



UNCUYO
UNIVERSIDAD
NACIONAL DE CUYO



Trabajo final Inteligencia Artificial 1 (2024)

Visión artificial y reconocimiento de voz

Universidad Nacional de Cuyo, Facultad de Ingeniería

Autor: Francisco Castel

2024/2025

Introducción

Visión Artificial

- Identificación de objetos mediante imágenes.
- Aplicaciones: automatización, robótica, diagnóstico médico.

Reconocimiento de Voz

- Conversión de audio en texto o comandos.
- Uso de coeficientes cepstrales y modelos de clasificación.

Desafíos abordados

- Reducir dimensionalidad.
- Clasificación robusta en condiciones controladas.

Resumen

Objetivo

- Combinar visión artificial y reconocimiento de voz para clasificar verduras, en concreto.
 - Berenjena
 - Camote
 - Papa
 - Zanahoria

Técnicas utilizadas

- KMeans para imágenes
- KNN para audios

Enfoque

- Extracción de características para cada tipo de dato
- Reducción de dimensionalidad para optimización y coherencia

Clasificador de Imágenes (No Supervisado)

- Algoritmo: **KMeans**.
- Características:
 - Momentos de Hu.
 - Color promedio.
- Entorno controlado:
 - Fondo constante.
 - Iluminación uniforme.

Clasificador de Audio (Supervisado)

- Algoritmo: **KNN**.
- Características:
 - MFCCs (media, máximo, mínimo, desviación estándar).
 - Espectro de Potencia
 - Entropía energética
 - RMS y duración.
- Entorno controlado:
 - Bajo nivel de ruido.

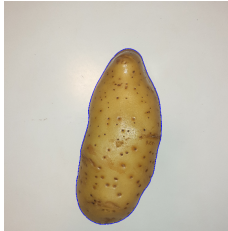
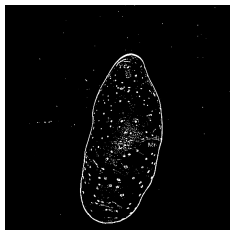
Preprocesamiento de imágenes

- Pasos

- 1 Desenfoque gaussiano
- 2 Escala de grises
- 3 Umbralización adaptativa
- 4 Operación morfológica

- Resultados

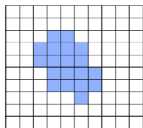
- Mejora en la detección de contornos
- Reducción de ruido sal y pimienta
- Características finales: Momentos de Hu y color promedio



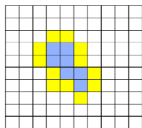
Filtro morfológico

Consiste en dos operaciones:

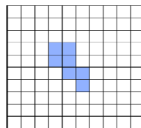
Erosión



(a)

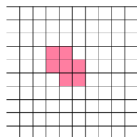


(b)

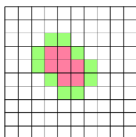


(c)

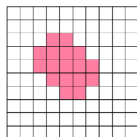
Dilatación



(a)



(b)



(c)

A partir de ellas se pueden lograr dos operaciones compuestas muy interesantes.


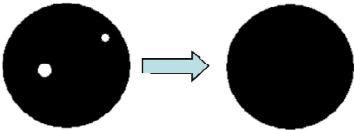
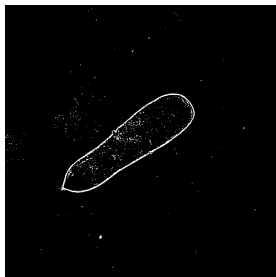
Apertura $I \circ H = (I \ominus H) \oplus H$	Cierre $I \bullet H = (I \oplus H) \ominus H$
	

Figura 3.- Ejemplos de operaciones de apertura (izquierda) y cierre (derecha).

Binarización adaptativa

La binarización adaptativa es una solución práctica al inconveniente de la selección de un umbral fijo para la totalidad de la imagen.

En lugar de tener un único umbral como la media de los dos valores predominantes, la media se calcula de forma local mediante una ventana definida.



Bibliography

Here is a reference to a source. For example, *Momentum Contrast for Unsupervised Visual Representation Learning by He et al. (2020)* [?] is a placeholder text for a bibliography entry.