

# Trabajo Final Inteligencia Artificial I – año 2018: Visión Artificial

Fernández, Gonzalo Gabriel

6 de diciembre de 2018

## Resumen

El presente trabajo sobre visión artificial fue realizado como trabajo final de la cátedra Inteligencia Artificial I, de la Facultad de Ingeniería de la Universidad Nacional de Cuyo.

Consiste en el diseño e implementación de un agente capaz de clasificar frutas en base a fotografías. Se realiza un breve estudio teórico del agente. En la implementación se analiza el procesamiento de imágenes en sus diferentes etapas: filtrado, segmentado y extracción de características. Para clasificar las imágenes se utilizan algoritmos de aprendizaje K-means y K-nn, y se hace un estudio de desempeño comparativo entre los dos algoritmos.

Como objetivo final se desea obtener un agente que en un entorno con condiciones controladas se desempeñe con un rendimiento por encima de las exigencias.

## 1. Introducción

### 1.1. Visión Artificial

“La **visión artificial** o visión por ordenador es una disciplina científica que incluye métodos para adquirir, procesar, analizar y comprender las imágenes del mundo real con el fin de producir información numérica o simbólica para que puedan ser tratados por un ordenador.” Enciclopedia Libre [1]. “En la visión por ordenador, intentamos describir el mundo que vemos en una o más imágenes e intentamos reconstruir sus propiedades, tal como la forma, iluminación y distribuciones de color.” Szeliski [3]

### 1.2. Problema a resolver

Se está desarrollando un **sistema de reconocimiento de frutas por visión artificial** con la intención de agilizar el proceso de cobro en las cajas de un supermercado. Para ello se pide desarrollar un agente prototipo que pueda **reconocer bananas, naranjas y limones** a partir de fotografías.

Las condiciones impuestas son que las **imágenes deben tomarse en escala de grises**, y que debe utilizarse los **algoritmos K-means y K-nn** para realizar la clasificación, y a partir de los resultados obtenidos sugerir uno de ellos para ser implementado.

Es recomendado una extracción de características mediante la aplicación de un análisis textural a través de una matriz de covarianza con un vector de características que se construya a partir de los autovalores de dicha matriz.

## 2. Especificación del agente

La descripción del agente y su entorno se realiza en base a la clasificación de Peter Norving [2], Capítulo 1.

### 2.1. Tipo de agente

El agente es **racional**: “En cada posible secuencia de percepciones, un agente racional deberá emprender aquella acción que supuestamente maximice su medida de rendimiento, basándose en las evidencias aportadas por la secuencia de percepciones y en el conocimiento que el agente mantiene almacenado.” Peter Norving [2]. Es **no omnisciente**, ya que no conoce el resultado de sus acciones, no posee total información de todo su entorno. El objetivo de utilizar algoritmos como K-means o K-nn es que el agente **aprenda** a lo largo de sus percepciones, y también que sea **autónomo**, y pueda apoyarse más en la experiencia adquirida gracias a sus propias percepciones que en el conocimiento inicial proveído por el diseñador.

El agente es un **agente que aprende**, debido al tipo de algoritmos que utiliza. La base de datos provee un “modelo” del mundo, y los algoritmos también actúan en base a una utilidad que se analiza más adelante.

## 2.2. Tabla REAS: Rendimiento, Entorno, Actuadores, Sensores

### ■ Tipo de Agente:

Computadora, microcontrolador o microprocesador en el cuál se encuentran los algoritmos entrenados, y los dispositivos de *entrada y salida*, o más correcto en éste análisis, *sensores y actuadores*: cámara fotográfica, monitor, etc.

### ■ Medidas de Rendimiento:

A la hora de implementar los algoritmos, se posee una base de datos con toda la información recopilada etiquetada por clase, en este caso fotografías de bananas, naranjas y limones. A esta base de datos se la divide en dos conjuntos: entrenamiento y testeo. El conjunto de testeo nos permitirá dar una medida de rendimiento del agente, siendo esta medida la cantidad de predicciones correctas de todos los elementos de testeo.

### ■ Entorno:

El entorno del agente será el área donde esté ubicada la fruta a la hora de tomar la fotografía, y la fruta en sí. Al entorno también lo integran todos los factores externos que se exponen más adelante, como luminosidad, fondo, calidad de la fruta, etc.

### ■ Actuadores:

El actuador del agente es sencillamente el medio por el que éste se comunicará con el usuario para darle a conocer su predicción. Puede ser un display, una consola, una señal luminosa o auditiva, o cualquier otro medio que resulte adecuado.

### ■ Sensores:

El sensor del agente es la cámara fotográfica que tomará la imagen de la fruta. Esto es así, si se considera la cámara fotográfica como parte del agente. Sin embargo, la imagen puede provenir de otras fuentes exteriores que no se consideren parte del agente. En este caso, podría considerarse como sensor las herramientas (librerías y sus funciones) que nos permiten procesar la imagen en el formato que se reciba (*PNG, JPEG*), en un formato de información con la que el agente puede operar.

## 2.3. Propiedades del entorno de trabajo

### ■ Parcialmente observable:

El entorno es parcialmente observable, ya que el agente está recibiendo una representación del mundo real, en dos dimensiones; una imagen que describe el entorno mediante un arreglo discreto de píxeles con un determinado valor. Esta limitación está asociada a lo explicado en la introducción, sección 1.1.

### ■ Estocástico:

Que el entorno sea parcialmente observable hace que desde el punto de vista del agente, el entorno parezca en gran medida estocástico. Debido a la gran diferencia entre la representación del entorno en la imagen y el mundo real, el agente nunca puede estar seguro de los hechos en el exterior.

### ■ Episódico:

El entorno es episódico ya que el estado del entorno en un determinado instante de tiempo no depende de la secuencia de estados previos. La cámara fotográfica tomará imágenes de las frutas en el orden que se le presenten, y este orden no tiene ninguna inferencia en el desempeño del agente.

### ■ Estático:

El entorno no cambia mientras el agente está deliberando. Una vez tomada la fotografía, comenzado el procesamiento el agente se aísla en cierto modo del entorno.

### ■ Discreto:

Que el entorno sea discreto está relacionado con que también es estático. El agente no interactúa en tiempo real con el entorno, sino solo en el corto lapso en que toma la fotografía.

### ■ Agente individual:

El entorno es agente individual ya que en todo el proceso de interacción agente-entorno, el único ente racional es el agente.

### 3. Diseño del agente

#### Referencias

- [1] Wikipedia la Enciclopedia Libre. *Visión Artificial*. 2018. URL: [https://es.wikipedia.org/wiki/Visi%C3%B3n\\_artificial](https://es.wikipedia.org/wiki/Visi%C3%B3n_artificial).
- [2] Stuart Russell y Peter Norving. *Inteligencia Artificial: Un Enfoque Moderno*. Pearson Educación S. A., 2004.
- [3] Richard Szeliski. *Computer Vision: Algorithms and Applications*. Springer, 2010.

# Índice

<b>1. Introducción</b>	<b>1</b>
1.1. Visión Artificial . . . . .	1
1.2. Problema a resolver . . . . .	1
<b>2. Especificación del agente</b>	<b>1</b>
2.1. Tipo de agente . . . . .	1
2.2. Tabla REAS: Rendimiento, Entorno, Actuadores, Sensores . . . . .	2
2.3. Propiedades del entorno de trabajo . . . . .	2
<b>3. Diseño del agente</b>	<b>3</b>