Universidad Rafael Landívar Campus Quetzaltenango Facultad de Ingeniería Inteligencia Artificial Ing. Dhaby Eugenio Xiloj

Proyecto detección de patrones por medio de redes neuronales.

Girón López, Oscar Juan Pablo 15312-15 Sánchez Castañeda, Norman Andrés 15038-15 Viernes 26 de abril 2019

# Investigación de la fruta

## Descripción de la fruta

La piña tiene forma ovalada, una corona y también una piel muy rugosa de color verdosa que se va volviendo amarilla y marrón a medida que la piña va madurando y desprende un olor dulce cuando ya está lista para comer. Es rica en hidratos de carbono, como éstos son de absorción lenta nos da energía durante más tiempo. La piña contiene minerales como el hierro, el magnesio, el yodo, el zinc o el manganeso, posee vitaminas A, del grupo B y C

## Fases de maduración de la piña post-cosecha

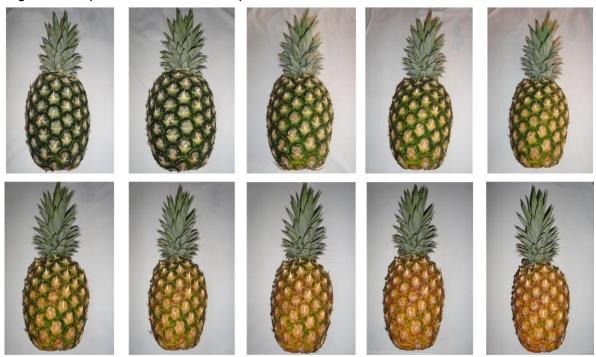
### Punto óptimo de consumo

Figura 1. Etapas de coloración de piña durante maduración.



Estado de madurez	Color	Descripción
Estado 1: (Días: 0)	Verde oscuro	Coloración verde en 75% o más de la piña, con ligera segmentación de color amarillo.
Estado 2: (Días: 4)	Verde-amarillo oscuro	Presenta un 25% a 50% de coloración verde amarilla.
Estado 3: (Días: 8)	Verde-amarillo	Presenta un 50% a 75% de coloración verde amarilla.
Estado 4: (Días: 12)	Amarillo-verde	Presenta un 75% a 90% de coloración amarillo verde.
Estado 5: (Días: 16)	Amarillo-verde	Presenta un 90% a 100% de coloración amarillo verde.

Figura 2. Etapas de maduración de piña extendidas.



## Punto después de consumo

Figura 3. Piña podrida, después de tiempo de consumo.



Después de los 16 días que es el punto de maduración para consumo, el fruto se tornará a un color amarillo-café en un 75% a 100%, y su corona (hojas) se tornarán color verde-amarillo-café.

#### Clasificación de calidad de la piña

La calidad de la piña se verifica mediante:

- Uniformidad de tamaño y forma.
- Firmeza.
- Libre de pudriciones.
- Ausencia de quemaduras de sol.
- Agrietamientos.
- Magulladuras.
- Hojas de la corona: color verde, longitud media y erguidas.

### La fruta se clasifica en 3 categorías:

- Categoría A o Extra: Frutos con peso superior a 1.5 kg, estás no deben tener defectos, salvo defectos superficiales muy leves siempre y cuando no afecten al aspecto general del producto
- Categoría B o 1: Frutos con peso entre 1 kg y 1.5 kg, estás poseen defectos leves de forma, defectos leves de coloración y defectos leves de cascara como rasguños o raspaduras.
- Categoría C o 2: Frutos con peso menor a 1 Kg, estas poseen defectos de forma, defectos de coloración y defectos de cáscara.

## Recopilación de información

Se obtendrán muestras de las fases de maduración de las piñas mostradas anteriormente, para tomar fotos con una cámara de 12 MP pixeles.

Se consultará y se podrán realizar las pruebas correspondientes con Ing. Luis Arturo Sánchez, quien será la persona en la que nos apoyaremos para poder tener datos más exactos en cuanto al nivel de maduración de las piñas.

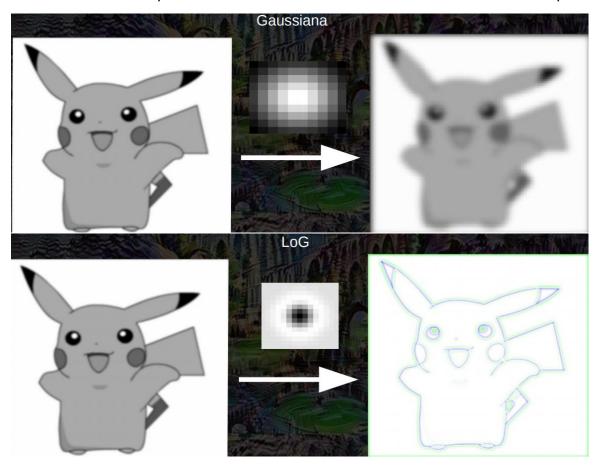
## Documentación del proyecto

#### Marco teórico

#### Redes convolucionales

Son redes utilizadas para completamente para el reconocimiento, detección, clasificación y procesamiento de imágenes y objetos. Son diseñadas para aprender relaciones entrada-salida donde la entrada es una imagen, como su nombre lo indica estas se basan en operaciones de convolución.

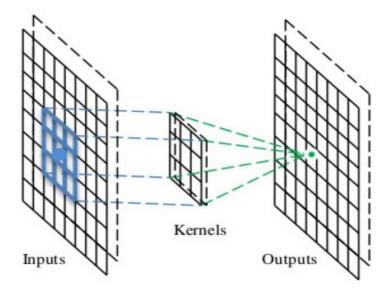
La convolución consiste en aplicar máscaras a las imágenes, esto para poder filtrarlas. Esto hace que visualmente se vean distorsionadas o cambien su aspecto.



Las imágenes que se utilizan como entrada se vectorizan, esto para que sean una sucesión lineal de pixeles, estos pasan por toda la red para que pueda ser entrenada.

En la convolución, cada pixel que se obtiene como salida es una combinación de los pixeles de entrada, las máscaras aplicadas representan la conectividad entre las capas sucesivas utilizadas para el entrenamiento.

Los pesos que se utilizan en el entrenamiento son representados por todos los valores existentes en las máscaras que se les aplican a las imágenes.



Las redes neuronales convolucionales se forman utilizando tres tipos de capas:

- Capas convolucionales
  - Utilizan máscaras
  - Filtran imágenes
- Capas pooling
  - La salida es el máximo de entrada en una ventana
  - Se utilizan para submuestrar
- Capas totalmente conectadas
  - Son aplicadas al final, en estas se pierde la infomación espacial

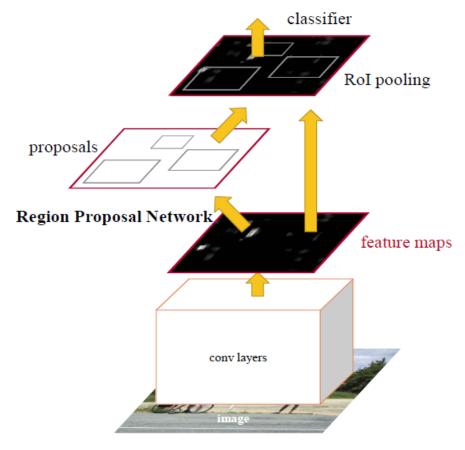
#### **Pooling**

Las imágenes de entrada al pasar las capas que se tenían en la red convolucional, también pasaban por capas de pooling o máx pooling.

El pooling se encarga de resumir características de las imágenes, encuentra el valor máximo en una ventana de muestra y pasa este valor como resumen de las características sobre esa área o región. Como resultado de esto obtenemos una reducción del tamaño de los datos por un factor igual al tamaño de la ventana de muestra sobre la cual se opera.

12	20	30	0			
8	12	2	0	2 × 2 Max-Pool	20	30
34	70	37	4		112	37
112	100	25	12			

El proceso de una red convolucional puede verse así con cada una de las capas utilizadas.

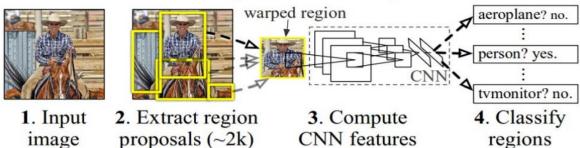


#### COCO

Para la elaboración del proyecto se utilizó un modelo pre-entrenado llamado COCO, el cual es un conjunto de datos de detección, segmentación y subtitulado de objetos a gran escala, dependiendo del entrenamiento. Entre sus principales características se encuentran:

- Segmentación de objetos.
- Reconocimiento de contexto.
- 1.5 millones de instancias de objetos.

# R-CNN: Regions with CNN features



En el caso de nuestro proyecto, al tratarse de un modelo Faster R-CNN(La cual trata de detección de objetos rápidos) que utiliza la red neuronal convolucional ResNet-101 (que permite clasificar las imágenes), la red neuronal de todo el modelo se modifica por medio un archivo config, en cual se establece cuantos objetos tiene que detectar, en donde están los archivos con los nombres de las etiquetas, donde debe de buscar los archivos con los ids, y para el entreno de la red neuronal se implementó lo siguiente:

- ➤ El número de épocas es de 1, mientras el número de pasos es de 500,000 ya que teniendo en cuenta que se puede dejar de entrenar en cualquier momento, y se creará un modelo entrenado hasta ese momento, logrando reanudarse desde donde se ha dejado de entrenar, para obtener un aprendizaje supervisado hemos establecido ese número de pasos.
- ➤ La constante de aprendizaje es 0.0002 para que el aprendizaje del entrenamiento de la red neuronal fuera más exacto y se pudiera obtener un mejor resultado al momento de ingresar una piña.
- Las imágenes de entrada son normalizadas o mapeadas, usando una conversión de los pixeles que esta tiene a 28x28:

$$n \times m = 28 \times 28$$

Por lo que por cada imagen se tienen una cantidad de 784 entradas.

La RNA cuenta con una capa de salida con un total de 7 salidas.

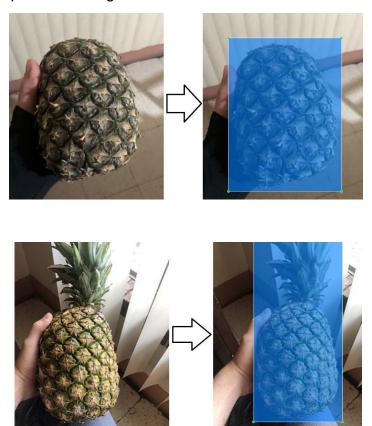
#### Datos de entrenamiento

Los datos de entrenamiento se manejaron por medio de la etiquetación de imágenes, y por medio del programa "labellmg", el cual crea un archivo xml, este contiene las coordenadas en las cuales se encuentra un rectángulo con el que se podrá distinguir el área con los datos que necesita aprender, como se muestran en las siguientes imágenes

Se consideraron 7 etiquetas o estados de una piña, los cuales son:

- Inmadura 16 dias para maduración
- Inmadura 12 dias para maduración
- Inmadura 8 dias para maduración
- Inmadura 4 dias para maduración
- Madura 8 dias en rango para consumo
- Madura 4 dias en rango para consumo
- Podrida

Imágenes de etiquetas de imágenes de entrenamiento:





#### Archivo xml:

Una vez obtenidos los archivos xml, se procede a hacer una conversión en archivos csv, en el cual pasará a ser un único archivo que contiene toda la información que proveen cada uno de los xml de cada imagen.

Cuando se tiene el archivo csv, se hará nuevamente una conversión pero en esta ocasión se convertirá en un archivo .record, este archivo contiene toda la información de las imágenes, las coordenadas y además contendrá la imagen como tal. Una vez teniendo los archivos record, la red estará lista para ser entrenada.

## Investigación

### Descripción del entorno:

Accesible: Se conocen los colores en las piñas, en este caso, cada una de los estados de la piña.

No Determinista: No depende del hecho anterior para dar un resultado.

Episódico: Primero deberá reconocer el archivo .record, luego analizar y encontrar la piña en la imagen, luego marcará la imagen de salida.

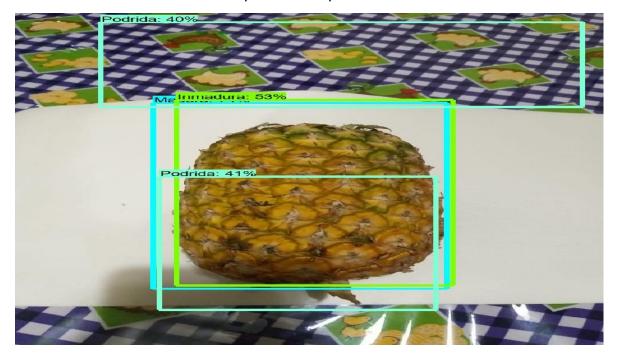
Discreto: Se conoce el número exacto de salidas, en el caso de la piña serán 7, dependiendo del porcentaje de colores que tenga cada una de las piñas.

Estático: No hay otro agente que intervenga.

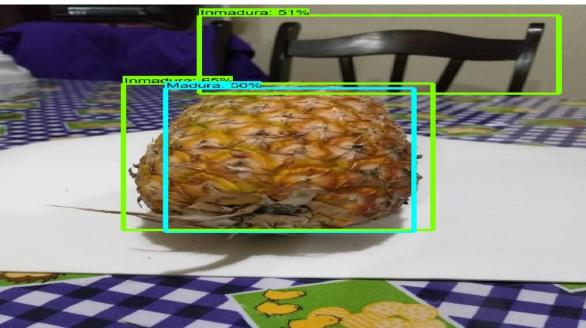
## Primera prueba:

La primera prueba se hizo después de haber cargado todas las imágenes de entrenamiento y de pruebas, se entrenó la red hasta los 364 pasos/steps y allí se creó un checkpoint y se guardó el modelo.

Los resultados obtenidos de la primera prueba no fueron los esperados, la cantidad de fotos y de pasos que se tenían no eran los suficiente para que el algoritmo aprendiera el camino correcto para reconocer las piñas, ya que al pasar imágenes este reconocía muchas formas que no eran piñas:







Por lo que se tuvo que tomar más fotos de las piñas de prueba, y de diferentes ángulos para que la red neuronal se fuera amoldando a lo requerido. Y así evitar posibles errores como el de las imágenenes, donde aparte de mostrar que la piña tiene los 3 grados de madurez, detecta que en el ambiente donde se encuentra la piña, existen más piñas aunque no se encuentren allí.

### Segunda prueba:

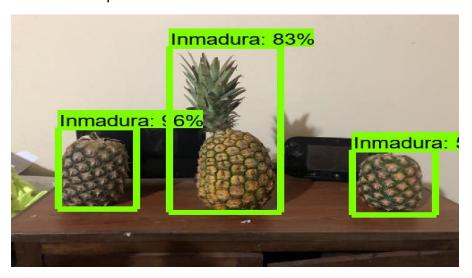
A pesar de que la red ya no encontraba otros objetos que no fueran piñas, ahora el problema se dio que no lograba determinar a qué grado de madurez pertenecía la piña, es decir, encontraba únicamente patrones que coincidieran con más de un grado de madurez



Tal y como se muestra en la imagen, la red neuronal no logra determinar si la piña está podrida o madura, ya que encuentra regiones que coinciden con las características de bien sea de una piña madura o podrida.

## Tercera prueba:

La tercera prueba fue la más exacta, ya que logró entrenar con más imágenes de pruebas y entrenamiento, y además, se ejecutaron una cantidad considerable más de pasos con respecto a las dos pruebas anteriores, por lo que demostró los resultados esperados:



## Cuarta prueba:

En las primeras tres pruebas realizadas se contaba con únicamente 3 etapas de la piña, las cuales eran: "Inmadura", "Madura" y "Podrida", después de ello se añadieron otras 4 etapas y se distribuyeron entre las existentes por lo que se obtuvo un total de 7 etapas, por lo que se volvió a entrenar la red desde el inicio.

La red con pocos pasos aprendió rápidamente a distinguir a la piña dentro del resto del ambiente, pero el distinguir la etapa en la que se encontraba la piña era un problema grande.



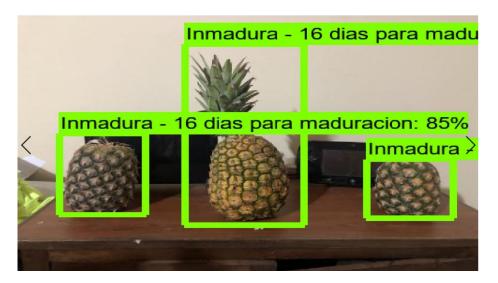


# Quinta prueba:

Como ya se conocía en promedio la cantidad de pasos necesarios para que la red reconociera en gran porcentaje las piñas y las etapas en las que se encontraban, se dejó entrenando la red hasta los 1783 pasos.







Como se puede observar en las imágenes, la red aprendió a reconocer los estados en los que se encontraban las piñas, los resultados son de un 85% efectivos, por supuesto se necesitarían más pasos y muchas más imágenes de cada uno de las etapas de la piña, pero los obtenidos fueron muy buenos, se tomaron fotos de piñas en varios lugares y el modelo utilizado permitió distinguira la piña entre todo ello.

Repositorio de imágenes en drive.

https://1drv.ms/u/s!AgGW0POBosC4jXAi-\_8i4\_ecm4zF?e=4tiNp1

Proyecto en Git Hub.

https://github.com/chuchumix/ProyectoIA