow pip.

4.2 Microsoft COCO

El conjunto de datos MS COCO (Microsoft Common Objects in Context) es un conjunto de datos a gran escala de detección de objetos, segmentación, detección de puntos clave y subtítulos. (COCO, 2021)

El conjunto de datos MS COCO tiene las siguientes características:

- Detección de objetos: cuadros delimitadores y máscaras de segmentación por instancia con 80 categorías de objetos
- Subtitulado: descripciones en lenguaje natural de las imágenes (véase MS COCO Captions),

- Detección de puntos clave: contiene más de 200.000 imágenes y 250.000 instancias de personas etiquetadas con puntos clave (17 posibles puntos clave, como ojo izquierdo, nariz, cadera derecha, tobillo derecho),

- Segmentación de imágenes de objetos: máscaras de segmentación por píxel con 91 categorías de cosas, como hierba, pared, cielo (véase MS COCO Stuff),

- Panorámico: segmentación completa de la escena, con 80 categorías de cosas (como persona, bicicleta, elefante) y un subconjunto de 91 categorías de cosas (hierba, cielo, carretera),

- Densidad de posiciones: más de 39.000 imágenes y 56.000 instancias de personas etiquetadas con anotaciones DensePose - cada persona etiquetada está anotada con un id de instancia y un mapeo entre los píxeles de la imagen que pertenecen al cuerpo de esa persona y un modelo 3D de plantilla. Las anotaciones están disponibles públicamente sólo para las imágenes de entrenamiento y validación. (Lin, y otros, 2015)

4.3 Deep SORT

Deep SORT es el framework de seguimiento de objetos más popular y uno de los más utilizados y elegantes en el estado del arte. Deep SORT es la evolución del algoritmo SORT - Simple Online Real-time Tracking, el cual es el método de detección de objetivos basado en Faster R-CNN, y el uso del algoritmo de filtro Kalman + algoritmo húngaro, que mejora en gran medida la velocidad del seguimiento de múltiples objetivos. Consta de 4 componentes básicos que son: (Wojke, Bewley, & Paulus, 2017) Detección, Estimación, Asociación, y Seguimiento de la creación y destrucción de la identidad. El núcleo son dos algoritmos: Filtro de Kalman y Algoritmo húngaro

Algoritmo de filtro de kalman

Dividido en dos procesos, pronóstico y actualización. El algoritmo define el estado de movimiento del objetivo como 8 vectores distribuidos normalmente.

- Predicción: cuando el objetivo se mueve, el fotograma de destino y la velocidad del fotograma anterior se utilizan para predecir la posición del fotograma de destino y la velocidad del fotograma actual.

- Actualización: el valor pronosticado y el valor observado, los dos estados de distribución normal se ponderan linealmente para obtener el estado actual de la predicción del sistema.

Asumiendo un modelo de velocidad constante y una distribución gaussiana, el algoritmo de Kalman puede estimar dónde está el objeto basándose en el modelo de su movimiento. Y la mejor parte del filtro de Kalman es que es

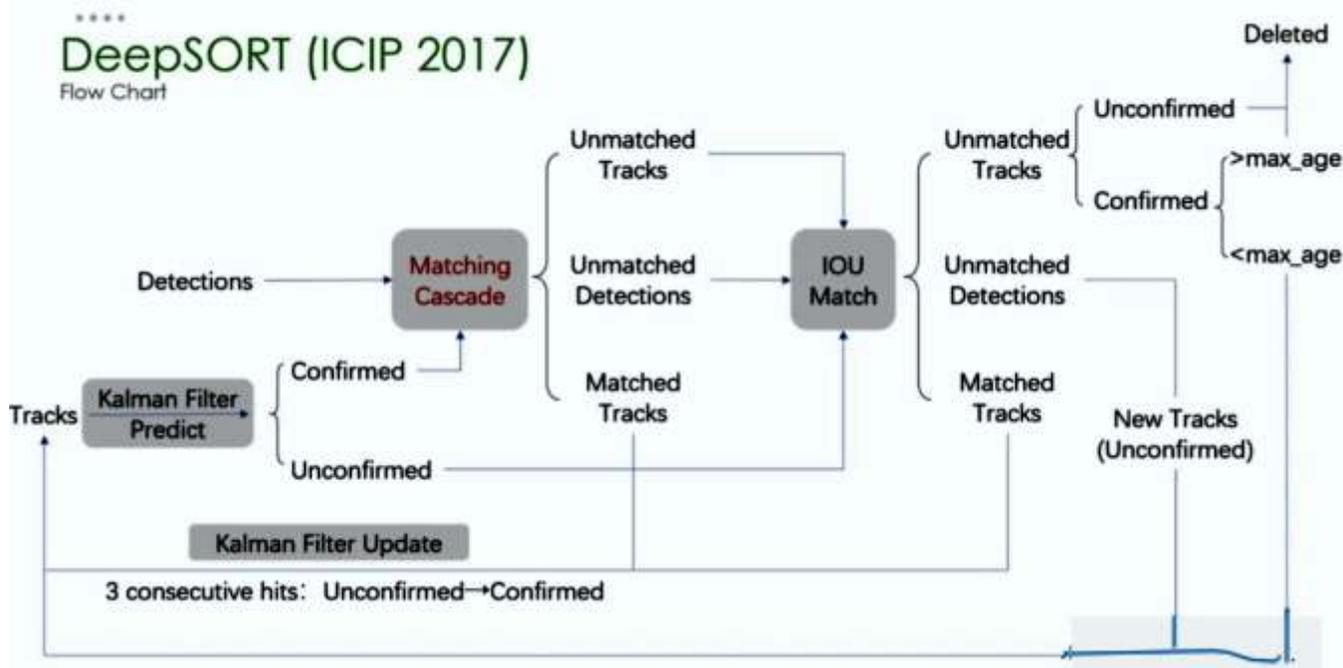
recursivo, lo que significa que tomamos las lecturas actuales, para predecir el estado actual, a continuación, utilizar las mediciones y actualizar nuestras predicciones.

Algoritmo Húngaro

Calcula la similitud y se obtiene la matriz de similitud de los dos fotogramas antes y después. El algoritmo húngaro resuelve el objetivo de coincidencia real de los dos fotogramas antes y después de resolver esta matriz de similitud. Esta parte de la biblioteca sklearn tiene una función correspondiente `linear_assignment` para resolver. (programador clic, 2021)

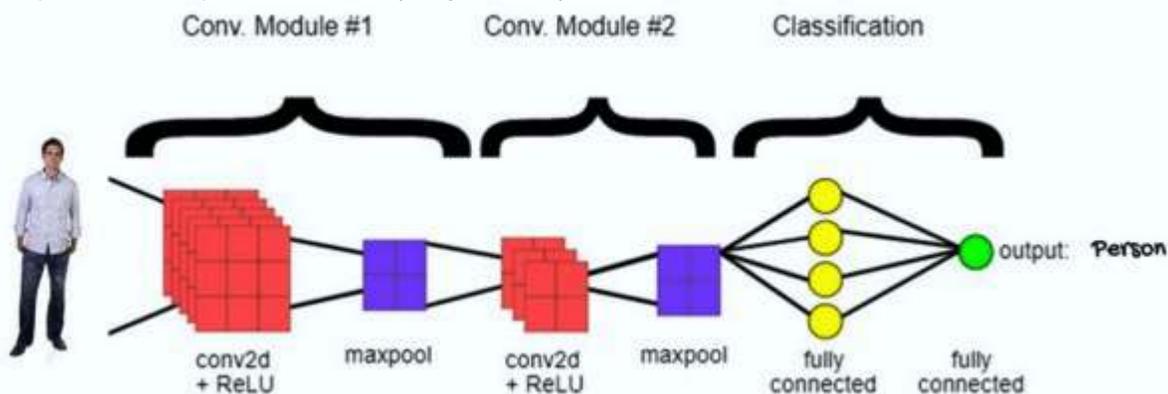
La característica más importante de DeepSort es integrar la Información de apariencia del objeto, usando el modelo de dominio ReID para extraer características, reduciendo el número de interruptores de ID. El diagrama de flujo general es el siguiente:

Figura 8: Arquitectura Deep SORT (programador clic, 2021)



Deep SORT construye un clasificador basado en nuestro conjunto de datos que se entrena meticulosamente hasta que alcanza una precisión razonablemente buena.

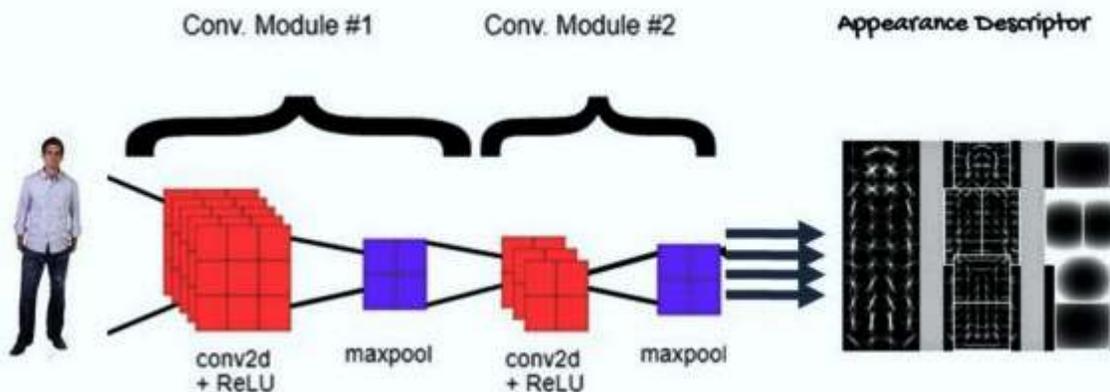
Figura 9: Apariencia Deep SORT Fuente:(Kanje, 2020)



Fuente: elaboración propia

A continuación, tomamos esta red y eliminamos la capa de clasificación final, dejando una capa densa que produce un único vector de características, a la espera de ser clasificado. Este vector de características se conoce como “descriptor de apariencia”.

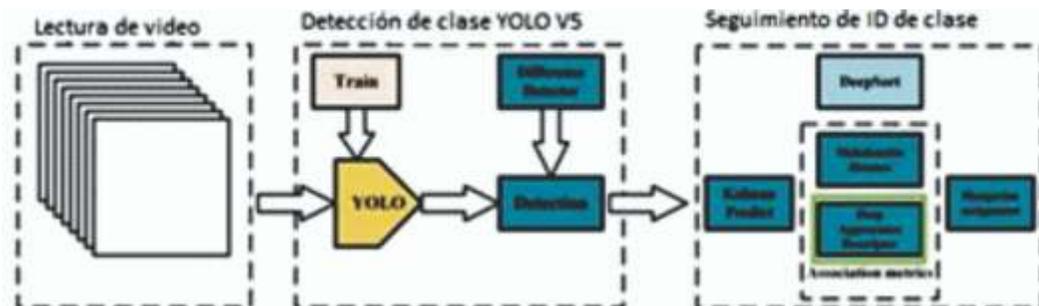
Figura 10: Descriptor de apariencia (Kanje, 2020)



Fuente: elaboración propia

Después de que el descriptor de apariencia obtiene los autores, se utilizan consultas de vecinos más cercanos en la apariencia visual para establecer la asociación medida-a-pista. La asociación medida-pista o MTA es el proceso de determinar la relación entre una medida y una pista existente. Para ello, utilizamos la distancia de Mahalanobis en lugar de la distancia euclidiana para la MTA.

Figura 11: Script



Fuente: https://github.com/julianchaux/Aforo_YOLO-V5_DeepSORT

5. Resultados

Se estudian e identifican los algoritmos a implementar en un ambiente de aprendizaje:

Luego de desarrollar los algoritmos se tienen los siguientes resultados en los ambientes de formación del SENA de Neiva:

Una vez finalizado el análisis de los algoritmos, se puede concluir que es una solución de bajo costo para el centro de formación y se ha desarrollado un modelo de bajo consumo, que detecta en tiempo real la ocupación de un espacio cerrado de los ambientes de aprendizaje y transmite los datos simultáneamente, el modelo cuenta con diferentes programas con los que se obtienen los datos, como la aplicación de reconocimiento y clasificación de imágenes y la aplicación de sensores

6. Referencias bibliográficas

COCO. (2021). Common Objects in Context COCO. Obtenido de <https://cocodataset.org/home>

Gorbalenya, A., Baker, S., Baric, R., de Groot, R., Drosten, C., Gulyaeva, A., . . . Ziebuhr, J. (2020). Severe acute respiratory syndrome-related coronavirus: The species and its viruses – a statement of the Coronavirus Study Group. *Nature Microbiology*, 536-544.

Kanjee, R. (31 de 08 de 2020). DeepSORT — Deep Learning applied to Object Tracking. Obtenidode:<https://medium.com/augmentedstartups/deepsort-deep-learning-applied-toobject-tracking-924f59f99104>

Lin, T.-Y., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., . . . Dollár, P. (2015). Microsoft COCO: Common Objects in Context.

Maiya, S. R. (2019). DeepSORT: Deep Learning to Track Custom Objects in a Video. Obtenido de:<https://nanonets.com/blog/objecttracking-deepsort/deep-sort>

OMS. (11 de 03 de 2020). Alocución de apertura del Director General de la OMS en la rueda de prensa sobre la COVID- 19 celebrada el 11 de marzo de 2020. Obtenido de Organización Mundial de la Salud: <https://www.who.int/es/directorgeneral/speeches/detail/who-director-general-opening-remarks-at-the-media-briefing-on-covid-19—11-march-2020>.

Yolov5. (2021). Github. Obtenido de <https://github.com/ultralytics/yolov5>

