



Universidad Nacional de Rosario  
Facultad de Ciencias Exactas, Ingeniería y Agrimensura  
Tecnicatura Universitaria en Inteligencia Artificial  
Procesamiento de Imágenes I - IA 4.4



---

**Universidad Nacional de Rosario -  
Facultad de Ciencias Exactas, Ingeniería y Agrimensura**

## **Tecnicatura Universitaria en Inteligencia Artificial**

**Asignatura:** Procesamiento de Imágenes

# **TRABAJO PRÁCTICO No 3**

**Alumnos:** Accurso, Agustín  
Barbarroja, Federico  
Cena, Lautaro

**Equipo Docente:** Gonzalo Sad  
Juan Manuel Calle  
Joaquín Allione



El objetivo del presente trabajo práctico consiste en desarrollar un sistema de visión artificial capaz de procesar archivos de video, detectar automáticamente el momento en que los dados se detienen tras una tirada y cuantificar el valor obtenido en sus caras superiores.

### **Detección de estabilidad**

El análisis frame a frame era algo ineficiente y complicado de hacer, por lo que se optó primero por usar una detección de quietud basada en la diferencia absoluta entre frames consecutivos.

- Se calcula la diferencia absoluta entre el frame actual y el frame anterior
- Si la media de la diferencia es menor a un umbral (en nuestro caso, 3.0) durante una ventana de tiempo sostenida, el sistema cambia al estado "**DETENIDO**" y dispara el análisis de la imagen.

### **Detección de contornos**

Para la detección de los candidatos a dados, se optó por un enfoque basado en bordes

1. **Filtrado:** Conversión a escala de grises y *Gaussian Blur* para reducir el granulado.
2. **Detección de Bordes:** Aplicación de *Canny* para obtener la estructura geométrica.
3. **Morfología:** Se aplica una operación de **Dilatación** para cerrar los contornos y generar "bloques" sólidos candidatos.
4. **Binarización:** *Thresholding* de Otsu para separar los objetos del fondo.

### **Fusión de dados**

Se identificó un problema crítico : una dilatación alta fusionaba dados cercanos, mientras que una dilatación baja fragmentaba un solo dado en dos mitad.

Para resolver esto, se implementó una **Fusión Condicional por Tamaño:**

1. El sistema detecta inicialmente todas las regiones posibles.
2. Se aplica un algoritmo de **Intersección de Bounding box** con una tolerancia de 10px.
3. Si dos cajas se solapan, se evalúa el **Área Resultante:**
  - Si **Area\_Combinada < MAX\_AREA\_DADO\_UNICO (11000px)**: Se asume que es un solo dado fragmentado y **se fusionan**.
  - Si **Area\_Combinada > MAX\_AREA\_DADO\_UNICO**: Se asume que son dos dados independientes tocándose y **se mantienen separados**.

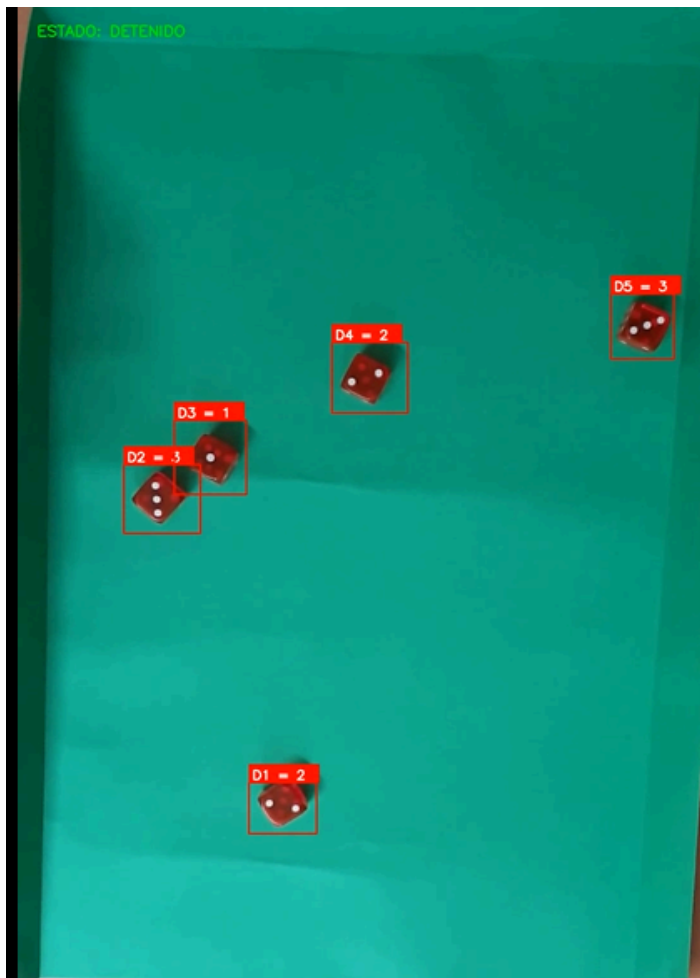
## Conteo de Pips

Para cada dado validado, se extrae la ROI y se aplica una umbralización estricta para aislar los puntos blancos. Se filtran los contornos internos por circularidad y área para descartar brillos o reflejos, sumando la cantidad de "pips" válidos detectados.

## Resultados

El sistema fue probado en las 4 secuencias de video.

- **Precisión:** Se logró una detección correcta del 100% de los dados en reposo, incluyendo casos de oclusión parcial y dados en contacto directo.
- **Velocidad:** El procesamiento se realiza en tiempo real, ya que el análisis pesado solo se ejecuta una vez por evento de estabilidad.
- **Visualización:** Se genera un video de salida que muestra el estado del sistema ("RODANDO" vs "DETENIDO"), las *bounding boxes* individuales, el ID del dado y su valor numérico.





## **Conclusión**

El trabajo práctico nos enseñó que a veces las técnicas puras de detección, como Canny o la Binarización, no son siempre totalmente eficientes en entornos complejos.

La implementación de la función de fusión condicional aplicada a los videos fue la clave para la distinción de objetos (dados) muy cercanos entre sí. Fue determinante para lograr la robustez del sistema final, permitiendo diferenciar correctamente entre un error de segmentación y una agrupación física de objetos.

El hecho de trabajar sobre escribiendo videos también fue algo complejo al principio, pero con el material de ayuda brindado pudimos superar las complicaciones.