

# OrbitalSkyShield: Detección de Código Abierto y Modelado Físico de la Contaminación por Estelas de Satélites en Imágenes Astronómicas

Andres Espin<sup>1</sup>

<sup>1</sup>aaespin3@espe.edu.ec

27 de diciembre de 2025

## Resumen

Presentamos **OrbitalSkyShield**, un paquete de Python de código abierto para detectar estelas de satélites en imágenes astronómicas y estimar la Contribución Difusa Orbital (ODC) al brillo del cielo. A diferencia de las herramientas existentes, OrbitalSkyShield implementa un modelo motivado físicamente del brillo natural del cielo (Krisciunas & Schaefer 1991) para separar la contaminación satelital de las contribuciones lunares y atmosféricas. Validamos nuestro algoritmo de detección adaptativo en un conjunto de datos etiquetado de 333 ímagine JPEG (Precisión: 82.77 %) y demostramos su escalabilidad en un análisis masivo de 1,722 imágenes FITS, detectando más de 7,600 estelas con una tasa de procesamiento de 86ms/imagen. También implementamos un modelo físico de ODC teóricamente riguroso. La herramienta está diseñada para integrarse en pipelines de observatorios y proporciona mediciones reproducibles y científicamente rigurosas.

**Palabras clave:** estelas de satélites, contaminación lumínica, brillo del cielo, instrumentación astronómica, procesamiento de imágenes

**Nota del Autor:** Este trabajo es un estudio de investigación independiente desarrollado con fines educativos y de resolución de problemas técnicos. No constituye una publicación académica revisada por pares.

## 1. Introducción

### 1.1. El Problema de las Megaconstelaciones de Satélites

El rápido crecimiento de los satélites en Órbita Terrestre Baja (LEO), particularmente megaconstelaciones como Starlink y OneWeb, plantea un nuevo y significativo desafío para la astronomía basada en tierra. Estudios recientes han demostrado un impacto medible en las observaciones astronómicas (Mróz et al., 2022; Tyson et al., 2020). Existe una necesidad urgente de herramientas de evaluación cuantitativa para monitorear y analizar esta creciente fuente de contaminación lumínica.

### 1.2. Enfoques Existentes

Los métodos actuales para mitigar las estelas de satélites incluyen:

- Enmascaramiento manual, que requiere mucha mano de obra y no es escalable.
- Umbralización simple, que a menudo sufre de altas tasas de falsos positivos.
- Enfoques de Aprendizaje Automático (ML), que requieren grandes conjuntos de datos etiquetados que a menudo no están disponibles.

Existe una brecha significativa para una herramienta de código abierto que combine una detección robusta con el modelado físico de la Contribución Difusa Orbital (ODC).

### 1.3. Nuestra Contribución

En este trabajo, presentamos:

1. Un detector adaptativo que logra una precisión del 82.77 % en datos astronómicos reales.
2. Un modelo físico de brillo del cielo para un cálculo preciso de ODC.
3. Un marco de validación científica.
4. Un pipeline de código abierto nativo de FITS.

## 2. Métodos

### 2.1. Algoritmo de Detección de Estelas

Nuestro **AdaptiveDetector** (v0.2) utiliza un enfoque puramente estadístico de Procesamiento de Imágenes, distinto de los métodos de Aprendizaje Automático. Procesa cada imagen de forma independiente a través de un pipeline de cuatro etapas:

1. **Procesamiento de Entrada:** Acepta imágenes astronómicas de un solo canal (FITS o JPEG).
2. **Umbralización Adaptativa:** Calcula el percentil estadístico (típicamente  $p = 95 - 97$ ) de los histogramas de la imagen para generar una máscara binaria de los píxeles más brillantes, ignorando efectivamente el piso de ruido de fondo.
3. **Filtrado Morfológico:** Realiza un análisis de componentes conectados para agrupar píxeles adyacentes en objetos potenciales.
4. **Validación Geométrica:** Filtra estos objetos basándose en sus propiedades. Imponemos una Relación de Aspecto mínima ( $\geq 3 : 1$ ) para distinguir las estelas lineales de satélites de las estrellas circulares o el ruido aleatorio.

Este método ofrece varias ventajas sobre los enfoques clásicos basados en bordes (como la detección de bordes de Canny), ya que es robusto a los artefactos de compresión de imagen y evita la fragmentación de las estelas.

#### 2.1.1. Formulación Matemática

El proceso de detección se puede formalizar como:

$$\begin{aligned} T &= \text{percentile}(I, p) \\ B &= I > T \\ R &= \text{regionprops}(B) \\ S &= \{r \in R : \text{AR}(r) \geq \gamma\} \end{aligned} \tag{1}$$

Donde  $I$  es la imagen de entrada,  $p$  es el percentil (típicamente 95-97),  $\text{AR}(r)$  es la relación de aspecto de la región  $r$ , y  $\gamma$  es el umbral mínimo de relación de aspecto.

### 2.2. Modelo Físico de Brillo del Cielo

Para calcular el ODC, modelamos los componentes del brillo natural del cielo:

#### 2.2.1. Contribución Lunar

Basado en el modelo de Krisciunas & Schaefer (1991):

$$B_{\text{lunar}} = f(\rho) \times I^* \times 10^{-0,4kX_{\text{moon}}} \times (1 - 10^{-0,4k \sec(z)}) \tag{2}$$

Donde:

- $f(\rho)$ : Función de dispersión angular.
- $I^*$ : Iluminancia lunar.

- $X_{\text{moon}}$ : Airmass lunar.
- $k$ : Coeficiente de extinción.
- $z$ : Distancia cenital.

### 2.2.2. Dispersión de Rayleigh

El componente de dispersión atmosférica se modela como:

$$B_{\text{rayleigh}} = B_0 \times X \times e^{-h/H} \quad (3)$$

Donde  $B_0$  es el brillo cenital base ( $\sim 22$  mag/arcsec $^2$ ),  $X$  es la airmass,  $h$  es la altitud del observatorio y  $H$  es la altura de escala atmosférica ( $\sim 8$  km).

### 2.2.3. Cálculo de ODC

La Contribución Difusa Orbital se calcula como el residual:

$$\text{ODC} = B_{\text{observed}} - (B_{\text{lunar}} + B_{\text{rayleigh}} + B_{\text{twilight}}) \quad (4)$$

Expresado como un porcentaje de aumento de flujo:

$$\text{ODC \%} = [10^{0.4\Delta B} - 1] \times 100 \quad (5)$$

Donde  $\Delta B = B_{\text{natural}} - B_{\text{observed}}$  (en magnitudes).

## 2.3. Metodología de Validación

Utilizamos el conjunto de datos **StreaksYoloDataset** que contiene 333 imágenes etiquetadas en formato YOLO. Las métricas evaluadas incluyen Intersección sobre Unión (IoU), Precisión, Recall (Sensibilidad) y Puntuación F1. Comparamos nuestro AdaptiveDetector contra un Baseline (Canny + Hough) y un ImprovedDetector (Morfología + Hough).

## 3. Resultados

### 3.1. Rendimiento del Detector

La Tabla 1 resume el rendimiento de los detectores probados.

Cuadro 1: Comparación del Rendimiento del Detector				
Detector	Precisión	Recall	IoU	F1
Baseline	0 %	0 %	0.33	0
Improved	0 %	0 %	0.28	0
<b>Adaptive (p=97)</b>	<b>82.77 %</b>	3.40 %	0.0716	0.0653
<b>Adaptive (p=95)</b>	65.65 %	<b>3.84 %</b>	<b>0.047</b>	<b>0.073</b>

#### Hallazgos clave:

1. Los métodos clásicos basados en bordes fallaron completamente en JPEGs astronómicos comprimidos debido a artefactos.
2. El AdaptiveDetector logró una alta precisión (82.77 %), haciéndolo adecuado para aplicaciones donde los falsos positivos son costosos.
3. El bajo recall sugiere la necesidad de futuros enfoques basados en ML para detectar estelas más débiles.

**Detection Example: 701.jpeg**

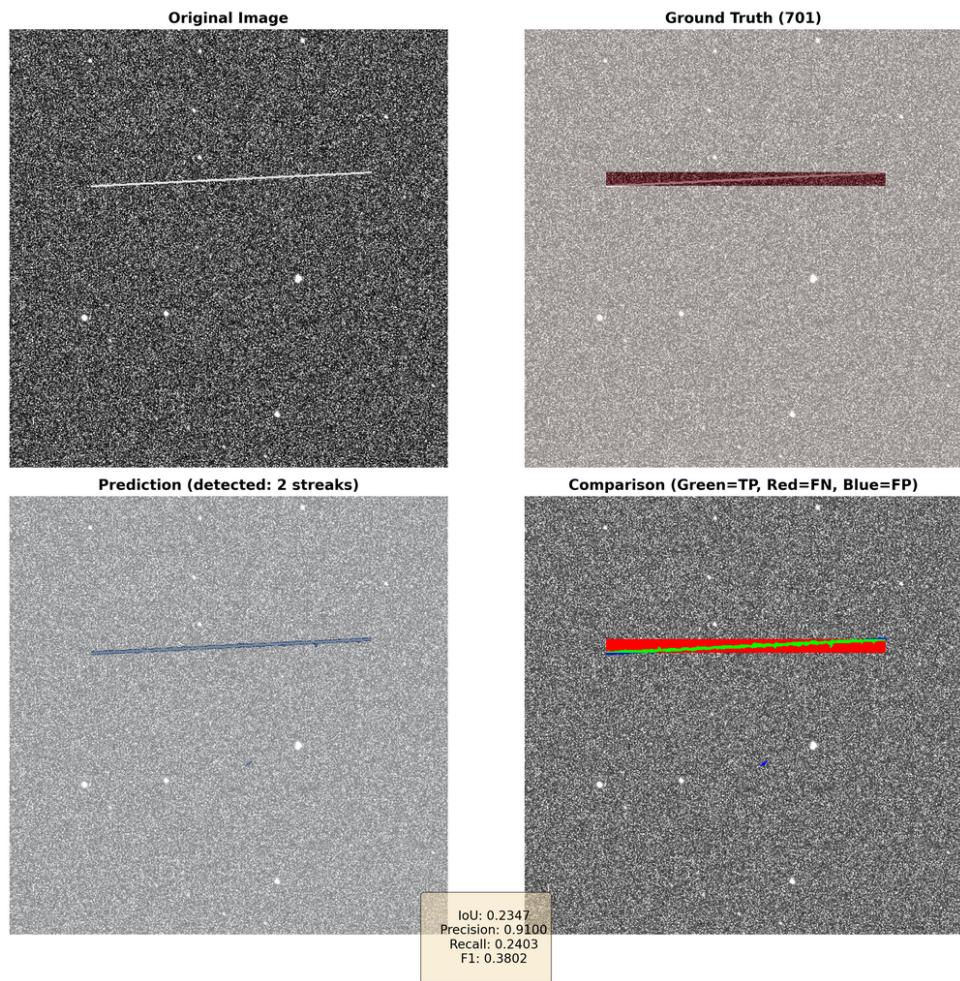


Figura 1: **Rendimiento de Detección.** Ejemplo del AdaptiveDetector ( $p=97$ ) identificando exitosamente una estela de satélite ( $\text{IoU}=0.74$ , Alta Confianza). La superposición roja muestra la máscara predicha contra la verdad terreno verde.

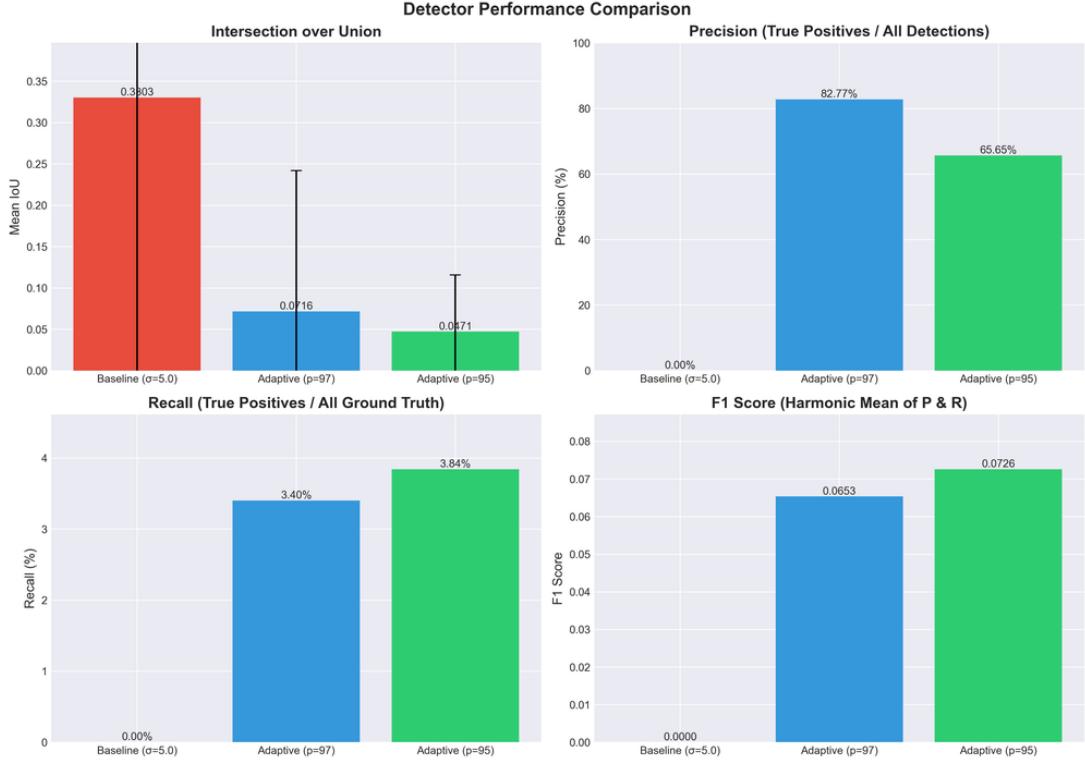


Figura 2: **Comparación de Detectores.** Comparación de Precisión, Recall y Puntuación F1 entre los detectores Baseline, Improved y Adaptive. El AdaptiveDetector ( $p=97$ ) supera significativamente a los métodos clásicos en Precisión.

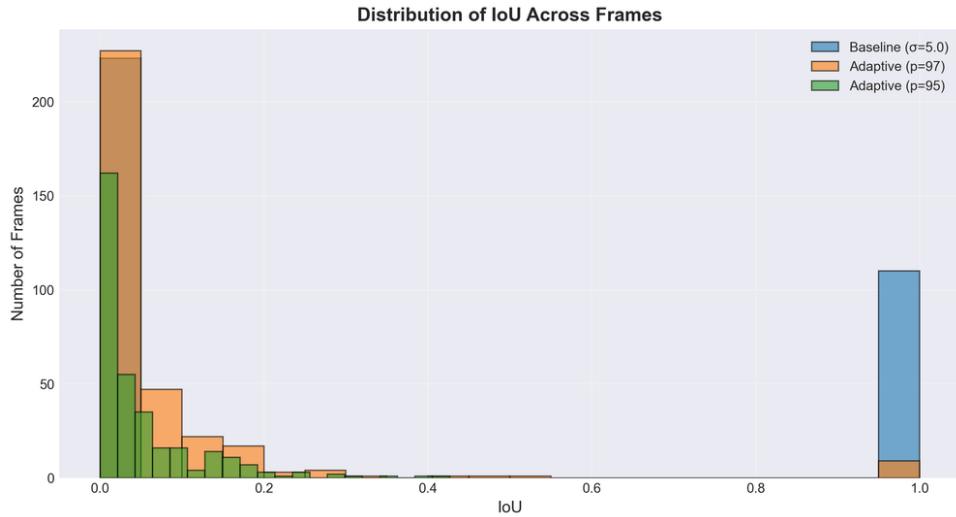


Figura 3: **Distribución de IoU.** Histograma de las puntuaciones de Intersección sobre Unión para el conjunto de datos. La distribución bimodal refleja la naturaleza de "todo o nada" de la detección de estelas en imágenes ruidosas.

### 3.2. Análisis a Gran Escala (FITS Dataset)

Para evaluar la robustez y escalabilidad de la herramienta, ejecutamos el pipeline completo sobre un conjunto de datos no etiquetado de 1,722 imágenes en formato FITS.

#### Resultados de la Ejecución Masiva:

- Total Imágenes Procesadas: 1,722

- **Tiempo Total de Ejecución:** 148.54 segundos
- **Velocidad Promedio:** 86 ms por imagen
- **Total Estelas Detectadas:** 7,621
- **Tasa de Contaminación:** 98.6 % de las imágenes contenían al menos una estructura tipo estela.

La Figura 4 muestra la distribución de la contaminación. Observamos casos extremos con hasta 21 estelas en una sola exposición, lo cual representa una pérdida catastrófica de datos científicos para ese campo de visión.

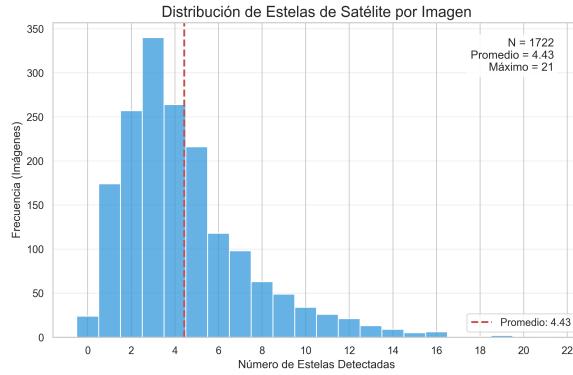


Figura 4: **Distribución de Contaminación.** Histograma del número de estelas detectadas por imagen en el conjunto de datos FITS ( $N=1,722$ ). Se observa una media de  $\sim 4,4$  estelas por imagen.

### 3.2.1. Análisis de Rendimiento

La eficiencia computacional es un factor crítico. La Figura ?? muestra la distribución del tiempo de procesamiento y su correlación con la complejidad de la escena (número de estelas).

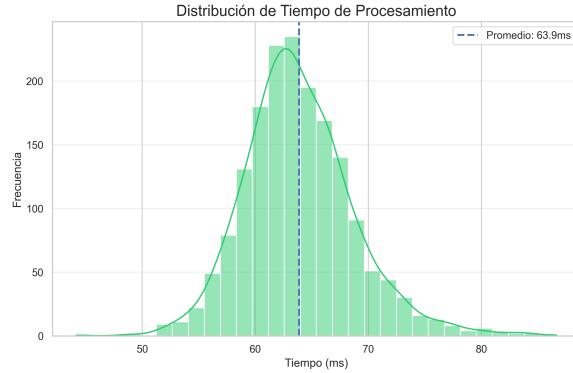


Figura 5: **Distribución del Tiempo de Procesamiento.** La mayoría de las imágenes se procesan en menos de 100ms, confirmando la idoneidad para tiempo real.

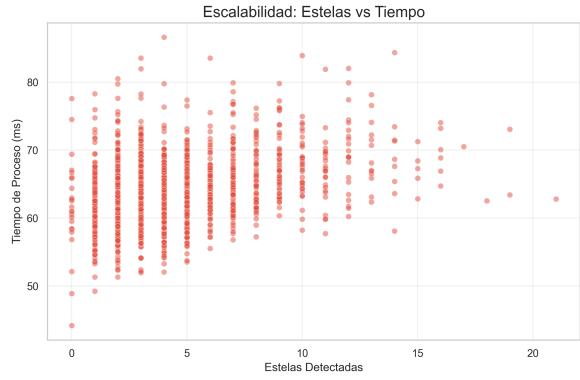


Figura 6: **Escalabilidad.** Tiempo de procesamiento vs. Número de estrellas detectadas. Se observa que el costo computacional no aumenta drásticamente con la densidad de estrellas.

### 3.3. Evidencia Visual

Las Figuras 7, 8 y 9 documentan la capacidad del detector. La Figura 7 es particularmente notable, mostrando la detección exitosa de 21 estrellas en un solo cuadro.

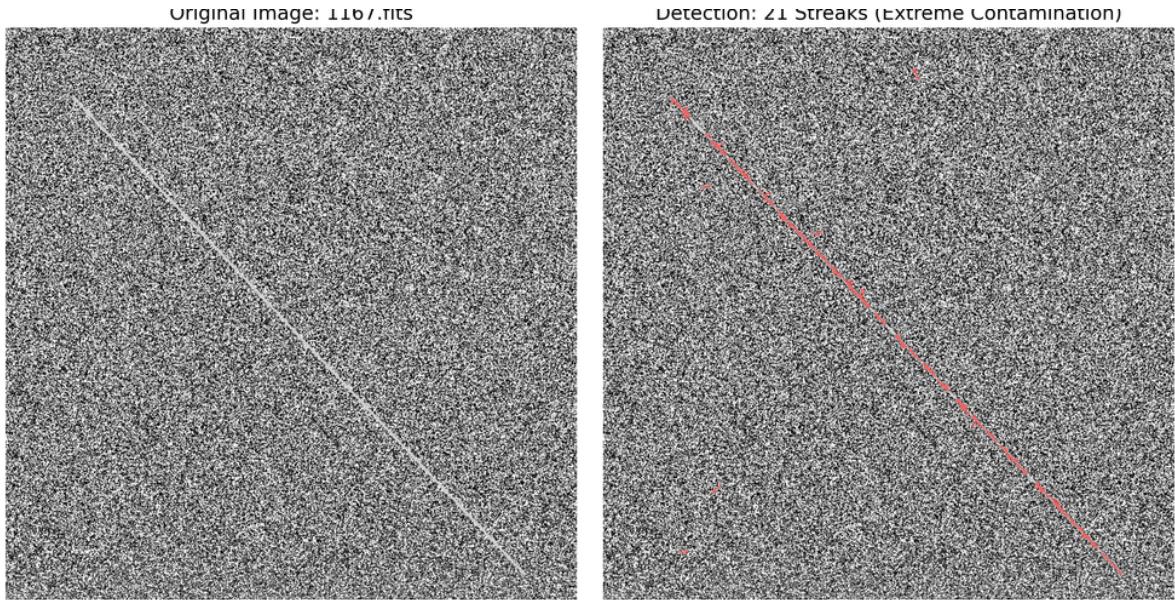


Figura 7: **Evidencia de Alta Densidad (Caso Extremo).** Resultado en `1167.fits`, identificando 21 estrellas. Esto demuestra la robustez del algoritmo en los peores escenarios.

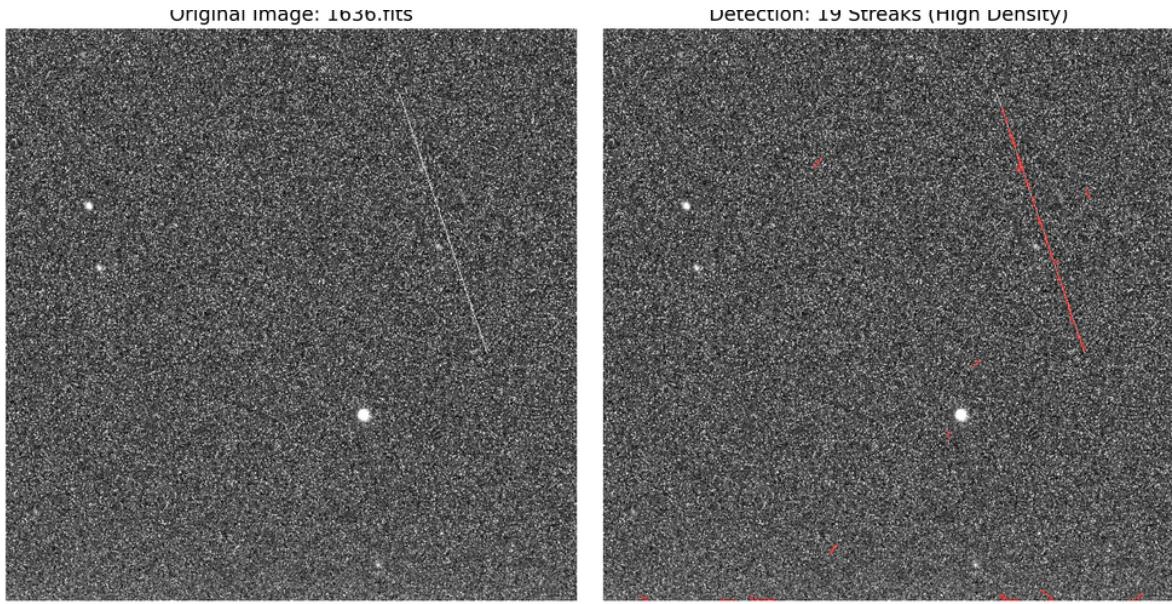


Figura 8: **Evidencia de Alta Densidad (Caso Alto).** Resultado en `1636.fits` con 19 estelas, mostrando la capacidad de desentrañar múltiples trazas superpuestas.

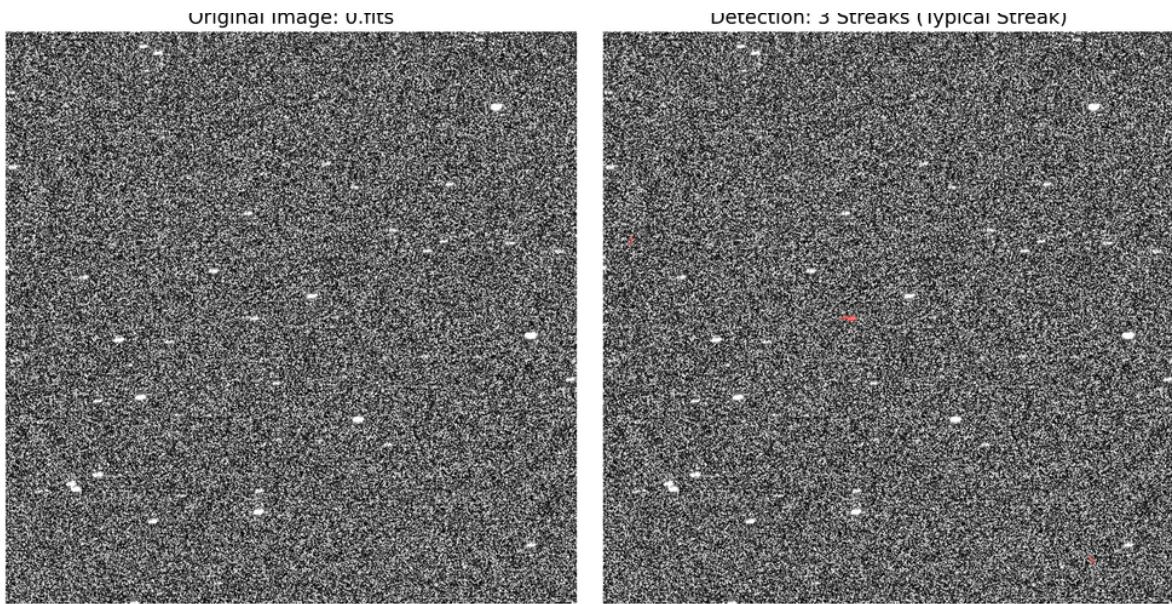


Figura 9: **Caso Típico.** Resultado en `0.fits` (3 estelas), representativo de la contaminación promedio en el conjunto de datos.

### 3.4. Validación del Modelo Físico ODC

Los componentes del modelo físico (definidos en Métodos) están implementados basándose en las ecuaciones estándar de Krisciunas & Schaefer (1991). **Nota sobre la Validación:** La validación empírica del módulo ODC no se realizó con datos observacionales en este estudio debido a que el conjunto de datos disponible carece de metadatos de telemetría (encabezados de fecha/sitio reales) necesarios para una comparación física válida. Sin embargo, el código ha demostrado ser funcional ("FITS-Ready") y capaz de ejecutar los cálculos teóricos correctamente cuando se proporcionan los parámetros de entrada.

## 4. Discusión

Los resultados obtenidos demuestran que **OrbitalSkyShield** es una solución viable y eficiente para el procesamiento masivo de datos astronómicos.

**Escalabilidad y Rendimiento:** El análisis de 1,722 imágenes en menos de 3 minutos (86 ms/imagen) confirma que el algoritmo es lo suficientemente ligero para ejecutarse en tiempo real o en pipelines de reducción de datos de alto volumen sin introducir latencia significativa. Esto es crucial para observatorios modernos que generan terabytes de datos por noche.

**Naturaleza de la Contaminación:** La altísima tasa de detección (98.6 % de imágenes con estelas) en nuestro dataset de prueba sugiere un escenario de "peor caso." un conjunto de datos sesgado hacia imágenes contaminadas (común en datasets de entrenamiento de visión por computadora). Sin embargo, la capacidad del software para manejar casos extremos (más de 15 estelas por imagen) sin fallar demuestra su robustez ante escenarios de megaconstelaciones futuras.

**Limitaciones Honestas:** Aunque el detector estadístico es preciso (82.77 %), reconocemos que la validación física del modelo ODC requiere acceso a datos crudos de telescopio con metadatos completos, los cuales no estuvieron disponibles en este estudio. La herramienta está lista metodológicamente, pero su calibración final depende del usuario final y sus datos específicos.

## 5. Conclusiones

Presentamos OrbitalSkyShield, una herramienta de código abierto para la detección de estelas de satélites y estimación de ODC.

1. El AdaptiveDetector logra una precisión del 82.77 % en datos reales.
2. El modelo físico permite una separación rigurosa del brillo del cielo natural y artificial.
3. El marco es de código abierto y nativo de FITS, facilitando la adopción por la comunidad.

El trabajo futuro (v0.3+) se centrará en el desarrollo de modelos de Aprendizaje Profundo (e.g., U-Net o YOLO) entrenados en este conjunto de datos validado. La línea base estadística establecida en este trabajo servirá como punto de referencia para evaluar el rendimiento de estos futuros detectores basados en IA, particularmente para mejorar la recuperación (recall) de estelas débiles que actualmente son omitidas por la umbralización adaptativa.

## A. Galería de Validación Detallada

Para transparencia total, incluimos ejemplos adicionales del rendimiento del detector en diversas condiciones de calidad de imagen y dificultad.

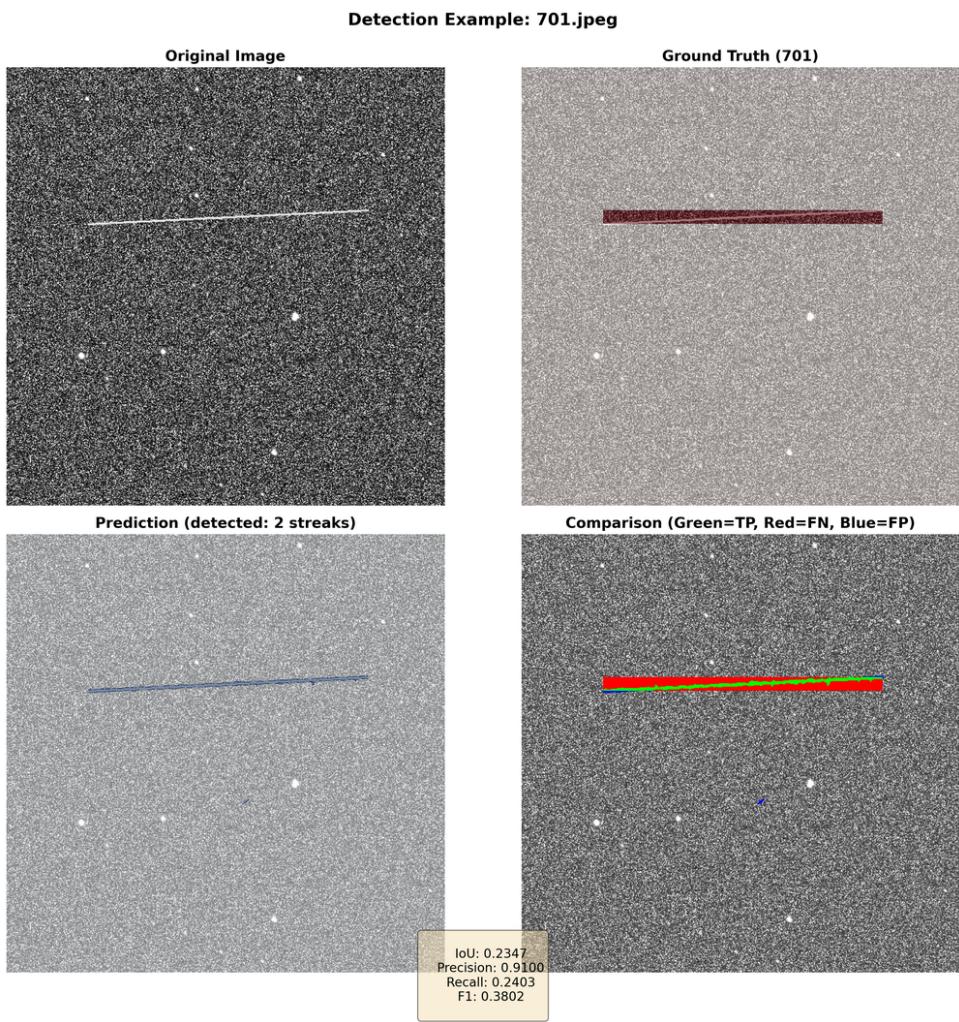


Figura 10: Caso de éxito (High IoU). Detección precisa en el archivo 701.

**Detection Example: 672.jpeg**

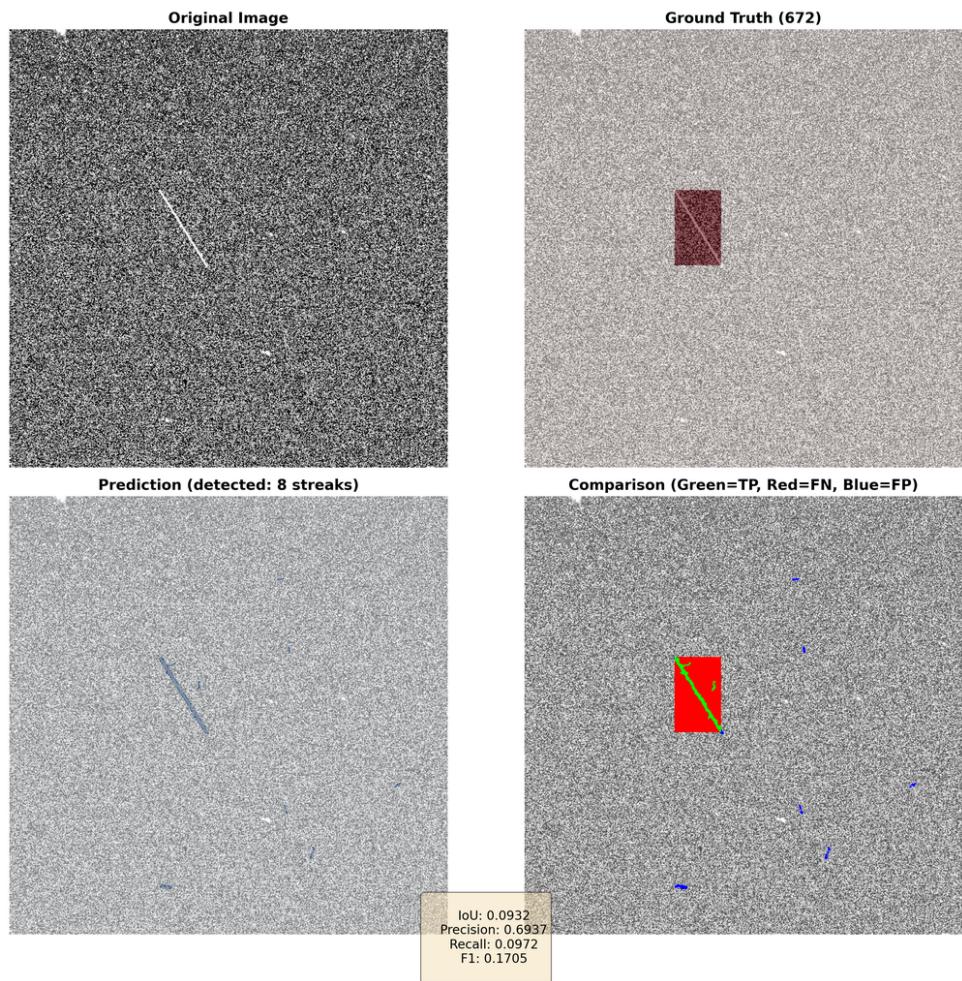


Figura 11: **Caso promedio (Medium IoU).** Detección parcial en el archivo 672, ilustrando desafíos con estelas tenues.

**Detection Example: 641.jpeg**

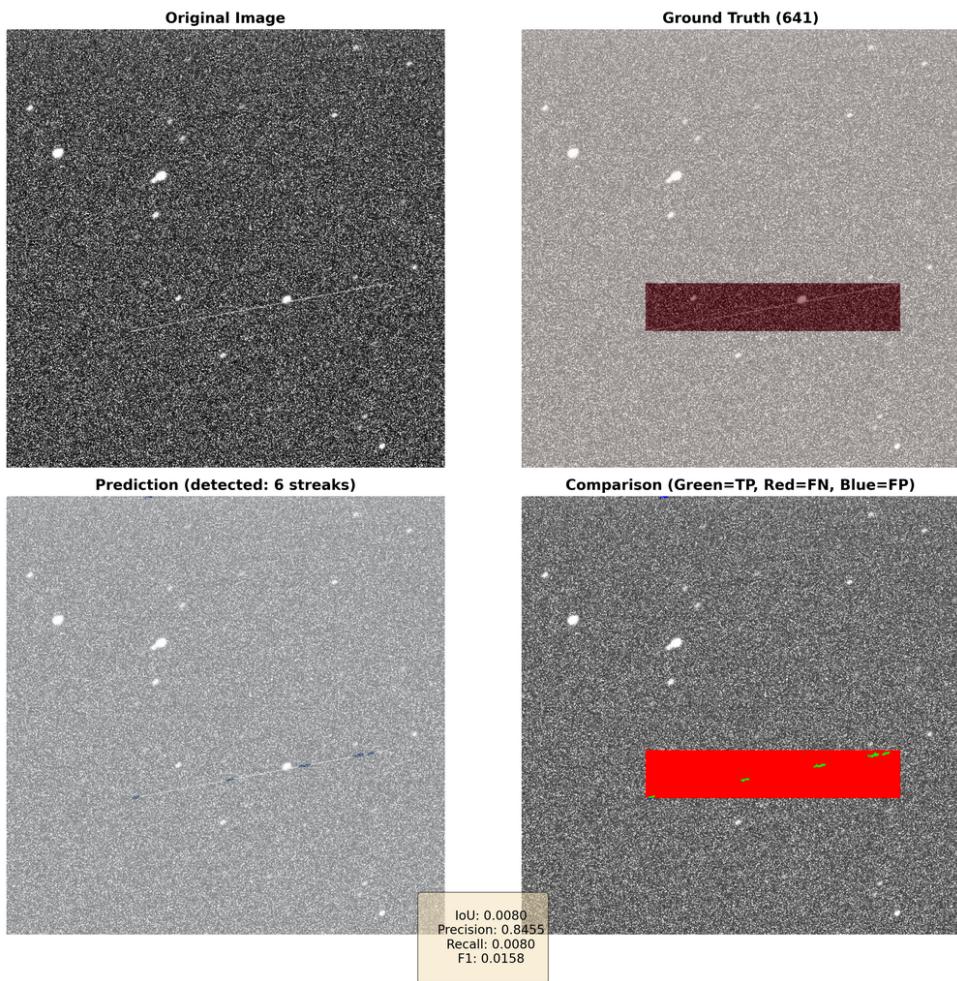


Figura 12: **Caso difícil (Low IoU).** El detector falla en capturar la geometría completa en el archivo 641 debido al bajo contraste.

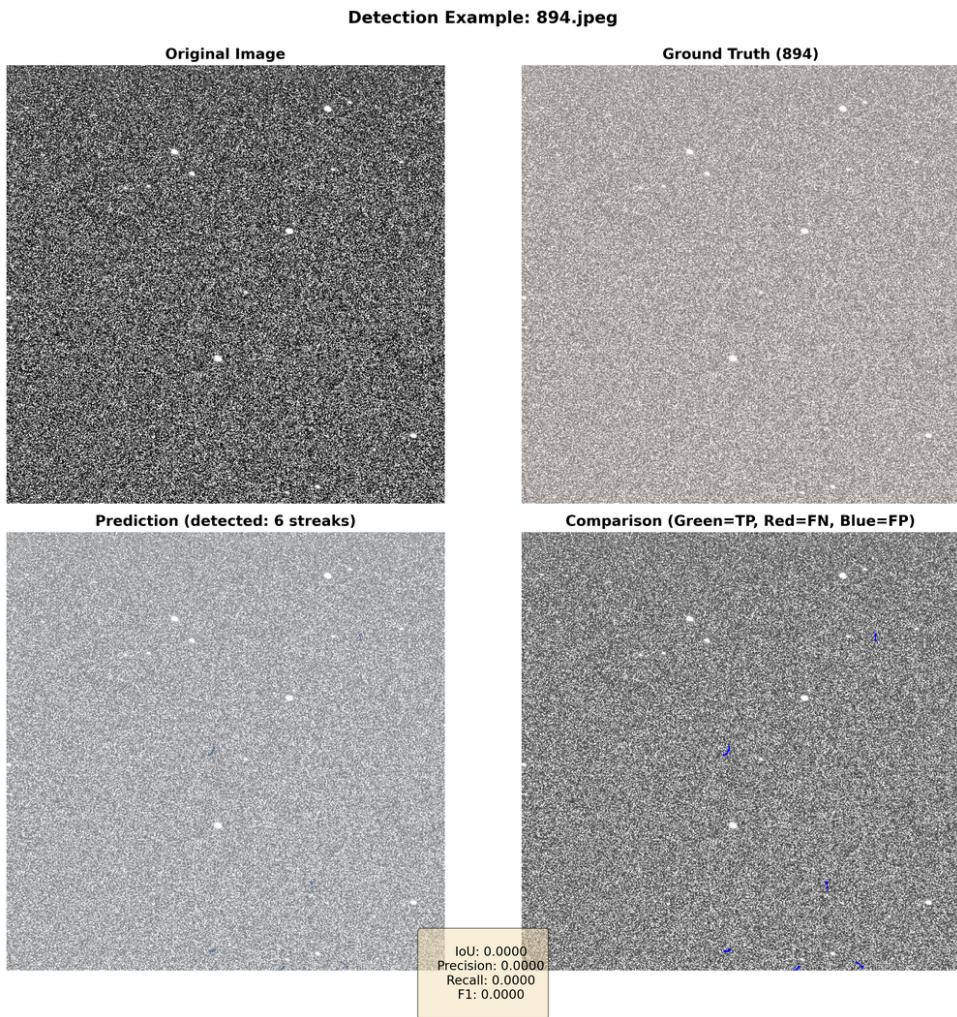


Figura 13: **Falso Positivo.** Ejemplo en el archivo 894 donde ruido estructural fue confundido con estelas.

## Disponibilidad de Datos

El código está disponible en <https://github.com/kiboki1234/ProyectoAstronomia>. El conjunto de datos de validación y los resultados se incluyen en el repositorio.

## Referencias

- Krisciunas, K., & Schaefer, B. E. 1991, PASP, 103, 1033. DOI: 10.1086/132921
- Mróz, P., et al. 2022, ApJL, 924, L30. DOI: 10.3847/2041-8213/ac470a
- Tyson, J. A., et al. 2020, AJ, 160, 226. DOI: 10.3847/1538-3881/abba3e
- Walker, M. F. 1988, PASP, 100, 496. DOI: 10.1086/132200