

Universidad Nacional de Rosario  
Facultad de Ciencias Exactas, Ingeniería y Agrimensura  
Tecnicatura Universitaria en Inteligencia Artificial

PROCESAMIENTO DE IMÁGENES 1



Informe:  
“Trabajo Práctico 3”

**Fecha:** 14/12/2025

**Autores:**

- CORRADINI, JULIO - Legajo: C-7525/6
- LEÓN, ESTEBAN - Legajo: L-3413/4
- SAAD, JUAN MANUEL - Legajo: S-5803/3
- ZIMMER, FEDERICO - Legajo: Z-1239/4

**Docentes:**

- SAD, GONZALO
- CALLE, JUAN MANUEL
- ALLIONE, JOAQUIN

2025

# **Problema 1 - Cinco dados**

## **1. Problemática a resolver**

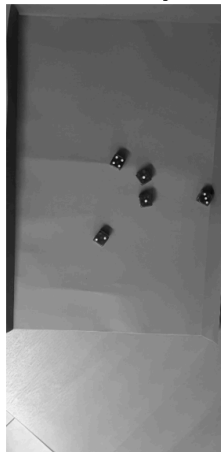
Las secuencias de videos con nombre tirada\_.mp4, corresponden a un conjunto de tiradas de 5 dados. A partir de ellos se deben abordar las siguientes consignas:

- Desarrollar un algoritmo que detecte automáticamente los frames en los cuales los dados se encuentran detenidos (en la Figura 1 se muestra uno de los frames en el cual los dados se encuentran estáticos luego de una tirada), y que muestre por terminal cada número obtenido.
- Generar videos (uno para cada archivo) donde los dados, mientras estén en reposo, aparezcan con su bounding box asociado, un nombre identificador y el valor obtenido.

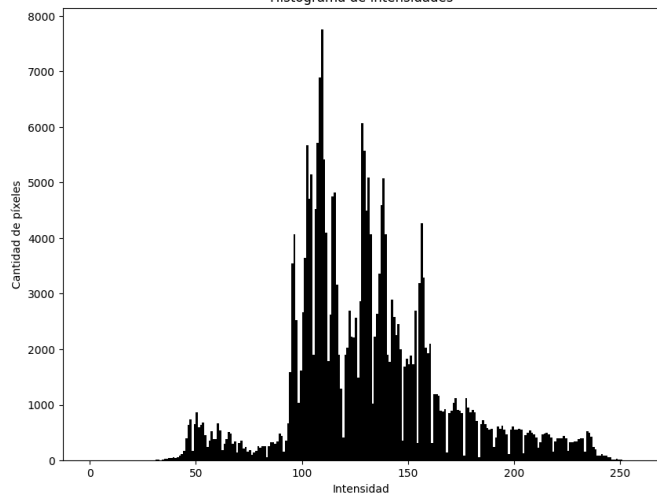
## **2. Análisis Preliminar y Desafíos**

Para comenzar la resolución del problema, se seleccionó manualmente un frame donde los dados se encontraban estáticos para analizar las propiedades de la imagen. Al convertir la imagen original a escala de grises, se observó que el fondo no presentaba una intensidad uniforme, mostrando una dispersión de brillo considerable que dificultaba la binarización.

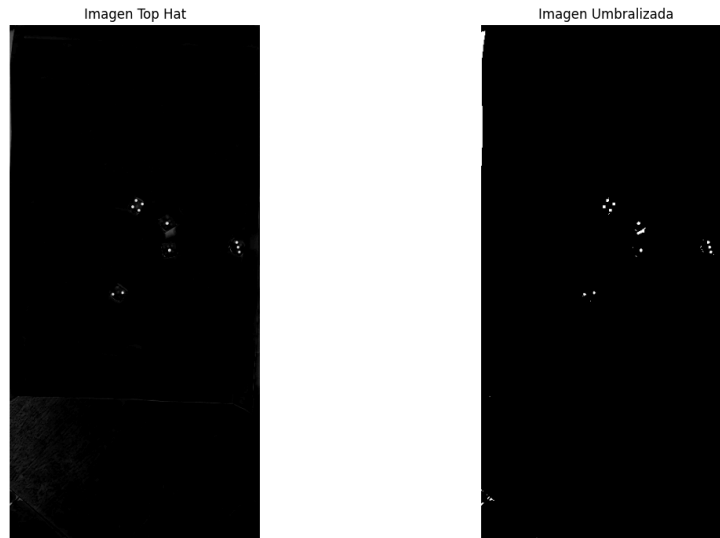
Frame en escala de grises



Histograma de intensidades



Para intentar corregir este problema de iluminación, se aplicó inicialmente TopHat. Sin embargo, esta técnica no arrojó los resultados esperados: al ser los bordes de los dados de una intensidad similar a la del fondo procesado, estos tendían a desaparecer o fusionarse con el entorno, perdiendo la estructura del objeto.

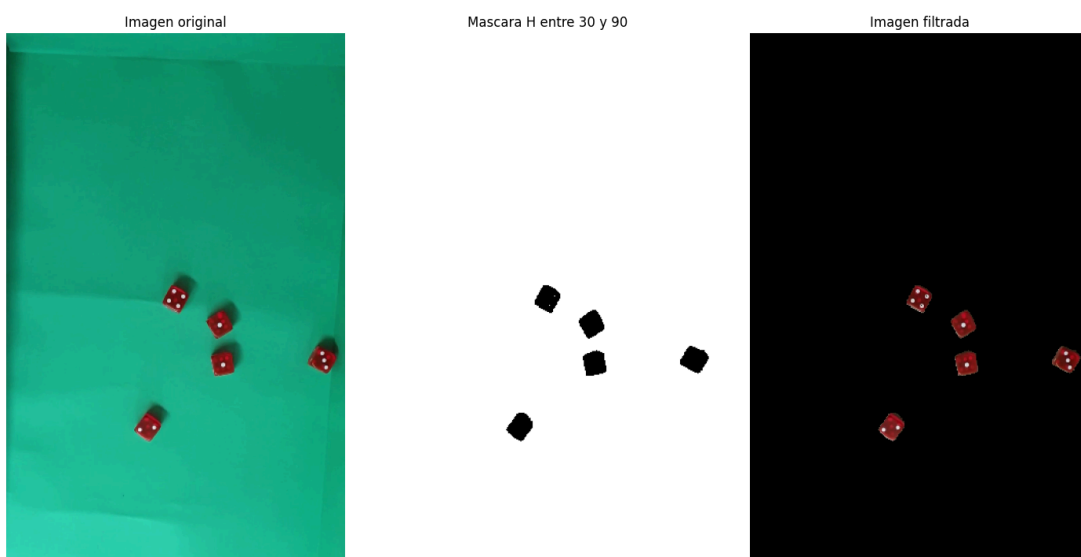


### 3. Metodología de Procesamiento

Dado el fallo de los métodos morfológicos iniciales, se optó por una estrategia basada en la segmentación por color y análisis de componentes conectados.

**3.1. Definición de ROI y Segmentación por Color:** Se detectó que existían zonas periféricas en el video que no aportaban información relevante. Por lo tanto, se implementó un recorte porcentual de la imagen para trabajar exclusivamente sobre la Región de Interés (ROI).

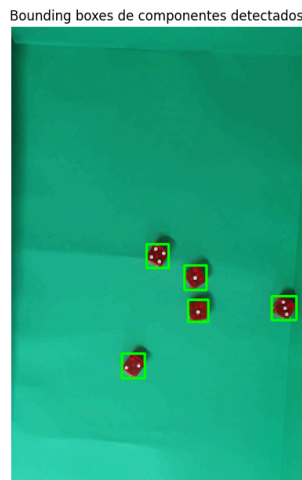
A continuación, se analizó el espacio de color. Al convertir la imagen a HSV, se identificó que el fondo poseía una tonalidad (Hue) verde característica. Esto permitió aplicar un filtro de paso de banda en el canal H (Hue entre 30 y 90) para generar una máscara. Al invertir esta máscara, logramos eliminar el fondo verde, aislando los objetos de interés.



**3.2. Detección de los Dados:** Con el fondo eliminado, se procedió a la detección de los dados:

- Convertimos la imagen filtrada a escala de grises.
- Binarizamos mediante Otsu.
- Aplicamos `connectedComponentsWithStats`.

Se implementó un filtro por área para descartar ruido pequeño. Solo se consideraron válidos aquellos frames donde se detectaron exactamente 5 componentes que cumplieran con el rango de área de un dado.

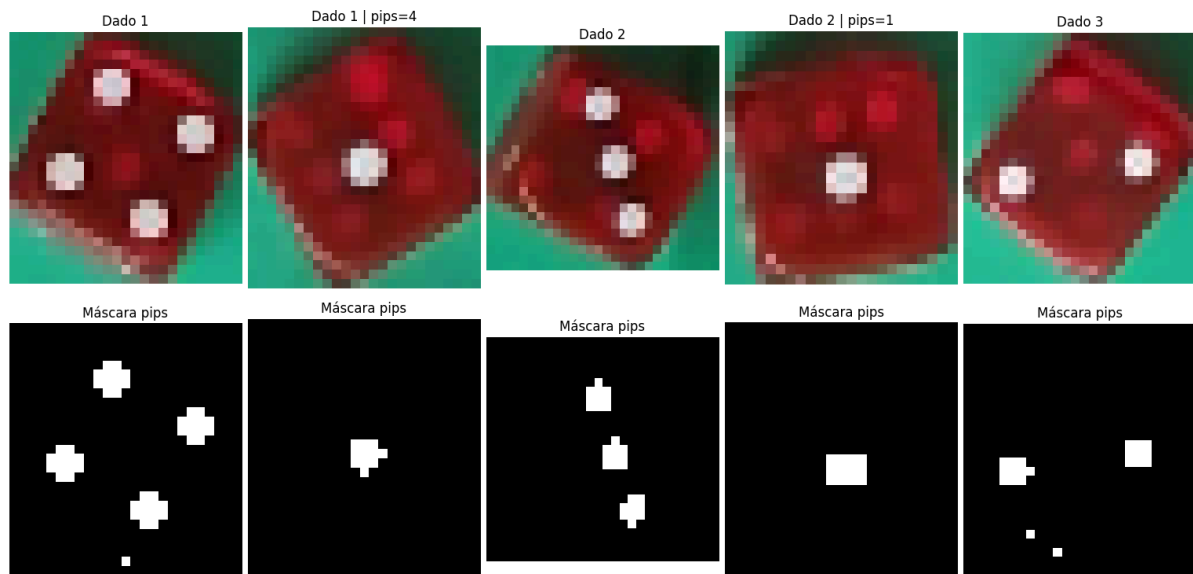


**3.3. Detección de Pips (Valores):** Una vez localizados los dados, se realizó un crop de cada uno utilizando las coordenadas de su bounding box. Para contar los pips, se aprovechó de las regiones más blancas y brillantes del objeto.

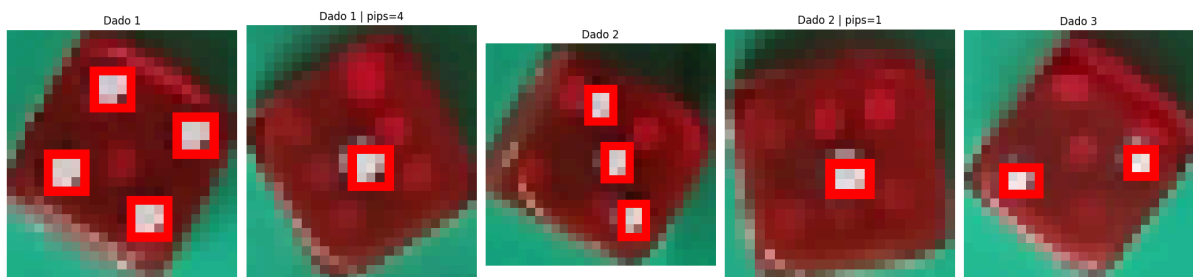
Se aplicó un filtrado en el espacio HSV buscando píxeles con baja saturación y alto valor (brillo):

- **Rango HSV:** Sat [0-60], Val [180-255].

Esto generó una máscara binaria donde los pips aparecen como píxeles blancos y el resto del dado como negro.



Gracias a este filtrado y a uno posterior por área, se eliminó el ruido interno del dado, permitiendo utilizar nuevamente componentes conectados para contar cuántos pips existían dentro de cada crop.



## 4. Análisis Temporal y Estabilidad

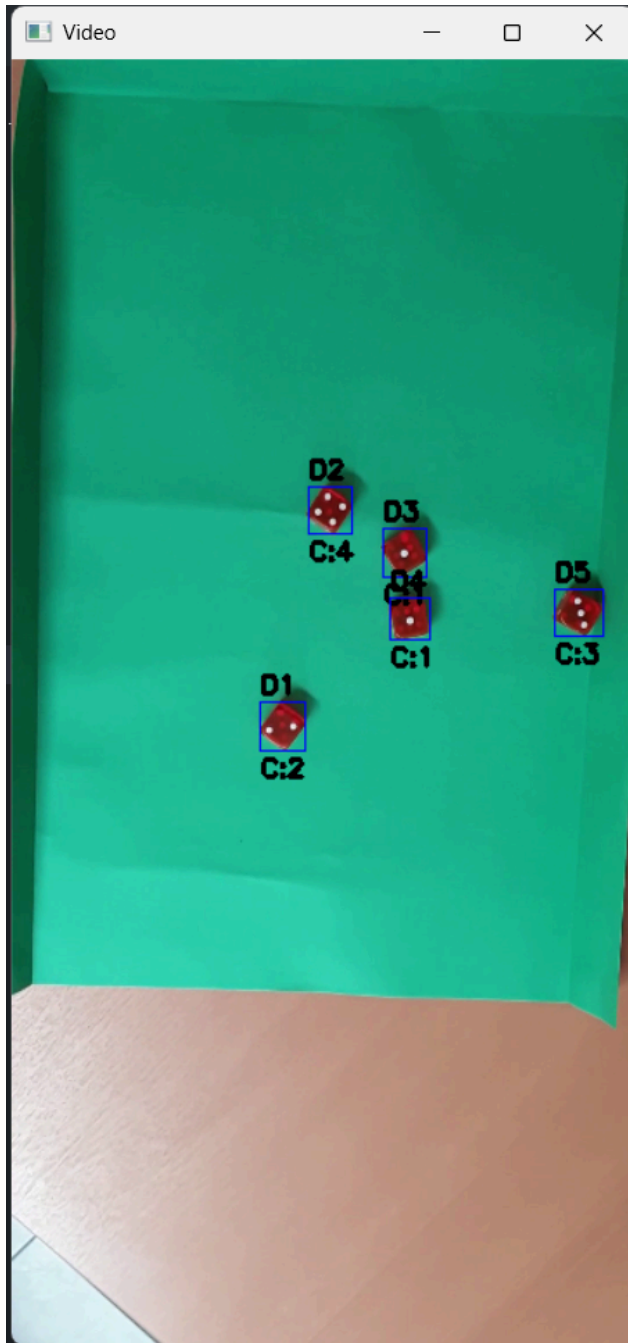
Uno de los problemas detectados en algunos videos fue el desenfoque de movimiento. Para solucionar esto, se implementó un algoritmo de verificación de estabilidad:

- Se analizan ventanas de 3 frames consecutivos.
- Se verifica que en los 3 frames existan 5 dados.
- Se comprueba que la cantidad de pips sea constante en el tiempo.
- **Tolerancia de Área:** Se definió un margen de error (10 píxeles) para el área de los dados. Si la variación de área entre frames consecutivos supera este umbral, se asume que el dado sigue moviéndose o está desenfocado.

El sistema identifica el primer frame estable y el último frame estable de la secuencia. La visualización de resultados se mantiene consistente entre estos dos puntos para facilitar la interpretación.

## 5. Resultados y Visualización Final

Finalmente, el sistema genera un video de salida y una tabla de resultados impresa por terminal. Sobre la imagen original, se dibujan los bounding boxes de cada dado, indicando su ID y el valor detectado.



El script exporta un video donde se visualizan estas detecciones únicamente cuando los dados se encuentran estáticos, garantizando una lectura limpia y precisa de la tirada.

## 6. Conclusiones

El desarrollo de este trabajo práctico permitió implementar un algoritmo de procesamiento de imágenes eficaz para la interpretación de tiradas de dados, logrando cumplir con los objetivos de detección, conteo y visualización de resultados.

A lo largo del proceso, se identificó que las técnicas tradicionales de corrección de iluminación (como Top-Hat) resultaban insuficientes debido a la complejidad del fondo y el bajo contraste de los bordes. No obstante, la adopción de una estrategia basada en la segmentación por color en el espacio HSV demostró ser una solución robusta para aislar correctamente los objetos de interés.

Asimismo, el desafío introducido por el desenfoque de movimiento en los videos fue resuelto exitosamente mediante el algoritmo de estabilidad temporal. Esta validación entre frames consecutivos garantizó que el procesamiento se realizara únicamente bajo condiciones óptimas, asegurando la precisión de los datos obtenidos y la correcta generación del video final.