

## Plataforma Detección de objetos en tiempo real

Raúl Moralejo<sup>1,2</sup>, Antonio Storni<sup>2</sup>.

<sup>1</sup> Escuela de Tecnología de la Información y las Comunicaciones  
Departamento de Básicas y Tecnológicas  
UNdeC - Universidad Nacional de Chilecito  
[romoralejo@gmail.com](mailto:romoralejo@gmail.com)

<sup>2</sup> GridTICs – Grupo en Tecnologías de la Información y las Comunicaciones  
Facultad Regional Mendoza  
Universidad Tecnológica Nacional  
[romoralejo@gmail.com](mailto:romoralejo@gmail.com), [antoniostorni@gmail.com](mailto:antoniostorni@gmail.com)

### RESUMEN

En los últimos años, los sistemas de visión artificial han evolucionado sustancialmente. Hoy en día, la visión artificial permite identificar personas, lugares y objetos en imágenes con una precisión equivalente o superior a la capacidad humana y con mayor velocidad y eficiencia. Los métodos que utilizan Deep Learning han obtenido un performance superior a otros enfoques en el campo del reconocimiento de imágenes. El objetivo de esta investigación es implementar herramientas que resuelvan problemas complejos del mundo real aplicando las últimas técnicas de visión artificial, utilizando software libre.

#### **Palabras clave:**

Visión artificial, machine learning, deep learning.

### CONTEXTO

El presente trabajo está inserto en el proyecto de investigación presentado en la Universidad Nacional de Chilecito, Departamento de Básicas y Tecnológicas. Escuela de Tecnología de la Información y las Comunicaciones, y en la Universidad Tecnológica Nacional, Facultad Regional Mendoza, Grupo GridTICs.

### 1. INTRODUCCION

Se puede definir la “Visión Artificial” como un campo de la “Inteligencia Artificial” que, mediante la utilización de técnicas adecuadas, permite la obtención, procesamiento y análisis de cualquier tipo de información espacial obtenida a través de imágenes digitales [1].

Las técnicas de visión artificial están presentes desde hace más de 40 años, pero no fué hasta principios de la década del 2000 que empezaron a aplicarse de modo masivo.

Los primeros algoritmos funcionan utilizando características codificando manualmente los objetos para hacer un reconocimiento, por ejemplo, se codificaban patrones para la nariz y los ojos para la detección de un rostro humano [2] [3].

Con el avance de las tecnologías de Machine Learning, y en especial Deep Learning, las herramientas de visión artificial han avanzado en su velocidad y precisión.

Actualmente, en vez de codificar manualmente las características de los objetos que queremos detectar, lo que necesitamos es

una gran cantidad de datos etiquetados para “entrenar” un clasificador basado en redes neuronales.

El gran avance en los últimos años en la precisión de las técnicas de Deep learning, se debe en parte debido al incremento en la capacidad de cómputo, y por otra parte, a la mayor disponibilidad de grandes sets de datos para entrenar las redes neuronales.

Por otra parte, las primeras aplicaciones de visión artificial, se hacía solamente una clasificación de imágenes, partiendo de una entrada de una imagen, podemos obtener un resultado textual de la clasificación de la imagen, por ejemplo, “perro”, “persona”, “bicicleta”.

Las nuevas técnicas de visión artificial no son sólo “clasificadores”, sino que también son “detectores”. Esto significa que además de la clase del objeto, podemos obtener su ubicación relativa a otros objetos y su tamaño, colocando en un cuadro delimitador [4][5].

Éstas técnicas de “detección” de objetos, fueron mejorando con los años, y actualmente podemos obtener detecciones de objetos sobre imágenes en un tiempo aproximado de 20 milisegundos, utilizando hardware de costo relativamente accesible. Esto implica que se puede hacer detección de video en tiempo real a una velocidad de 50 cuadros por segundo [6].

## 2. LINEAS DE INVESTIGACION y DESARROLLO

Los ejes principales son visión artificial, machine learning, deep learning, software de base libre (lenguaje de programación, framework, base de datos, sistema operativo) para el desarrollo de plataforma.

## 3. RESULTADOS OBTENIDOS/ESPERADOS

El objetivo principal del proyecto es generar una plataforma para la detección de objetos en tiempo real para distintos ámbitos.

Existen ya modelos entrenados y listos para usar, para reconocer objetos comunes, como, “persona, auto, bicicleta, camión”. En el caso de necesitar el reconocimiento de objetos específicos que no se encuentren en los modelos pre-entrenados, sólo se necesitaría contar con un set de datos etiquetados suficientemente grande y diverso, con imágenes de los objetos que queremos que el sistema reconozca para re-entrenar la red.

### La plataforma podrá ser usada de la siguiente manera:

Configurando una o varias cámaras IP como entrada, definir o acotar un área para aplicar la detección de objetos en tiempo real, la clase de objeto que queremos detectar (por ejemplo: “persona”) y a través de un sistema de reglas condicionales, ejecutar distintas acciones que podrían ser: encender alarma, activar enviar un email o un mensaje instantáneo, aumentar o disminuir una variable de control, etc.

### Beneficios de uso

Los beneficios de uso podrían dividirse en 3 categorías:

#### Medio Ambientales

El proyecto ayudará a disminuir las emisiones de CO2 al mejorar la fluidez del tránsito urbano.

#### Sociales

Puede ayudar a las personas con alguna discapacidad o en zonas en donde la inseguridad es un peligro latente.

### Técnicos

El proyecto contribuirá a expandir el uso de técnicas de deep learning en la comunidad.

## Casos de uso

### Industria:

Definir una zona segura o insegura para que transiten las personas.

- Reconocer objetos permitidos y no permitidos en una cinta transportadora.

### Seguridad:

- Definir zonas y horarios en los cuales si se detecta una persona, disparar una alarma.
- Detectar el rostro de personas autorizadas o alertar en caso de rostros no autorizados.

### Ciudades:

- Sincronizar y optimizar los semáforos dependiendo de la cantidad de autos o peatones que haya en un momento dado.
- Generar datos estadísticos de movimiento de personas, vehículos, etc.

### Transporte:

- Generación de un sistema para alertar o frenar un auto o tren en caso de que una persona se encuentre en el camino.

- Framework de Deep Learning desarrollado por Google, Tensorflow [7].
- Red neuronal convolucional pre-entrenada YOLO [8].
- Para comando de la red neuronal y procesamiento de imágenes, el lenguaje de programación Python [9].
- Para plataforma de administración web, el framework Django/Python [10].
- Base de datos PostgreSQL [11].
- Sistema operativo Linux, Ubuntu Server [12].

## Principales pantallas del aplicativo

### Carga de usuarios

Se pueden cargar usuarios y grupos, con permisos por usuario y por grupo.



### Carga de cámaras



## Ejemplos de detección

### Ejemplo de Detección de personas y camión

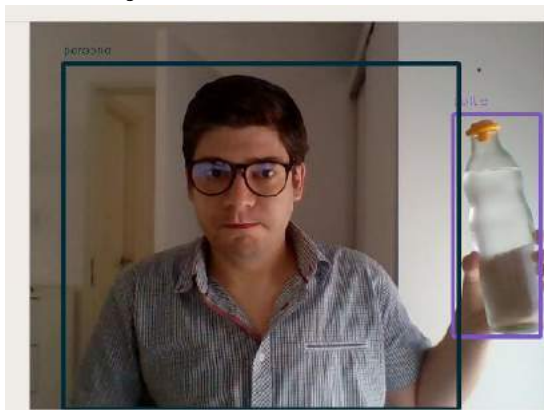
## Tecnología de información utilizada



**Ejemplo de detección de automóvil y persona**



**Ejemplo de detección de persona y objeto (botella)**



#### 4. FORMACION DE RECURSOS HUMANOS

Uno de los principales objetivos del proyecto es la capacitación de los recursos humanos.

La meta como investigadores es la capacidad para realizar investigación científica, generar conocimiento y facilitar la transferencia de tecnología a la sociedad.

Este proyecto de investigación posibilita la colaboración inter-institucional y la ejecución entre grupos de I+D.

Para lograr estos objetivos se dispone del siguiente personal:

- 1 Investigador formado.
- 2 Investigadores de apoyo.

Adicionalmente se realizarán:

Dictado de cursos, seminarios y conferencias.  
Promoción, coordinación y asistencia técnica de tesinas para alumnos de la carrera de grado de ingeniería y licenciatura en sistemas de la UNdeC.

Promoción, coordinación, dirección y asistencia técnica a tesis doctorales, postgrado y/o maestría.

Presentación de trabajos en congresos y reuniones científicas/técnicas.

Publicación de trabajos en revistas con/sin referato.

#### 5. BIBLIOGRAFIA

- [1] VISIÓN ARTIFICIAL | Edgard Martinez  
[http://www.academia.edu/6084688/VISI%C3%93N\\_ARTIFICIAL](http://www.academia.edu/6084688/VISI%C3%93N_ARTIFICIAL)
- [2] Viola Jones Algorithm.  
<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.10.6807>
- [3] HOG Algorithm.  
<http://lear.inrialpes.fr/pubs/2006/DTS06/eccv2006.pdf>
- [4] FAST R-CNN (2015)  
<https://arxiv.org/abs/1504.08083>

[5] YOLO (2015)  
<https://arxiv.org/abs/1506.02640>

[6] YOLO Darknet implementation.  
<https://pjreddie.com/darknet/yolo>

[7] <https://www.tensorflow.org/>

[8] <https://pjreddie.com/darknet/yolo/>

[9] <https://www.python.org/>

[10] <https://www.djangoproject.com/>

[11] <https://www.postgresql.org/>

[12] <https://www.ubuntu.com/>