

## **Visión Artificial**

Sergio Andrés Mejía Tovar

Aldemar Yamid Ramírez Ávila

## **Proyecto Final**

### **Clasificación de Basuras**

### **Entrega final**

Presentado a:

Andrea del Pilar Rueda Olarte



Pontificia Universidad Javeriana

Facultad de Ingeniería

Departamento de Ingeniería de Sistemas

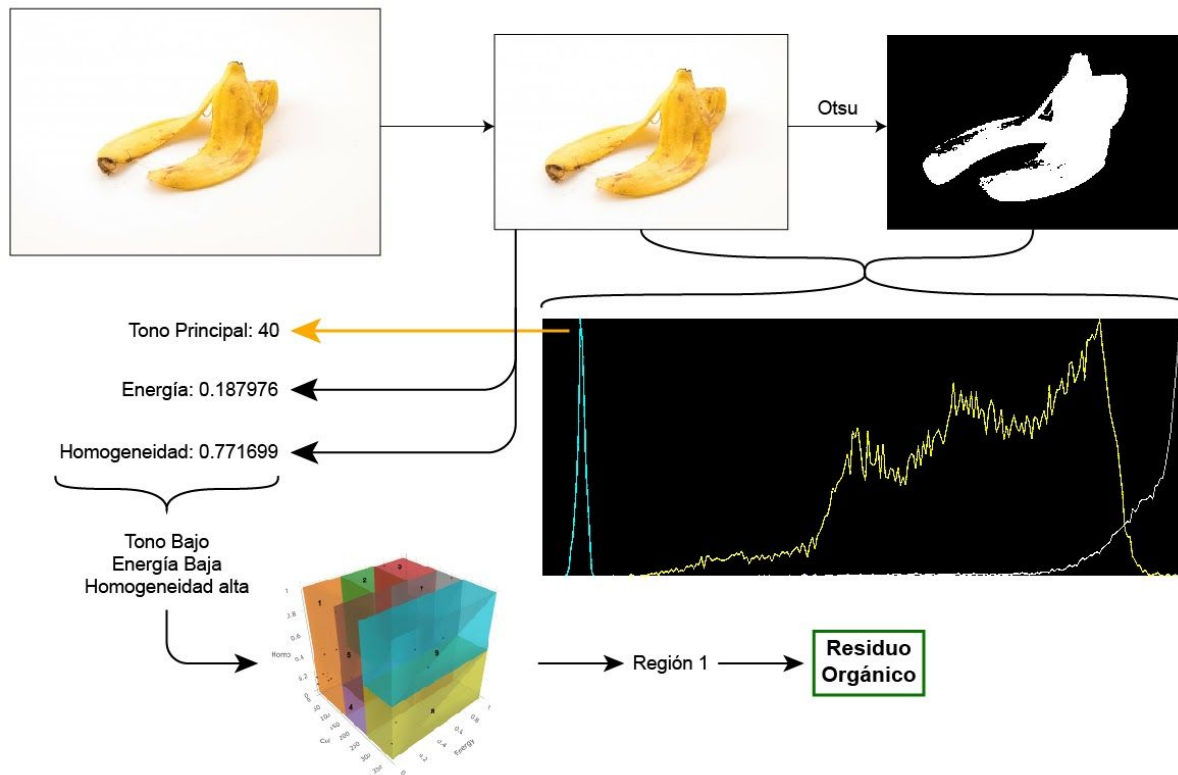
Mayo de 2020

<b>Introducción</b>	<b>2</b>
<b>Recolección de Imágenes</b>	<b>2</b>
<b>Segmentación</b>	<b>3</b>
<b>Extracción de características (histograma y textura)</b>	<b>3</b>
Cálculos de histograma	3
Cálculos de textura	4
<b>Caracterización</b>	<b>5</b>
Definición del vector de características	5
Definición de las regiones	7
Clasificación de las regiones	8
<b>Resultados Finales</b>	<b>9</b>
<b>Conclusiones</b>	<b>10</b>

## 1. Introducción

El presente documento muestra el proyecto del curso de Visión Artificial, el cual pretende realizar una caracterización de objetos que representan diversos residuos, con el fin de determinar si estos son biodegradables, reciclables, etc. por medio de únicamente la foto del objeto. Para esto se implementaron técnicas de segmentación, generación de histogramas y caracterización. En las etapas tardías del proyecto se refinó el alcance a caracterizar un residuo únicamente como orgánico o inorgánico

Cada uno de los pasos se describe a lo largo del documento, argumentadas con las decisiones tomadas. A continuación se muestra una imagen del proceso general:



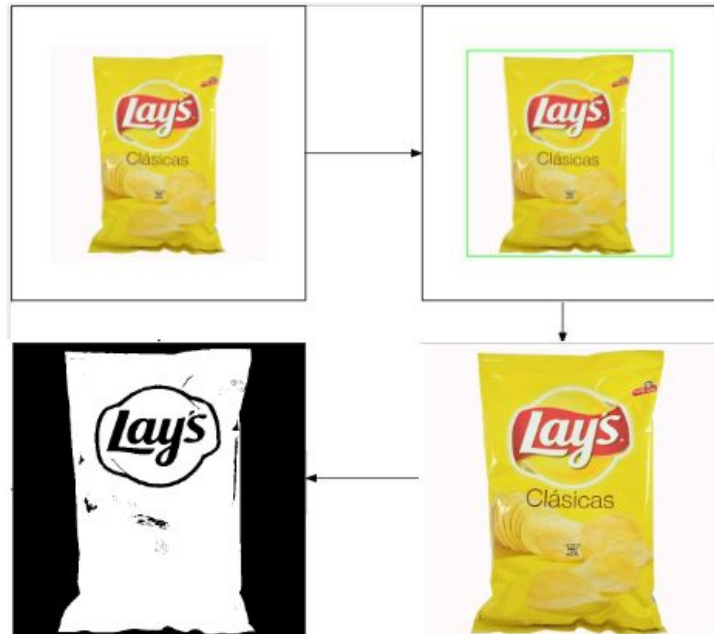
## 2. Recolección de Imágenes

Para las imágenes se usaron aquellas que cumplieran con un contraste alto entre el objeto y el fondo, donde el objeto debería ser más oscuro que el fondo. Para simplicidad de la caracterización se usaron imágenes con un fondo blanco.

Se buscaron imágenes que tuvieran productos de consumo cotidiano además que sean desechables como el caso de los paquetes de comida, biodegradables como el caso de las frutas y reciclables como el caso de la bolsa de basura y cartón.

### 3. Segmentación

El primer paso del proceso es una segmentación inicial del objeto del fondo, utilizando un proceso de binarización con el método de Otsu. Para esto, se recortó inicialmente la imagen al 70% de su alto y ancho y luego se generó la imagen binarizada, guardándose en memoria.



*Proceso de Segmentación de las Papas Lay's*

El objetivo de realizar esta segmentación es poder trabajar con los valores únicamente dentro de la sección segmentada, ignorando el blanco del fondo de cada imagen.

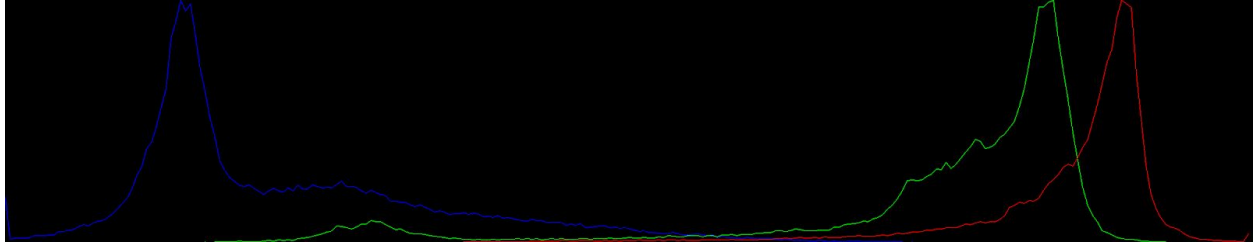
### 4. Extracción de características (histograma y textura)

Después de haber segmentado la imagen, se le hicieron dos tratamientos independientes. El primero fue generar un histograma en diferentes espacios de color, y el segundo fue calcular la matriz de coocurrencia para el cálculo de texturas.

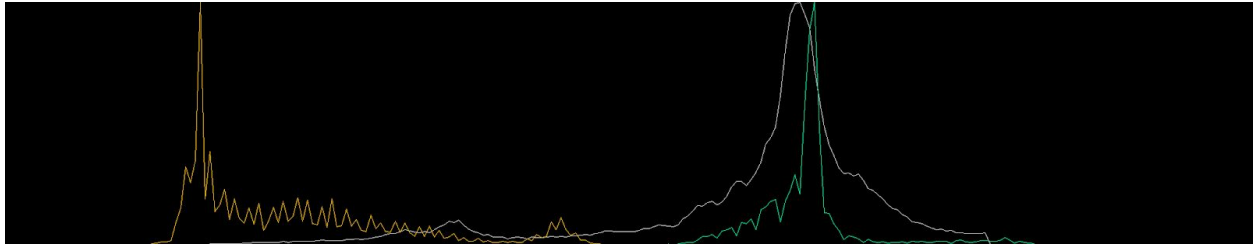
#### 4.1. Cálculos de histograma

Para la generación de los histogramas de cada imagen se trabajaron 3 espacios de color: RGB, HSV y YCrCb. Esto con el fin de analizar cada uno y observar cuál es el que ofrece mayor información en cuanto a usabilidad para caracterizar más adelante las imágenes. Cabe recordar que estos histogramas fueron realizados utilizando únicamente la región segmentada.

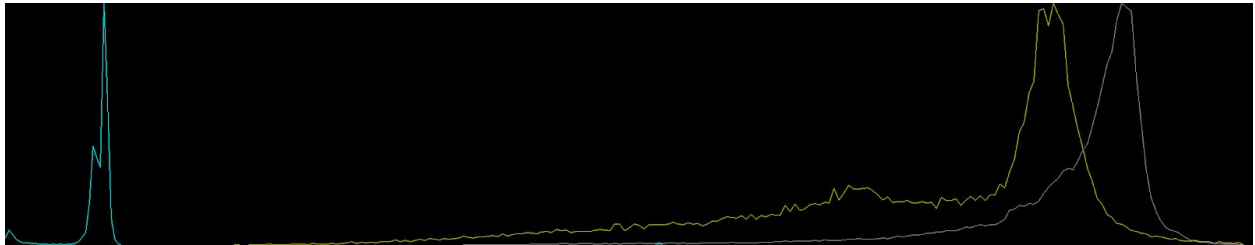
A continuación se muestran los histogramas de la imagen de prueba del paquete de papas Lay's amarillo.



*Histograma RGB de las Papas Lays*



*Histograma YCrCb de las Papas Lays. El canal Y se ve en amarillo, Cr en gris y Cb en cyan.*



*Histograma HSV de las Papas Lays. El canal H se ve en cyan, el S en amarillo y V en gris.*

Basados en estos resultados, se decidió trabajar principalmente con el espacio de color HSV, y trabajar específicamente con el canal de Tono (Hue) para posteriores análisis. Específicamente se halló el valor del Tono más repetido en la imagen (el pico más alto en el histograma) y se guardó este valor.

## 4.2. Cálculos de textura

Para los cálculos de la textura se tomó el texel como la imagen completa de la región segmentada con un ángulo 0 y una distancia de 1 pixel, con el fin de analizar los resultados de los diferentes tipos de texturas se calculó los casos de Energía, Media, Contraste, Homogeneidad y Entropía.

Así, se definió la matriz de coocurrencia del siguiente modo:  $P_{0^{\circ},1}(a, b)$

A continuación se muestra los resultados para una parte del grupo de imágenes elegidas.

	<b>Energía</b>	<b>Media</b>	<b>Contraste</b>	<b>Homogeneidad</b>	<b>Entropía</b>
<b>prueba0.jpg</b>	0.0542168	215.952	6364.4	0.205788	2.42094
<b>prueba1.jpg</b>	0.187976	243.539	1105.69	0.771699	1.55115
<b>prueba2.jpg</b>	0.442491	179.14	5379.29	0.338983	3.62245
<b>prueba3.jpg</b>	0.430178	231.618	4175.94	0.691545	1.55291
<b>prueba4.jpg</b>	0.796196	686.563	10720.8	0.903647	4.08411
<b>prueba5.jpg</b>	0.145916	205.626	8028.6	0.445926	2.45287
<b>prueba6.jpg</b>	0.555983	242.862	2582.56	0.787207	1.13444
<b>prueba7.jpg</b>	0.589851	243.676	1439.6	0.790028	1.0622
<b>prueba8.jpg</b>	0.330917	226.427	2035.97	0.651798	1.92717
<b>prueba9.jpg</b>	0.746989	162.727	8026.44	0.145143	4.04182
<b>prueba10.jpg</b>	0.295405	202.353	7285.38	0.566115	2.12932
<b>prueba11.jpg</b>	0.881773	166.434	4076.39	0.34869	3.24175
<b>prueba12.jpg</b>	0.340186	188.318	9407.15	0.270756	3.45991
<b>prueba13.jpg</b>	0.190738	196.572	13589.3	0.248252	3.45352
<b>prueba14.jpg</b>	0.374456	219.032	4147.69	0.639052	1.73668
<b>prueba15.jpg</b>	0.743176	162.05	4746.24	0.367668	3.2369
<b>prueba16.jpg</b>	0.455329	239.277	1840.94	0.830218	1.16074

## 5. Caracterización

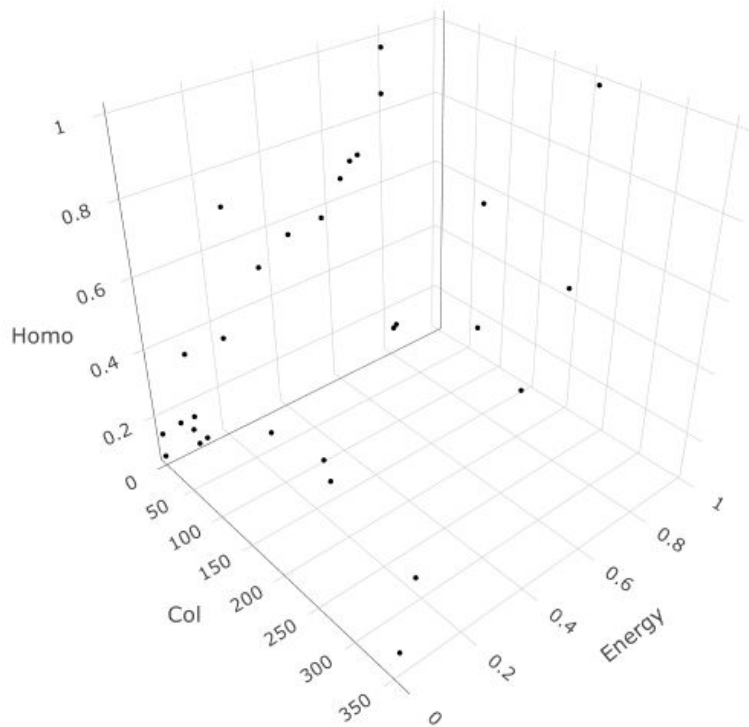
La siguiente gran etapa del proceso fue la caracterización de imágenes, que es a lo que se quiere llegar al final del proyecto. El objetivo de esta etapa es definir para cada imagen unas características en valores numéricos, y basadas en estas poder etiquetar la imagen como residuo orgánico o residuo no orgánico.

### 5.1. Definición del vector de características

Después de la extracción de características previamente realizada, se decidió en tomar 3 características para la caracterización del objeto: El Tono más repetido, el valor de Homogeneidad de la textura y el valor de la Energía de la textura. Se decidió en utilizar estas tres características por las siguientes razones:

- Se decidió usar el Tono puesto que al hallar el tono más repetido para cada imagen se observó 3 grupos claramente definidos en los rangos  $[0,150)$ ,  $[150,250)$  y  $[250,360]$ .
- Para poder definir las otras dos características, se hallaron todas las medidas de textura (Entropía, Energía, Homogeneidad, Contraste, Media) y se calculó el coeficiente de variación para cada una de estas. Se tomaron las dos características con mayor coeficiente de variación, resultando en el uso de la Energía y la Homogeneidad.

Así, se pudo graficar un punto en tres dimensiones para cada una de las imágenes probadas (tamaño de la muestra = 45). Los puntos se muestran en el siguiente gráfico:



*Gráfico de los vectores de características para cada imagen del set de pruebas*

## 5.2. Definición de las regiones

Al observar los puntos previamente graficados, se tuvieron dos observaciones importantes:

1. Los puntos se agrupaban principalmente en la sección de color bajo [0,150), por lo que esta región se puede dividir con mayor precisión en los otros dos ejes.
2. Hay secciones del gráfico que no contienen puntos, por lo que se pueden descartar estas regiones del plano cartesiano para la identificación de regiones.

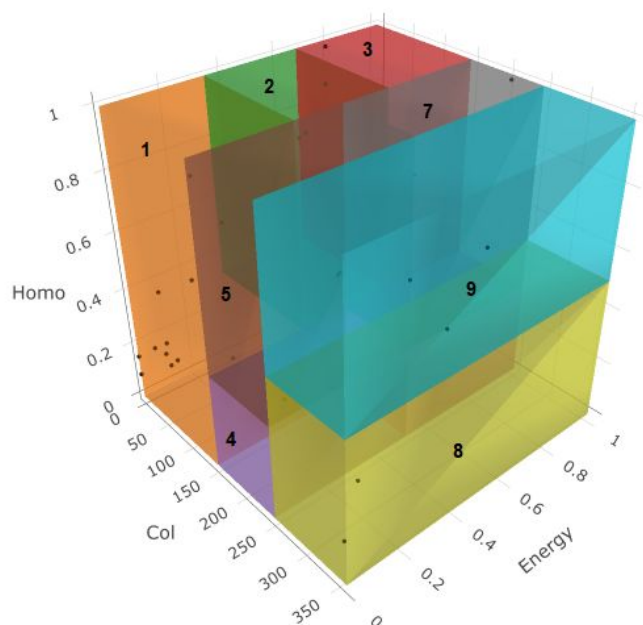
Con esto en mente, se decidió generar 9 regiones en el plano, descritas del siguiente modo:

Región	Tono	Energía	Homogeneidad	Intervalo
1	Bajo	Baja	Arbitraria	$C = [0, 150), E = [\min E, \min E + (\max E - \min E)/3], H = [\min H, \max H]$
2	Bajo	Media	Media-Alta	$C = [0, 150), E = [\min E + (\max E - \min E)/3, \min E + 2 * (\max E - \min E)/3], H = [\min H + (\max H - \min H)/3, \max H]$
3	Bajo	Alta	Media-Alta	$C = [0, 150), E = [\min E + 2 * (\max E - \min E)/3, \max E], H = [\min H + (\max H - \min H)/3, \max H]$
4	Medio	Media-Baja	Baja	$C = [150, 250), E = [\min E, \min E + (\max E - \min E)/2], H = [\min H, \min H + (\max H - \min H)/3]$
5	Medio	Media-Baja	Media-Alta	$C = [150, 250), E = [\min E, \min E + (\max E - \min E)/2], H = [\min H + (\max H - \min H)/3, \max H]$
6	Medio	Media-Alta	Baja	$C = [150, 250), E = [\min E + (\max E - \min E)/2, \max E], H = [\min H, \min H + (\max H - \min H)/3]$
7	Medio	Media-Alta	Media-Alta	$C = [150, 250), E = [\min E + (\max E - \min E)/2, \max E], H = [\min H + (\max H - \min H)/3, \max H]$
8	Alto	Arbitraria	Baja	$C = [250, 360], E = [\min E, \max E], H = [\min H, \min H + (\max H - \min H)/2]$
9	Alto	Arbitraria	Alta	$C = [250, 360], E = [\min E, \max E], H = [\min H + (\max H - \min H)/2, \max H]$

Tabla 1. Definición de las regiones.



De manera gráfica, estas regiones se pueden observar en el siguiente gráfico:



*Gráfico de las regiones definidas*

### 5.3. Clasificación de las regiones

Puesto que se generaron más regiones que etiquetas se necesitan, se procedió a clasificar estas regiones en una de las dos características deseadas: Orgánicos e Inorgánicos. Esto se hizo de modo completamente manual, observando la región que el algoritmo le asignaba a cada imagen del set de prueba. Después de observar los resultados se llegó a la siguiente clasificación:

- Orgánicos: Regiones 1 y 2
- Inorgánicos: Regiones 3, 4, 5, 7 y 8

Lo primero que es necesario decir es que varias de estas regiones no tienen resultados perfectos, principalmente la región 1; esto se debe a que tanto imágenes orgánicas como inorgánicas fueron asignadas a esta región. Sin embargo, puesto que la mayor cantidad de imágenes eran orgánicas, se decidió clasificar como región orgánica.

Cabe aclarar además un par de aspectos en relación a esta división. La primera es que la región 6 no aparece en ninguna de la clasificaciones. Esto es debido a que ninguna imagen fue asignada a esta región. Por otra parte, la región 9, que tampoco aparece en la clasificación, se dejó como indeterminada, puesto que habían objetos de ambas características en ella, a diferencia de la región 1 que la mayoría de las imágenes catalogadas en esta región eran orgánicas.

## 6. Resultados Finales

Basado en lo anterior, se pudo clasificar cada imagen en uno de los dos tipos de residuo que nos interesó. A continuación mostramos las imágenes del set de prueba en su región correspondiente.



*Imágenes Orgánicas*



*Imágenes Inorgánicas*



*Imágenes Indeterminadas*

Se logra observar que si bien esta caracterización no es perfecta, se logra obtener ciertos grupos definidos. Así, el porcentaje de aciertos fue del 66.67% en relación a un etiquetado manual.

## **7. Conclusiones**

Si bien los resultados fueron variables en cuanto al etiquetado de las imágenes, se observó que en general hubo una buena clasificación de los objetos, dándonos cuenta que uno de los factores que más influenció fue la diferencia de color entre los objetos. Sin embargo, se notó que los valores de Homogeneidad y Entropía aportaron una caracterización más fina, teniendo objetos que si fuera solo por color se encontrarían en el otro grupo.

Además, durante el proceso se tuvo la idea de clasificar por la forma general del objeto, siguiendo la idea de objetos más rectangulares tienden a ser inorgánicos. Sin embargo, no se logró implementar esta funcionalidad, puesto que la irregularidad de las formas no ayudaba a la correcta caracterización de polígonos.

Para terminar, se logró evidenciar que es posible extraer múltiples características de la imagen que permitan a una caracterización de objetos sin necesidad de entradas adicionales, y que este trabajo queda abierto al refinamiento, adición de herramientas para mejorar la clasificación, como redes neuronales o más características de forma, curvatura, etc.