



Universidad Nacional de Catamarca

Facultad de Tecnologías y Cs. Aplicadas

TRABAJO FINAL
DISEÑO DE SISTEMA DE
DETECCION DE RESIDUOS.

INGENIERIA ELECTRONICA

DEPARTAMENTO DE ELECTRONICA
Maximio Victoria No 55 - Tel/Fax: 0383-4435112 Int. 171
San Fernando del Valle de Catamarca (CP 4700). Argentina.

PRESENTA

Autor: Federico Leandro Diaz Brizuela
MUNº: 00659

Director: Julián Antonio Pucheta

-Año 2022-

-Año 2022-

Resumen

Este trabajo de grado presenta el desarrollo de un sistema de control para el reconocimiento y clasificación de residuos reciclables. Para el desarrollo nos encontramos con la necesidad de proponer un sistema electrónico para la clasificación de residuos sólidos urbanos, teniendo en cuenta que en la actualidad el manejo de los residuos no se procesa de forma correcta.

El estudio tiene como objetivo implementar a través de una red neuronal un algoritmo de inteligencia artificial que reconozca y diferencie diferentes objetos. La inteligencia debe ser capaz de clasificar un objeto, permitir el seguimiento de uno de ellos por medio de una cámara en tiempo real.

En este proyecto creamos diferentes algoritmos para comprender el funcionamiento principal de las redes neuronales y luego implementaremos y analizaremos los diferentes tipos de redes neuronales de YOLO, con el fin de evaluar cuál de ellos es el más fiable y el más adecuado para la arquitectura de detección de residuos. Los mejores resultados serán basados en el mAP (mean average precision) y su velocidad en FPS. Las aplicaciones de este algoritmo son ideales para un sistema autónomo equipado con una GPU y capaz de reconocer y rastrear objetos en tiempo real.

Introducción

1. Motivos

El Machine Learning juega un papel cada vez más importante en la vida cotidiana de hoy en día. Sus aplicaciones son múltiples e incluyen diferentes sectores, desde el militar, la medicina hasta la seguridad o la gestión del tráfico.

El Deep Learning es una subcategoría del Machine Learning. Se refiere a la función y estructura de un cerebro que se conoce como red neuronal artificial. Con la introducción de las CNN (Convolutional Neuronal Network) y el aumento de la capacidad de cálculo de los procesadores, se ha producido un aumento exponencial del uso de las inteligencias artificiales con técnicas de Deep Learning y especialmente en la categoría de la Computer Vision y, por tanto, en el reconocimiento y seguimiento de objetos en imágenes o vídeos.

2. Objetivos

Objetivo General

Diseñar un sistema automático de reconocimiento y clasificación de residuos reciclables (plástico, vidrio, papeles y metal).

Objetivo Especifico

- Evaluar las diferentes tecnologías en la detección y clasificación de imágenes para el reconocimiento de materiales (plástico, vidrio, papel y metal).
- Diseñar el emulador de un sistema de clasificación y detección de residuos.
- Implementar el prototipo del sistema de control Optimo mediante redes Neuronales para la identificación y clasificación de los residuos.
- Visualizar en la pantalla, indicando información sobre el material del residuo.
- Implementar simulación del sistema y seguimiento mediante programa como Colab, Github, Matlab o GNU Octave.
- Implementar comunicación entre la cámara con las Redes Neuronales Inteligentes.

3. Metodología de Trabajo

Para alcanzar los objetivos planteados, se estableció una metodología conformada por cinco etapas. Cada etapa estuvo acompañada por sus respectivas actividades. Estas etapas fueron:

ANÁLISIS Y REUNION DE INFORMACION

En esta etapa se reunirá la información acerca de todos los actores del sistema. Para ello se realizará las siguientes actividades:

- Observación del comportamiento en los puntos ecológicos.
- Definición de la muestra de residuos a trabajar. De acuerdo a la información obtenida en la actividad de observación, se realizará un análisis de los residuos encontrados en los puntos ecológicos y se definirá la muestra de trabajo, teniendo en cuenta forma y tamaño del residuo.
- Elaboración de código del sistema de detección de redes neuronales. Con base a los requisitos del sistema se realizará la selección de librerías y programas necesarios para cumplir con los objetivos planteados.
- Experimentación de Redes Neuronales Inteligentes Artificiales para la detección de los la muestra.

DISEÑO Y ESTRUCTURACIÓN

De acuerdo con los resultados obtenidos en la etapa de análisis e investigación se procederá al diseño físico y lógico del sistema. En esta etapa se realizarán las siguientes actividades:

- Diseño gráfico estructural del sistema.
- Diseño de la ubicación de las cámaras.
- Realización del diagrama de flujo con redes neuronales inteligentes del sistema.

DESARROLLO E IMPLEMENTACIÓN

En esta etapa se realizará el montaje físico del sistema de control de acuerdo con el diseño predefinido en la etapa anterior. Las actividades realizadas en esta etapa serán:

- Construcción estructural de las redes neuronales, en esta parte se construirá la estructura con ayuda de programas como Sublime Text 3, TensorFlow, Colab y GitHub.
- Adecuación del sistema de entrada. En esta actividad se analizará la imagen obtenida de muestra con una serie de filtros y artificios matemáticos.
- Elaboración de una base de datos, en esta etapa se recopilará una cantidad considerable de imágenes con las cuales serán comparadas con la imagen obtenida de muestra.
- Tratamiento de imágenes de base de datos, se realizará un tratamiento de la base de datos, marcando los bordes y girando las imágenes para adaptar la imagen captada en cualquier posición.
- Detección de la ubicación de la muestra con respecto a los otros residuos.

PRUEBAS

Finalmente, se realizarán las pruebas de funcionalidad del sistema de detección mediante Redes Neuronales Inteligentes y las pruebas de efectividad de este. Lo que se hará en esta etapa es:

- Pruebas de componentes del sistema.
- Pruebas con cada elemento a detectar.
- Prueba de eficacia

DOCUMENTACIÓN

Paralelamente a las diferentes etapas de la metodología, se realizarán la documentación pertinente para cada actividad, la cual está contenida en el presente documento.

Marco Teórico

¿Qué es Reciclar?

El reciclaje es un proceso cuyo objetivo es convertir residuos en nuevos productos o en materia prima para su posterior utilización.

Gracias al reciclaje se previene el desuso de materiales potencialmente útiles, se reduce el consumo de nueva materia prima, además de reducir el uso de energía, la contaminación del aire (a través de la incineración) y del agua (a través de los vertederos), así como también disminuir las emisiones de gases de efecto invernadero en comparación con la producción de plásticos.

El reciclaje es un componente clave en la reducción de desechos contemporáneos y es el tercer componente de la regla de las 3R (Reducir, Reutilizar y Reciclar).

Los materiales reciclables son muchos, incluyen todo el papel, cartón, el vidrio, los metales ferrosos y no ferrosos, algunos plásticos, telas y textiles, maderas y componentes electrónicos. En otros casos no es posible llevar a cabo un reciclaje debido a la dificultad técnica o alto coste del proceso, de modo que suele reutilizarse el material o los productos para producir otros materiales y se destinan a otras finalidades, como el aprovechamiento energético.

¿Qué son las Redes Neuronales Inteligentes Artificiales?

Las redes neuronales artificiales están basadas en el funcionamiento de las redes de neuronas biológicas. Las neuronas que todos tenemos en nuestro cerebro están compuestas de dendritas, el soma y el axón: Las dendritas se encargan de captar los impulsos nerviosos que emiten otras neuronas. Estos impulsos, se procesan en el soma y se transmiten a través del

axón que emite un impulso nervioso hacia las neuronas contiguas.

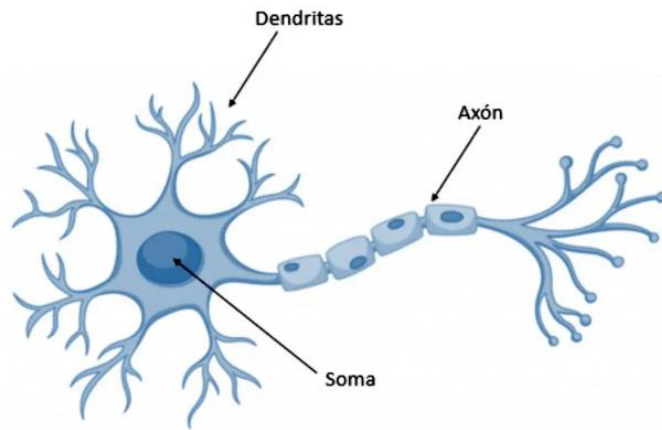


FIGURA N°6-Neurona Física.

A nivel esquemático las redes neuronales artificiales serían:

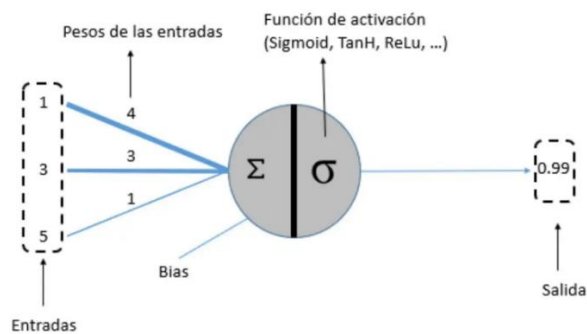


FIGURA N°7- Neurona Artificial

En el caso de las neuronas artificiales, la suma de las entradas multiplicadas por sus pesos asociados determina el “impulso nervioso” que recibe la neurona. Este valor, se procesa en el interior de la célula mediante una función de activación que devuelve un valor que se envía como salida de la neurona.

Del mismo modo que nuestro cerebro está compuesto por neuronas interconectadas entre sí, una red neuronal artificial está formada por neuronas artificiales conectadas entre sí y agrupadas en diferentes niveles que denominamos capas.

Una capa es un conjunto de neuronas cuyas entradas provienen de una capa anterior (o de los datos de entrada en el caso de la primera capa) y cuyas salidas son la entrada de una capa posterior.

En la siguiente imagen podemos ver una red con cuatro capas:

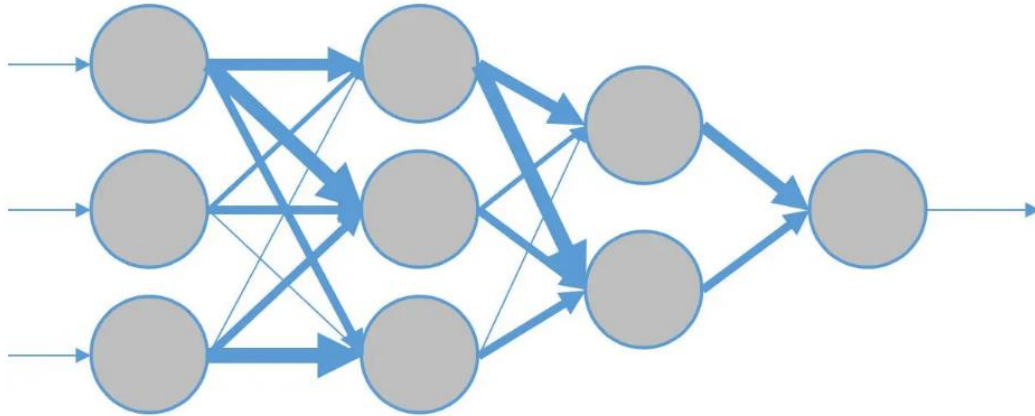


FIGURA N°8-Capas de Neuronas.

Las neuronas de la primera capa reciben como entrada los datos reales que alimentan a la red neuronal. Es por eso por lo que la primera capa se conoce como capa de entrada. La salida de la última capa es el resultado visible de la red, por lo que la última capa se conoce como la capa de salida. Las capas que se sitúan entre la capa de entrada y la capa de salida se conocen como capas ocultas ya que desconocemos tanto los valores de entrada como los de salida.

Las necesidades actuales van mucho más allá de la simple clasificación o localización en imágenes estáticas, hoy en día la necesidad es tener un análisis en tiempo real: ¡nadie querría estar a bordo de un coche autónomo que tarda varios minutos (o incluso sólo segundos) en reconocer las imágenes.

YOLO es mucho más preciso que los otros algoritmos, y aunque es un poco más lento, sigue siendo uno de los algoritmos más rápidos que existen.

Por esta razón en este PI, como primer algoritmo a ser probado se decidió usar YOLO.

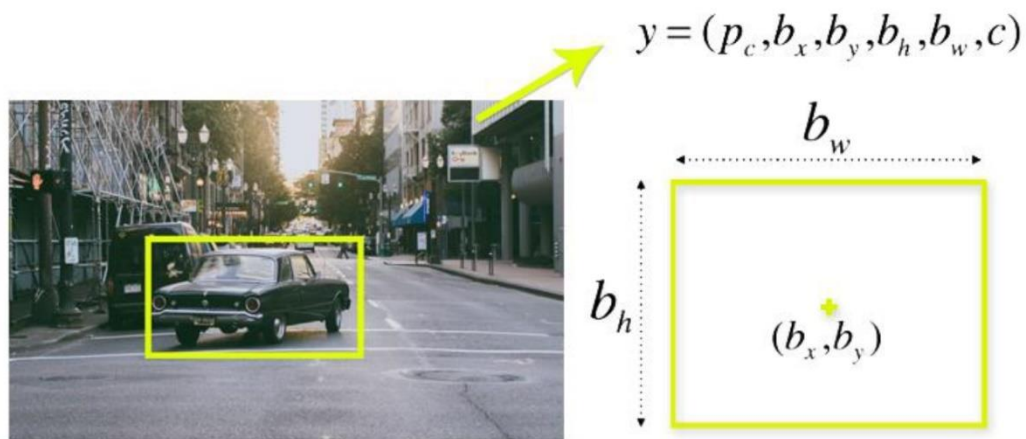
1. Yolo

YOLO (You Only Look Once), es una red para la detección de objetos. La tarea de detección de objetos consiste en determinar la ubicación en la imagen donde están presentes ciertos objetos, así como clasificarlos. Los métodos anteriores para esto, como el R-CNN y sus variaciones, usaban una 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.

Cada celda de la matriz predice los límites de los cuadros B y las probabilidades de clase C . La predicción del cuadro delimitador tiene 5 componentes: $(x, y, w, h, \text{confianza})$. Las coordenadas (x, y) representan el centro del box, en relación con la ubicación de la celda de la cuadrícula. Estas coordenadas se normalizan para que caigan entre 0 y 1. Las dimensiones de la caja (w, h) también se normalizan a $[0, 1]$, en relación con el tamaño de la imagen:



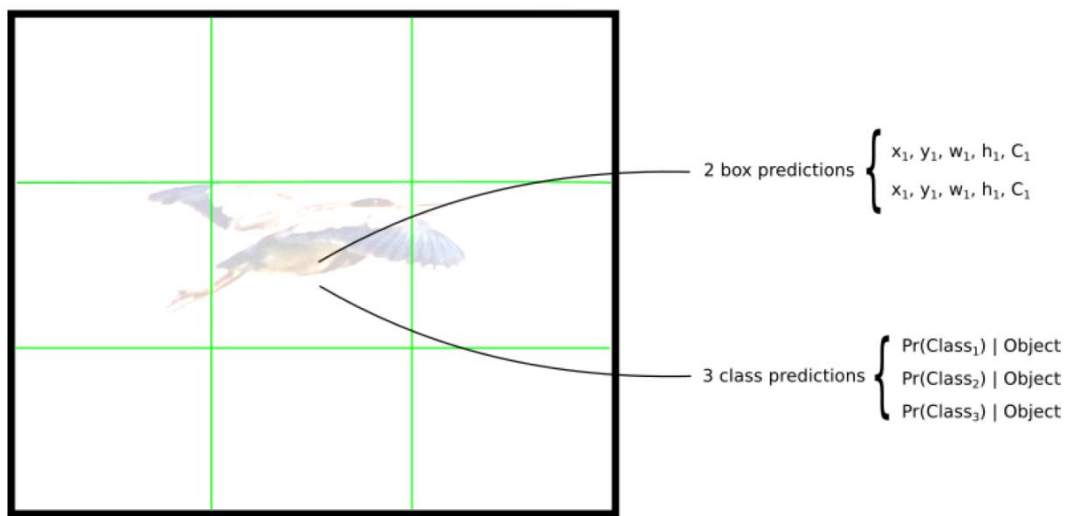
Cada celda de la matriz hace B de esas predicciones, por lo que hay en *total*

$S \times S \times B * 5$ salidas relacionadas con las predicciones de los boxes delimitadores.

También es necesario predecir las probabilidades de la clase, $Pr(Clase(i) | Objeto)$. Esta probabilidad está condicionada en la celda de la matriz que contiene un objeto.

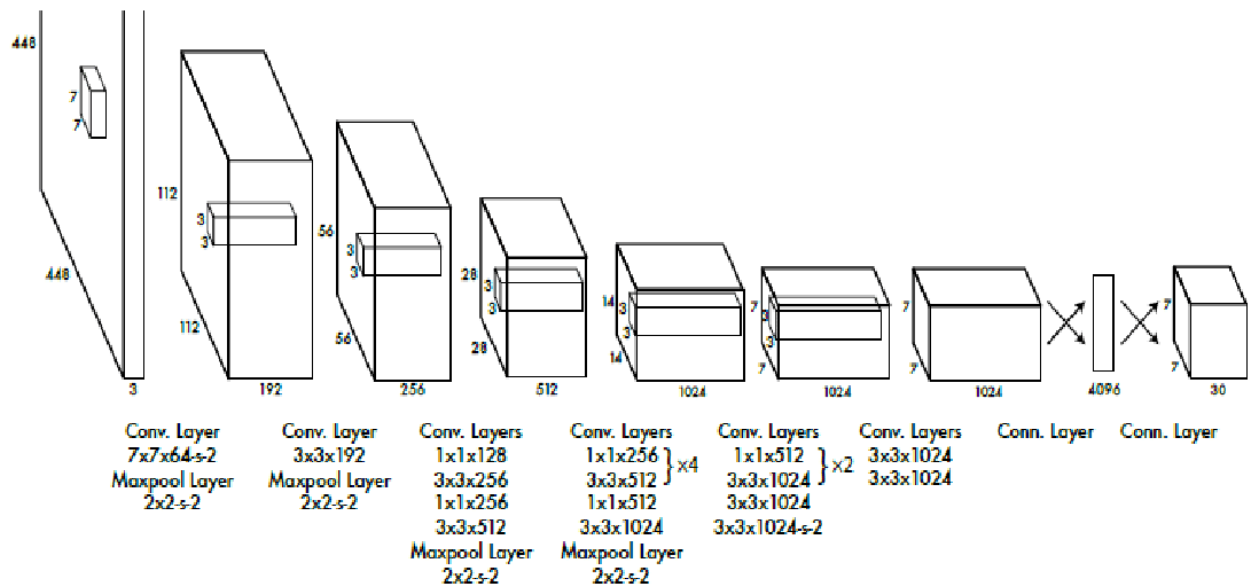
En la práctica, significa que si no hay ningún objeto presente en la celda de la matriz, la función de pérdida no la penalizará por una predicción de clase errónea. La red sólo predice un conjunto de probabilidades de clase por celda, independientemente del número de casillas B. Eso hace que las probabilidades de clase $S \times S \times C$ en total.

Añadiendo las predicciones de clase al vector de salida, obtenemos un tensor $S \times S \times (B * 5 + C)$ como salida.



2. La Red

La estructura de la red se parece a la de una CNN normal, con capas convolucionales y de pooling máximo, seguidas por 2 capas fully conectad al final:



Algunos comentarios sobre la arquitectura:

- Nótese que la arquitectura fue creada para ser usada en el dataset de VOC de Pascal, donde los autores usaron $S=7$, $B=2$ y $C=20$. Esto explica por qué los mapas de características finales son 7×7 , y también explica el tamaño de la salida ($7 \times 7 \times (2 \times 5 + 20)$). El uso de esta red con un tamaño de cuadrícula o un número diferentes de clases podría requerir el ajuste de las dimensiones de las capas.
- Las secuencias de capas de reducción 1×1 y capas convolucionales 3×3 se inspiraron en el modelo GoogLeNet (Inception).
- La capa final utiliza una función de activación lineal. Todas las demás capas usan un RELU con fugas ($\Phi(x) = x$, si $x > 0$; $0.1x$ en caso contrario).