

Trabajo de Detección mediante IA

El trabajo realizado representa el inicio de una serie de proyectos de automatización del sector de reciclaje, con el objetivo de lograr procesos seguros, eficientes y con la posibilidad de contar con datos en tiempo real, de forma remota. Como proyecto futuro se podrá elaborar una maquina a escala industrial con una etapa de post selección con tecnología que separe de manera segura para disminuir el error de clasificación.

Una segunda etapa del proyecto consiste en establecer un sistema de detección mediante cámaras, implementando análisis de imagen mediante IA. Con el objetivo de que el proceso de detección sea menos aparatoso y obtener una mayor velocidad de detección en grandes cantidades, obtener de manera más rápida una mayor cantidad de elementos detectados, y al estar hablando de elementos reciclados muchas veces al obtener el contacto entre los elementos y los sensores, provocaría un mayor desgaste. Para esta etapa se sugiere utilizar redes neuronales Inteligentes prediseñadas, como lo es la red de YOLO, y con los datos obtenidos poder hacer cálculos y gráficos de la Media, Mediana, entre otras variables que ayudan a un mejor estudio de la situación actual.

Aplicando IA con Yolo, se puede separar la sección modular de la maquinaria de clasificación, como lo son las cámaras y el área de procesamiento de imágenes, el cual deberá contar con un diseño particular para poder aislar esta área con respecto a la sección de clasificación hidráulica. Consiste en elaborar una red neuronal inteligente sabiendo que son un modelo computacional evolucionado, compuesto por un conjunto de unidades, llamadas neuronas artificiales, conectadas entre sí para transmitirse señales. La información de entrada atraviesa la red neuronal (donde se somete a diversas operaciones) produciendo unos valores de salida.

Cada neurona está conectada con otras a través de enlaces. En estos enlaces el valor de salida de la neurona anterior es multiplicado por un valor de peso. Estos pesos en los enlaces pueden incrementar o inhibir el estado de activación de las neuronas adyacentes. Del mismo modo, a la salida de la neurona, puede existir una función limitadora o umbral, que modifica el valor resultado o impone un límite que no se debe sobrepasar antes de propagarse a otra neurona. Esta función se conoce como función de activación. Las redes neuronales actuales suelen contener desde unos miles a unos pocos millones de unidades neuronales. El objetivo será la creación de una red de clasificación, el cual en nuestro caso será clasificado entre 4 (cuatro) categorías posibles (plástico, lata, orgánico y vidrio) estas serían nuestra salida de la RNI (red neuronal inteligente) y la misma marcara a cuál de la categoría creé que pertenece.

En la clasificación de imagen se tiene que convertir dicha imagen en blanco y negro, y adjudicarle un valor numérico a cada píxel (el numero 0 seria para totalmente negro y el numero 255 para totalmente blanco). Tomando una imagen de 100 X 100 pixeles, resultan 10.000 pixeles, cada uno con un valor entre 0 y

255, colocando cada píxel en una neurona distinta, formando así la información de entrada.

Por lo tanto, el clasificador basado en la Red Neuronal Inteligente va a estar compuesta por una entrada de 10.000 neuronas y una salida de 4 indicadores posibles, como se observa en la Figura 0.1, que son los materiales del residuo (orgánico, vidrio, plástico o lata). Se agregan 2 capas ocultas de 50 neuronas cada una, luego se puede agregar o quitar capas neuronales y de esta forma, aprueba y error, analizar su precisión.

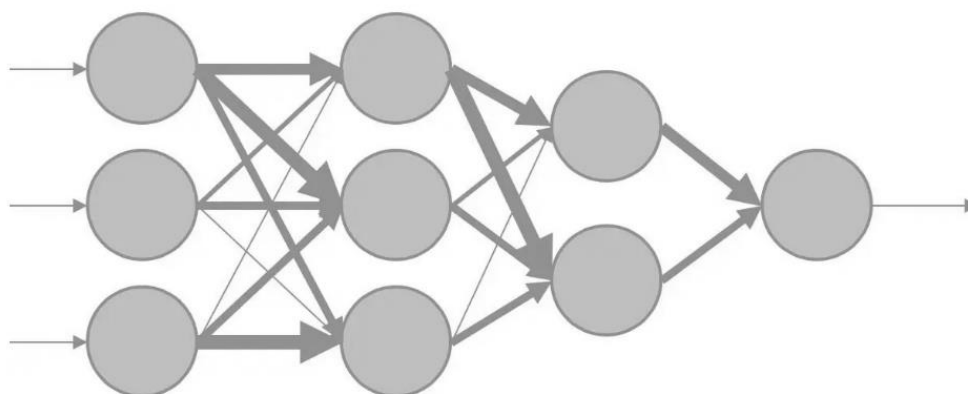


Figura 0.1. Red neuronal Inteligente.

Luego se procede a entrenar a la red, para la correcta detección de los residuos. Este entrenamiento se realiza mediante la obtención de imágenes y etiquetados de la misma, luego creando un sistema de validación se verifica si la Red Neuronal Inteligente se encuentra en condiciones de realizar la clasificación aceptable por el usuario.

También, se puede implementar Una red Neuronal inteligente Pre-Entrenada. Hay dos tipos de modelos de detección de objetos: detectores de objetos de dos etapas y detectores de objetos de una etapa. La arquitectura de los detectores de objetos de una sola etapa (como YOLO) se compone de tres componentes: Backbone, Neck y Head para hacer predicciones densas como se muestra en la figura a continuación. Ver Figura 0.2.

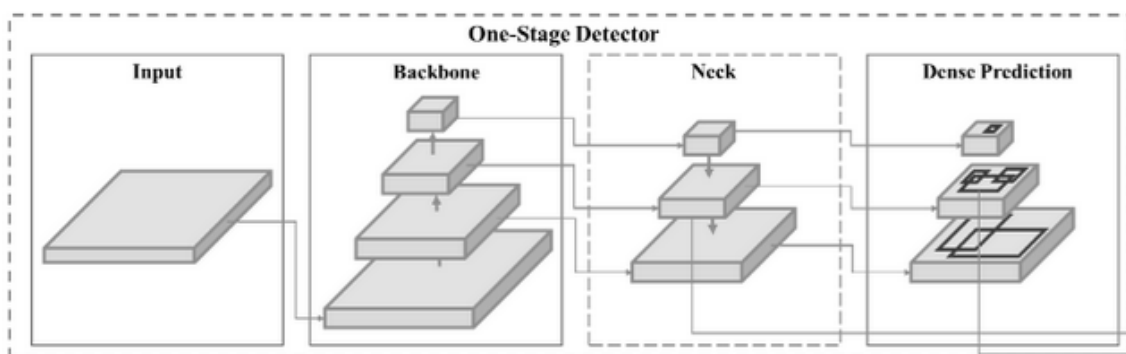


Figura 0.2. Arquitectura del detector de una sola etapa.

La columna vertebral (Backbone de la Figura 0.2) es una red pre-entrenada que se utiliza para extraer una representación rica en funciones para las imágenes. Esto ayuda a reducir la resolución espacial de la imagen y aumenta su resolución de características. El cuello del modelo (Neck de la Figura 0.2) se utiliza para extraer pirámides de características. Esto ayuda al modelo a generalizar bien a objetos de diferentes tamaños y escalas. La cabeza modelo (Dense Prediction la Figura 0.2) se utiliza para realizar las operaciones de la etapa final. Aplica cuadros de anclaje en mapas de características y representa el resultado final: clases, puntajes de objetividad y cuadros delimitadores.

YOLO (You Only Look Once) **¡Error! No se encuentra el origen de la referencia.**, es una red para la detección de objetos, donde la tarea de detección de objetos consiste en determinar la ubicación en la imagen donde están presentes ciertos objetos, y clasificarlos. Los métodos anteriores para esto, como el R-CNN y sus variaciones, usaban un pipeline para realizar esta tarea en múltiples pasos. Esto puede ser lento de ejecutar y también difícil de optimizar, porque cada componente individual debe ser entrenado por separado. YOLO, lo hace todo con una sola red neuronal.

Así que se toma una imagen como entrada, pasa a través de una red neuronal que se ve similar a una CNN normal, y se obtiene un vector de box delimitadores y predicciones de clase en la salida.

La imagen de entrada se divide en una matriz de células $S \times S$. Por cada objeto presente en la imagen, se dice que una celda de la matriz es "responsable" de predecirla. Esa es la celda en la que se encuentra el centro del objeto. Como se puede observar en la Figura 0.3

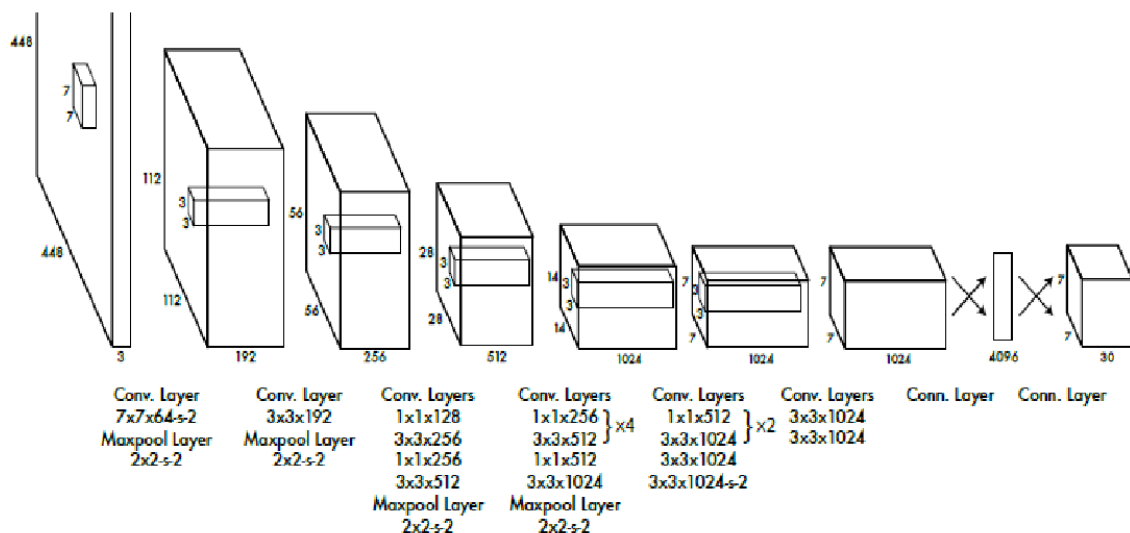


Figura 0.3. Etapas de Análisis.

Cabe mencionar que YOLOv5 fue lanzado con cinco tamaños diferentes en los cuales se puede observar su rendimiento en la Figura 0.4 que se realizó mediante los valores obtenidos por la Tabla 0.1.

n para modelo de tamaño extrapequeño (nano).

s para modelo de tamaño pequeño.
m para modelo de tamaño mediano.
l para modelo de gran tamaño
x para modelo de tamaño extragrande

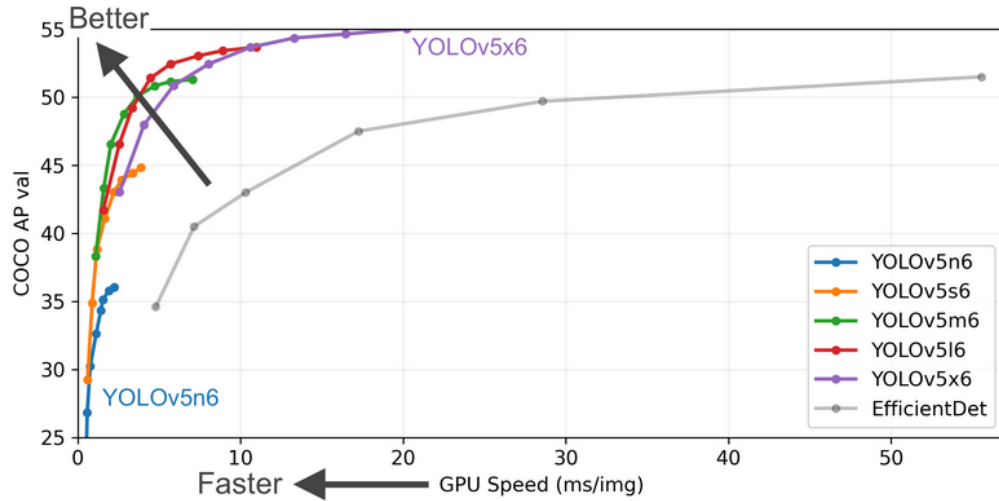


Figura 0.4. Rendimiento de los modelos de diferentes tamaños YOLOv5.

Model	size (pixels)	mAP ^{val} 0.5:0.95	mAP ^{val} 0.5	Speed CPU b1 (ms)	Speed V100 b1 (ms)	Speed V100 b32 (ms)	params (M)	FLOPs @640 (B)
YOLOv5n	640	28.0	45.7	45	6.3	0.6	1.9	4.5
YOLOv5s	640	37.4	56.8	98	6.4	0.9	7.2	16.5
YOLOv5m	640	45.4	64.1	224	8.2	1.7	21.2	49.0
YOLOv5l	640	49.0	67.3	430	10.1	2.7	46.5	109.1
YOLOv5x	640	50.7	68.9	766	12.1	4.8	86.7	205.7
YOLOv5n6	1280	36.0	54.4	153	8.1	2.1	3.2	4.6
YOLOv5s6	1280	44.8	63.7	385	8.2	3.6	12.6	16.8
YOLOv5m6	1280	51.3	69.3	887	11.1	6.8	35.7	50.0
YOLOv5l6	1280	53.7	71.3	1784	15.8	10.5	76.8	111.4
YOLOv5x6	1280	55.0	72.7	3136	26.2	19.4	140.7	209.8
+ TTA	1536	55.8	72.7	-	-	-	-	-

Tabla 0.1. modelos YOLOv5

En este caso se realizaron experimentos con un conjunto de datos de entrenamiento de elaboración propia con más de 100 imágenes por elemento. Mientras mayor sea la cantidad de elementos, mejor se hará la detección y será más exacto el resultado de la predicción deseada. En las imágenes siguientes se muestran algunos resultados obtenidos en esta investigación.

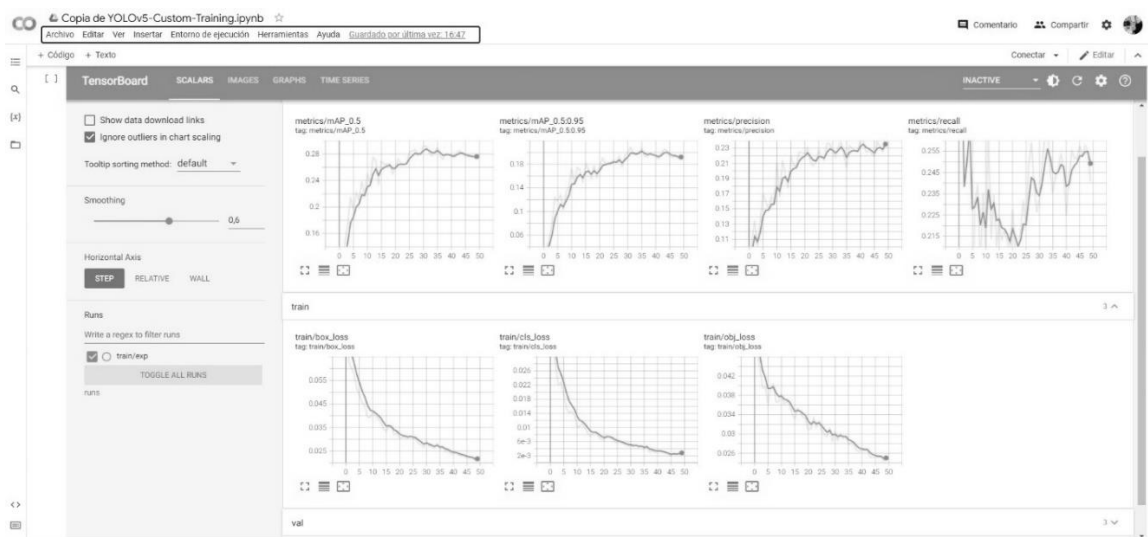


Figura 0.5. Graficas de parámetros de detección de Latas (ROC).

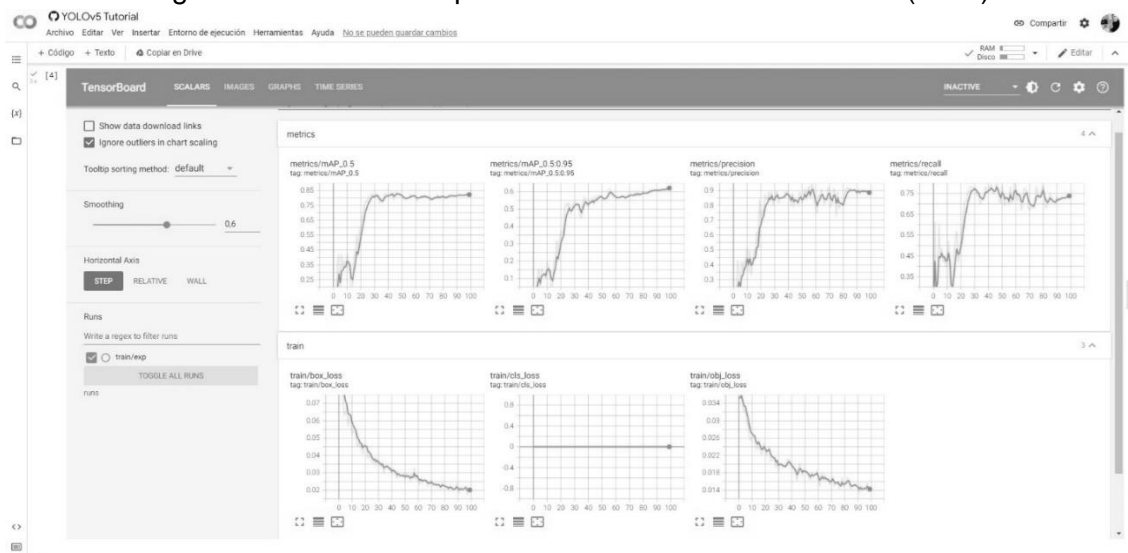


Figura 0.6. Graficas de parámetros de detección de Plástico (ROC)

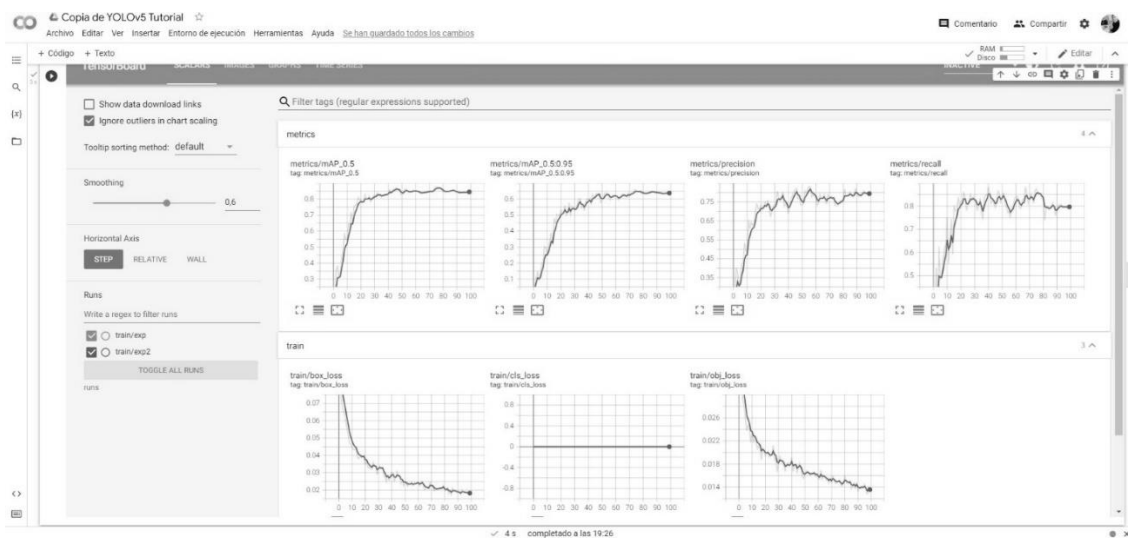


Figura 0.7. Graficas de parámetros de detección de Vidrio (ROC)

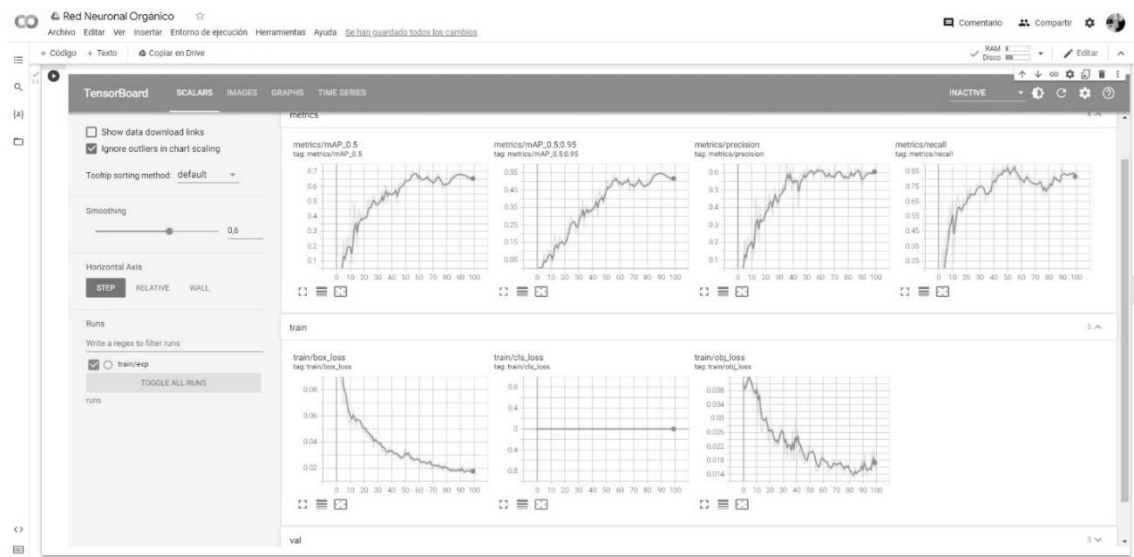


Figura 0.8. Graficas de parámetros de detección de Papel (ROC)

Una tercera etapa contempla la automatización del sector de separación de los residuos, implementando brazos robóticos para la correcta separación. Una cuarta etapa sería el análisis de imagen para un testeo de calidad de los productos reutilizables, de esta forma al separar estos elementos se podrá establecer cuáles serán reubicados y cuáles serán eliminados.

Ventajas

- La Implementación de esta tecnología aumenta la capacidad de clasificación en grandes cantidades de elementos.
- Se puede clasificar cualquier tipo de elementos y con una mayor precisión con respecto a sus características.
- Este sistema cuenta con una mayor vida útil sin mantenimiento, ya que no se encuentra en contacto con ningún elemento.
- Cuenta con un mayor margen de mejoras tanto en su velocidad de detección como su precisión.

Desventajas

- Elevado costo de Mantenimiento.