

TRABAJO PRÁCTICO INTEGRADOR N°3

IA4.4 Procesamiento de Imágenes

Tecnicatura Universitaria en Inteligencia Artificial

Trabajo realizado por:
Lautaro Florenza, Katia Otterstedt y Sebastián Palacio

Problema 1 - Cinco dados

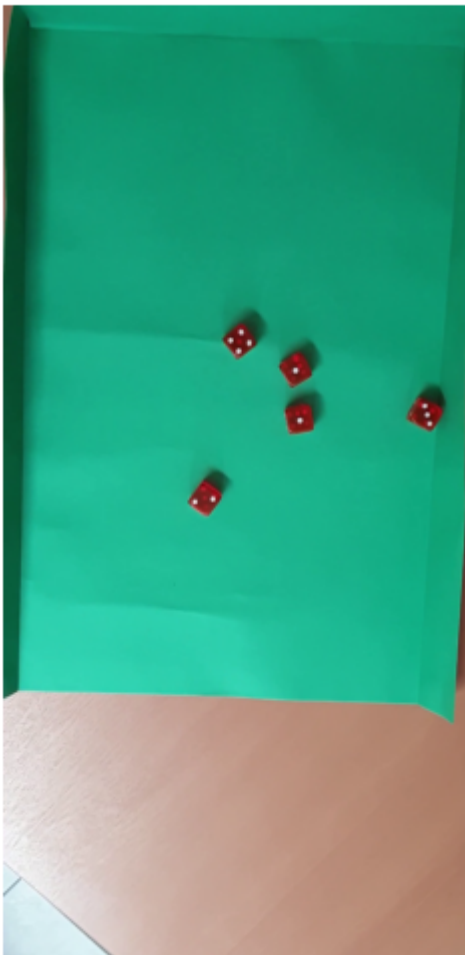
Introducción

En este trabajo se procesan videos que registran tiradas de **cinco dados rojos** sobre un fondo verde.

Los objetivos son detectar automáticamente un **frame donde los dados se encuentren detenidos**, contar el valor de cada dado y generar un **video de salida** donde se identifique el valor de cada dado mientras está en reposo.

La solución fue implementada utilizando técnicas clásicas de Procesamiento Digital de Imágenes, priorizando robustez, simplicidad e interpretabilidad, sin recurrir a machine learning.

Original



Descripción general de la solución

Para cada video, el sistema analiza **solo una porción avanzada del mismo** (aproximadamente entre el 45 % y el 95 % del total de frames), donde es más probable que los dados ya hayan terminado de rodar.

Esto reduce el tiempo de procesamiento y evita detecciones incorrectas durante la tirada.

En cada frame candidato se intenta:

- Detectar **exactamente cinco dados**
- Contar los puntos visibles en cada uno
- Verificar que todos los valores estén dentro del rango válido (1 a 6)

Entre los frames que cumplen estas condiciones, se selecciona el **más confiable**, priorizando aquel con mayor cantidad de dados detectados.

El proyecto también posee **visualización interactiva de las etapas intermedias** para el usuario, para mostrar las etapas de procesamiento.

Segmentación y detección de dados

La detección de los dados se basa en una **segmentación por color**, aprovechando que los dados son rojos y el fondo tiene menor componente roja.

Para ello se construye un **mapa de rojez**, Este mapa se calcula operando sobre los canales del espacio BGR mediante la siguiente fórmula aritmética aplicada pixel a pixel:

$$\text{MapaRojez} = R - \max(G, B)$$

Donde R, G y B representan la intensidad de cada canal. Esta operación permite desacoplar la información cromática de la intensidad luminosa:

que resalta las zonas donde el rojo predomina sobre los otros canales.

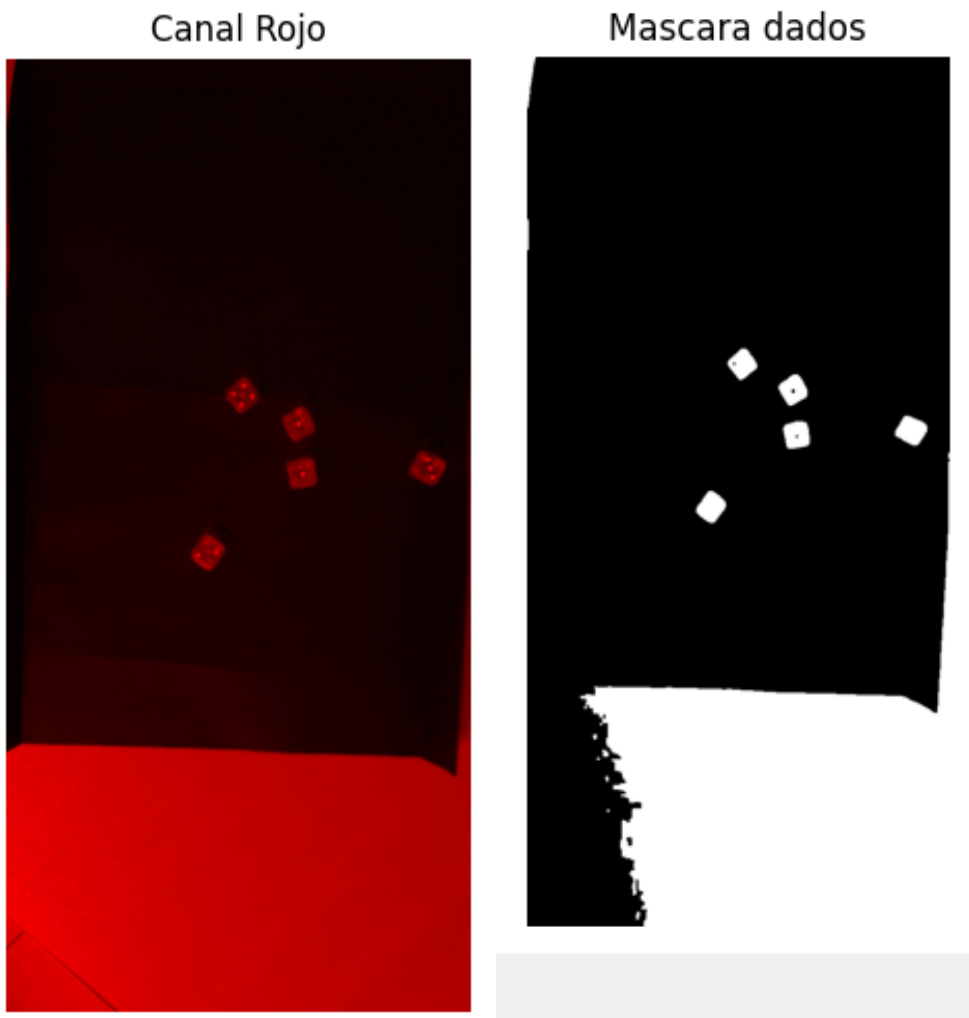
Sobre este mapa se aplica:

- Un umbral automático (cuando es posible)
- Un umbral fijo como respaldo en casos difíciles
- Operaciones morfológicas simples para eliminar ruido

Luego se detectan componentes conectadas y se filtran usando criterios geométricos básicos, como:

- Relación de aspecto cercana a un cuadrado
- Área coherente con el tamaño esperado de un dado

Si se detectan más de cinco candidatos, se seleccionan los cinco más consistentes según su tamaño.



Conteo de puntos

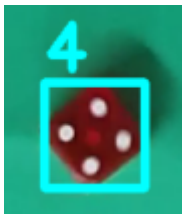
Una vez detectados los dados, se recorta la región correspondiente a cada uno y se realiza el conteo de puntos.

El procedimiento es el siguiente:

- Conversión a escala de grises
- Umbral adaptativo para manejar variaciones de iluminación
- Eliminación de bordes para evitar falsas detecciones
- Filtrado de regiones pequeñas o irregulares

Se consideran válidos únicamente los puntos que presentan un tamaño adecuado y una forma aproximadamente circular.

La cantidad de regiones válidas detectadas corresponde al valor del dado.



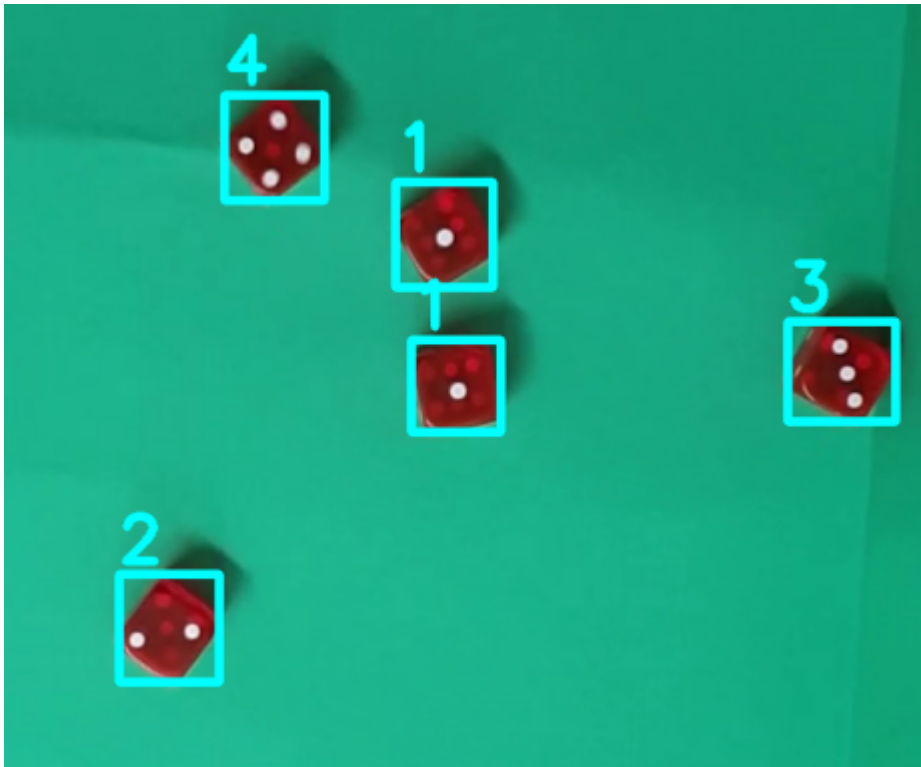
Validación y selección del frame final

Un frame se considera válido **únicamente** si:

- Se detectan **exactamente cinco dados**
- Todos los dados tienen un valor entre **1 y 6**

Este criterio evita aceptar frames con dados todavía en movimiento o con conteos incorrectos.

Entre los frames válidos, se selecciona el que presenta el **mayor puntaje total**, lo que ayuda a descartar detecciones parciales o inestables.



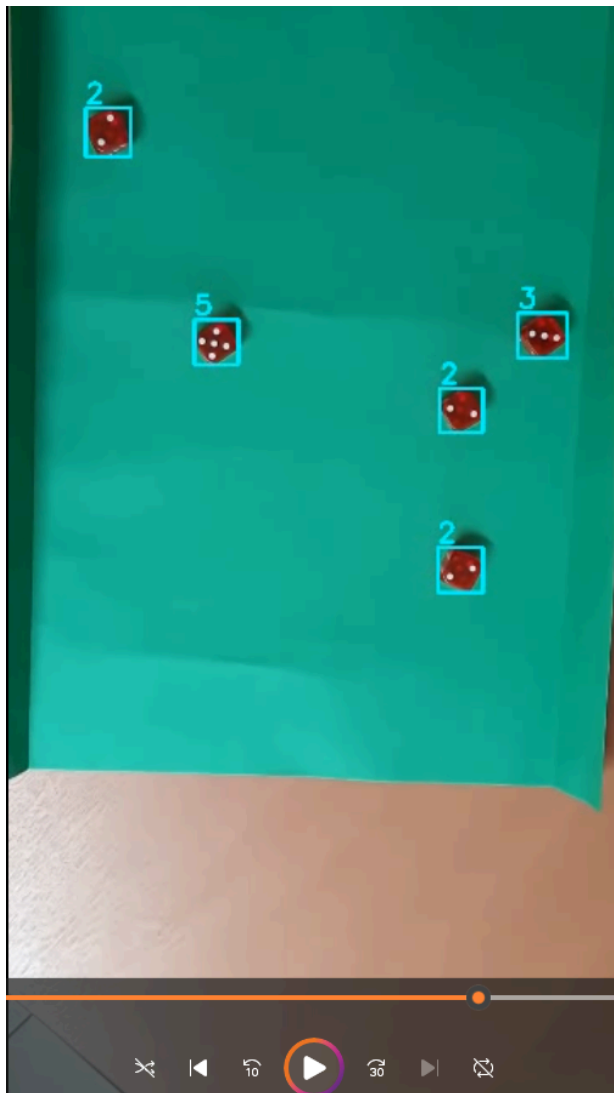
Generación del video de salida

Como etapa final, el sistema puede generar un **video procesado**, donde los dados aparecen identificados únicamente cuando se encuentran estables.

Para esto se utiliza un seguimiento simple basado en:

- Comparación de posiciones entre frames consecutivos
- Un umbral de distancia para decidir si un dado permanece quieto

El video de salida visualiza los **rectángulos delimitadores (bounding boxes)** y el valor numérico de cada dado. Además, este se genera en una resolución reducida para optimizar tanto el tamaño del archivo final como el tiempo de procesamiento.



Conclusión

Se desarrolló un sistema automático capaz de detectar **cinco dados** en videos de tiradas, identificar el momento en que se encuentran suficientemente estables y contar de forma confiable los puntos visibles en cada uno.

El enfoque, basado exclusivamente en técnicas clásicas de procesamiento de imágenes, resultó **robusto, eficiente e interpretable**, cumpliendo con los requisitos del trabajo sin necesidad de utilizar aprendizaje automático.