



# **INFORME TRABAJO PRÁCTICO 3**

## **PROCESAMIENTO DE IMÁGENES**

**ESTEVA MATIAS (E-1253/1)**

**PRIETO TOBIAS (P-5260/4)**

**TECNICATURA UNIVERSITARIA EN INTELIGENCIA ARTIFICIAL  
2025**

## Introducción

Durante la cursada de la materia, adquirimos diversas herramientas de procesamiento digital de imágenes. Inicialmente, trabajamos con imágenes generadas en un entorno controlado para comprender el funcionamiento básico de estas herramientas. Posteriormente, aplicamos estos conocimientos a imágenes capturadas en el mundo real, donde los patrones son más complejos y menos lineales. Con la incorporación de nuevas técnicas, logramos extraer el análisis inicial al ámbito real, implementando la detección de objetos, formas y colores. Este proceso nos ha llevado al punto actual.

Especificamente, en este último trabajo, el desafío fue aplicar todo el repertorio de herramientas y conocimientos a una secuencia de imágenes reales, es decir, a un video capturado en la vida real. Esto supuso una prueba integral de todo lo aprendido hasta el momento.

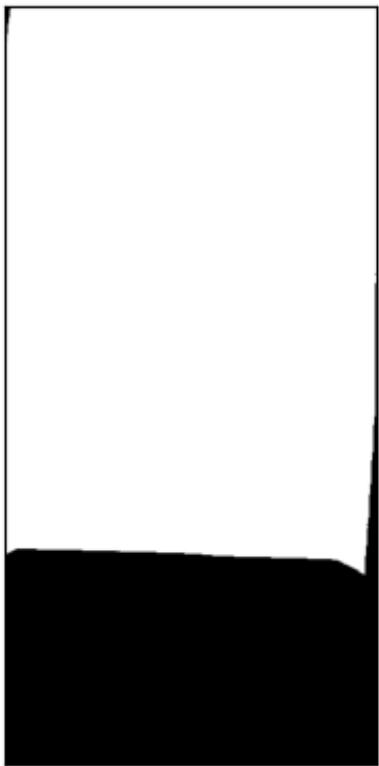
## Problema 1 | Cinco dados

En este ejercicio final, se nos brindan una serie de videos de tiradas de dados, concretamente se trata de cuatro videos grabados desde una posición estática de cámara donde se ve como se arrojan cinco dados de color rojo sobre una mesa que posee una gran superficie de color verde.

El objetivo de este trabajo es poder identificar cuál es el puntaje obtenido por el usuario en cada tirada de dados, detectando inicialmente los instantes donde los mismos se encuentran estáticos para poder llevar a cabo el conteo de forma correcta, es por ello que para la resolución del problema consideramos crucial utilizar la mayor cantidad de información presente en los videos a nuestro favor, a continuación detallamos las etapas que transcurren en cada una de las detecciones para poder lograr nuestro cometido.

### Recorte de la región de interés.

Como dijimos antes, los videos fueron tomados desde la misma posición fija, es por ello que para simplificar el análisis posterior decidimos realizar un recorte de la imagen centrándonos en la región de interés, la cual se trata de la superficie verde sobre la que se arrojan los dados, es por ello que mediante el uso del espacio de color HSV, somos capaces de crear una máscara detectando el color verde del fondo, simplificando un poco la cantidad de información inicial del planteo.



Máscara binaria



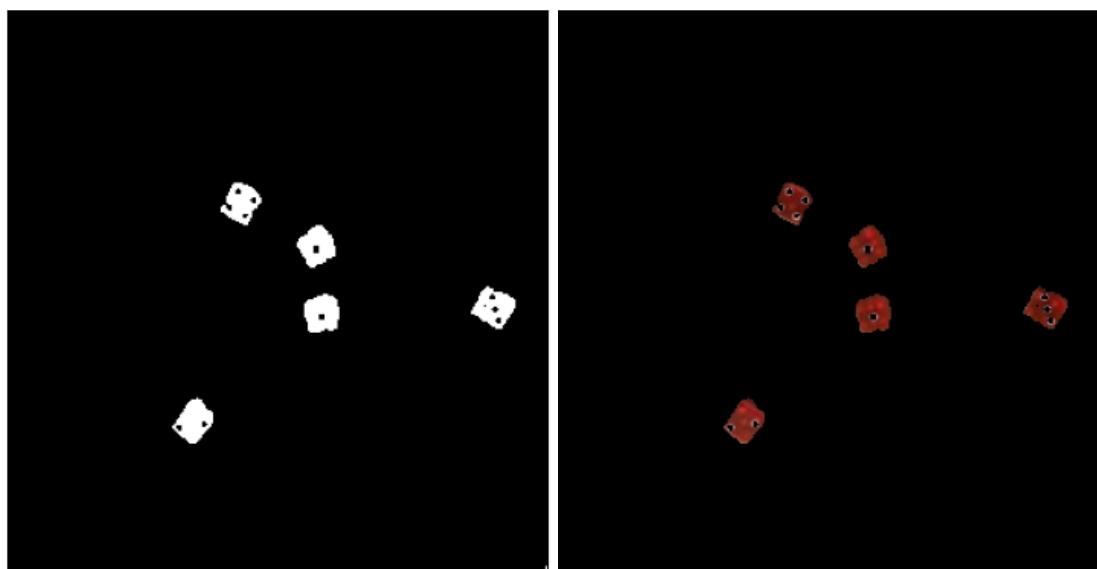
Mascara en video

## Detección de frames estáticos

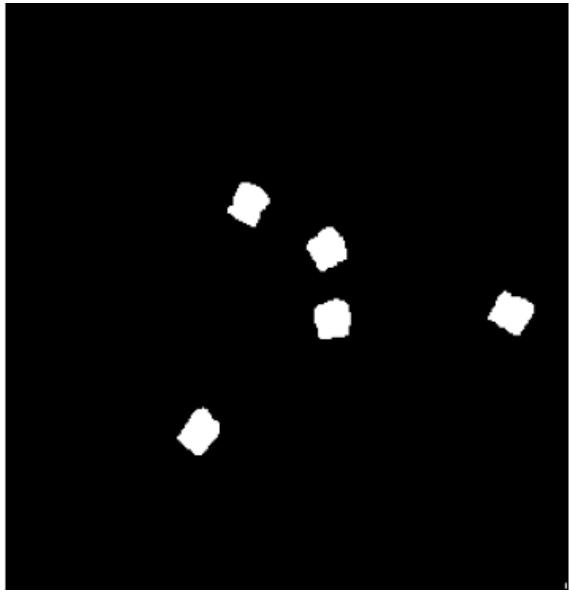
Continuando con la resolución, una vez que tenemos recortada la región de interés, debemos ser capaces de poder detectar cuando los datos se encuentran por así decirlo en su estado final, es decir, quietos luego de ser arrojados, es por ello que para lograr dicha detección creamos una función cuyo objetivo es recorrer todos los frames, previamente recortados, transformándolos a escala de grises y aplicando un desenfoque gaussiano para reducir ruido, una vez hecho eso calcula la diferencia absoluta entre el frame actual y el anterior, teniendo como referencia de movimiento un umbral definido arbitrariamente para detectar o no dicha estabilidad. Luego si el video se mantiene 'quieto' por más de una cierta cantidad de frames, es decir otro umbral, devuelve una lista de números de frame donde los datos están estáticos.

## Procesamiento de frames estáticos

Siguiendo con la salida de la función anterior, una vez que obtenemos los frames donde hay una estabilidad entre cuadros, procedemos a analizarlos. Nuestro objetivo aquí es encontrar en la imagen la ubicación de los dados rojos y luego ser capaces de contabilizar el puntaje obtenido de cada uno, por ello iniciamos con una detección del color de los mismos mediante el formato HSV, donde el color rojo ocupa los extremos del mismo por lo que creamos dos índices y luego los unificamos con un operador 'or'. Además debemos detectar una saturación alta del color para descartar automáticamente otras tonalidades más lavadas del mismo como podrían ser las presentes en la mano por ejemplo. Al finalizar dicho proceso obtenemos una máscara como la siguiente:

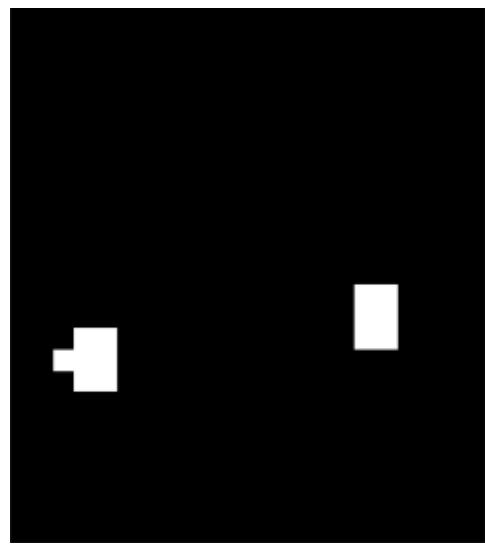
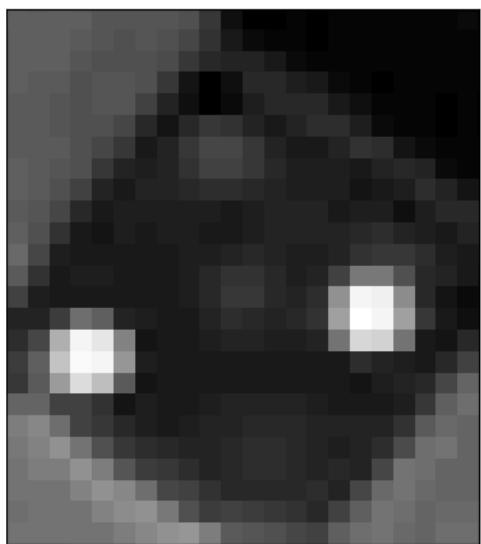


Observando los resultados podemos apreciar que detectamos correctamente todos los dados, pero particularmente algunos presentan algunas irregularidades en cuanto a formas no cerradas, para unificar ello y realizar una correcta detección de los puntos obtenidos, es que realizamos un paso previo aplicando morfología, particularmente de cierre, para poder dilatar y luego erosionar esas irregularidades y así contar con formas más uniformes, obteniendo el siguiente resultado:



## Análisis interno

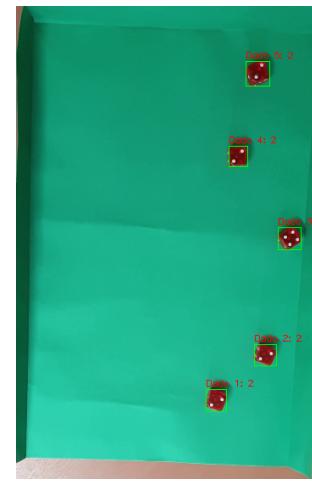
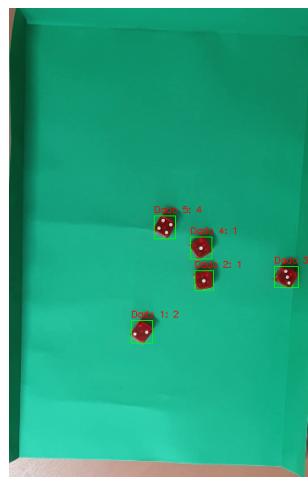
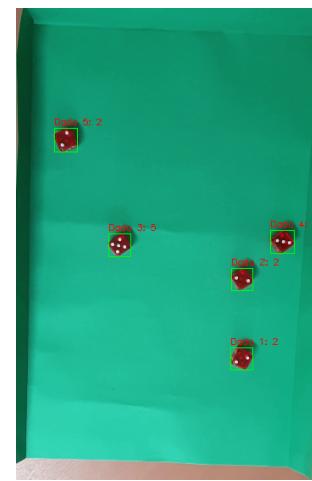
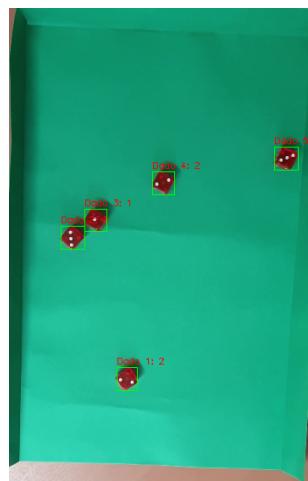
Una vez que sabemos ubicar a cada uno de los dados en la imagen, procedemos a contar en cada uno de ellos la puntuación obtenida y es allí donde realizaremos un conteo mediante componentes conectadas, pero antes se recorta la región del dado pero usando la imagen en escala de grises, debido a que ya no nos interesa el color porque sabemos que estamos trabajando sobre un dado, y además nos es muy útil porque al trabajar en dicha escala los puntos blancos de los dados son de los valores más elevados, por lo que podemos aplicar un umbralado para binarizar la región y quedarnos únicamente con lo que nos interesa, es decir todo lo que tenga un valor más alto del umbral se vuelve blanco (los puntos) y todo lo demás se vuelve negro (el cuerpo del dado).



Teniendo la segunda imagen podemos ahora sí aplicar la función encargada de detectar componentes conectados para contabilizar los puntos obtenidos en el mismo.

## Generar video

Una vez detallado toda la etapa de análisis y procesamiento, debemos guardar el nuevo video mostrando la detección automática de los dados y sus puntos, para ello se lee el video original y se prepara un objeto VideoWriter para guardar el resultado, luego itera por todos los frames para ver si el frame actual está en la lista de frames quietos, si lo está lo procesa para obtener los valores, dibuja los rectángulos alrededor de los dados y escribe el texto con el número del dado asociado, por último guarda el frame procesado en el nuevo archivo de video. A continuación adjuntamos capturas de la detección realizada en el momento de estabilidad de frames:



## Conclusiones

A modo de conclusión de este trabajo práctico y de la asignatura, podemos afirmar que, además de adquirir valiosas herramientas para el futuro, hemos tomado conciencia de la enorme complejidad del mundo que nos rodea. Destacamos cómo conocimientos de fundamentos matemáticos y herramientas aparentemente "simples", vistos previamente en la carrera, pueden aplicarse de manera efectiva para resolver diversos problemas relacionados con el procesamiento de imágenes. Es relevante notar que utilizamos estas aplicaciones constantemente en nuestra vida diaria sin ser conscientes de su presencia, abarcando desde la detección de bordes en una imagen sencilla, pasando por la identificación automática de formatos de formularios, hasta la detección automatizada de carriles en un vehículo.