

Detección Eficiente de Fuego mediante Fidelidad Espacial: Desafiando el Paradigma de Deep Learning con Modelos Parsimoniosos

(Efficient Fire Detection via Spatial Fidelity: Challenging the Deep Learning Paradigm with Parsimonious Models)

Juan David Gueixpal Ayala
Facultad de Ingeniería Eléctrica y Electrónica (FIEE)
Universidad Veracruzana
Veracruz, México

Resumen—Este estudio presenta un enfoque de ingeniería de características incremental e iterativa para la detección de incendios en conjuntos de datos pequeños ($N = 90$). Se evaluaron características espectrales (HSV, YCbCr), espaciales (Canny Edges) y morfológicas (Hu, FFT).

En la primera fase del estudio, el preprocesamiento de reducción de dimensionalidad (reescalado a 256×256) limitó la precisión al 65 %, ocultando características estructurales clave; en este escenario, el análisis espectral dominó por completo, usando la morfología solo para casos especiales. Al utilizar las imágenes en su resolución original, se logró disparar el rendimiento a un 77 % de precisión.

Los resultados identificaron una dualidad crítica entre la Crominancia Roja (Cr) y la Densidad de Bordes en el canal de Saturación. Contrario a la tendencia actual de utilizar arquitecturas de deep learning y grandes volúmenes de datos, este estudio sugiere fuertemente que un modelo de machine learning clásico (árbol de decisión) alimentado con datos de fidelidad espacial adecuada puede ser suficiente en escenarios restringidos para tareas de discriminación robusta. Se observa un sobreajuste medio (Training: 90 %, Test: 77 %) derivado de la escasez de datos, pero el rendimiento valida que la fidelidad espacial es superior en importancia a la complejidad algorítmica. Esto asegura la viabilidad de sistemas de monitoreo de bajo costo cuyo rol sea el de asistentes de filtrado para operadores humanos, optimizando la asignación de recursos en la vigilancia forestal.

Index Terms—Computer Vision, Fire Detection, Edge Computing, Spatial Resolution, Parsimony, Feature Engineering.

I. INTRODUCCIÓN

La vigilancia de incendios forestales enfrenta un problema crítico de escalabilidad. Mientras que el costo de los sensores visuales ha disminuido, el estado del arte en visión artificial ha tendido hacia arquitecturas de deep learning con librerías como YOLO y utilización de recursos como las CNNs que demandan recursos computacionales significativos (GPUs) y alto consumo energético [1]. Esta dependencia tecnológica dificulta el despliegue masivo en zonas remotas y países subdesarrollados donde la energía y los recursos están limitados.

El desafío radica en que los modelos ligeros tradicionales suelen sufrir de altas tasas de falsos positivos al confundir

el fuego con objetos estáticos en un espectro similar (luces artificiales, atardeceres).

Se propone que el fuego posee una “firma de rugosidad” específica en su pureza espectral (saturación) que es detectable en resoluciones medias, sin necesidad de la potencia de cómputo de una red neuronal convolucional y los datos que esta necesita para mantenerse.

El objetivo es demostrar que un modelo simple, alimentado con datos de fidelidad espacial adecuada, puede funcionar como un sistema de alerta temprana eficiente, no con un enfoque de automatización de detección, sino con una arquitectura *human-in-the-loop* donde la IA no funge como un juez absoluto; en su lugar, cumple un rol de facilitador de los escenarios peligrosos para la validación de un ser humano, enfoque fundamentado puramente en que los incendios no descansan, y los seres humanos sí.

II. ESTADO DEL ARTE

La detección automática de incendios ha evolucionado significativamente desde los sistemas basados en sensores físicos (humo, calor) hacia sistemas de visión por computadora, capaces de proporcionar alertas más tempranas y mayor información contextual. La literatura actual se puede dividir en dos paradigmas principales: los métodos clásicos basados en características y el enfoque contemporáneo basado en aprendizaje profundo (Deep Learning).

II-A. Métodos Clásicos: Espacios de Color y Reglas Físicas

Históricamente, la detección de fuego por visión artificial se ha fundamentado en la extracción manual de características (*hand-crafted features*), aprovechando las propiedades físicas y espectrales de las llamas. Un trabajo seminal en esta área es el de Chen et al. [2], quienes propusieron uno de los primeros modelos robustos basados en el análisis cromático y dinámico. Su método utiliza el modelo de color RGB para identificar píxeles de fuego mediante reglas de decisión estáticas y dinámicas, combinadas con la detección de movimiento para diferenciar el fuego de objetos estáticos con colores similares.

Posteriormente, diversos autores expandieron este enfoque explorando espacios de color alternativos como HSV y YCbCr, argumentando que estos separan la luminancia de la crominancia de manera más efectiva que RGB, permitiendo aislar la componente de fuego incluso bajo variaciones de iluminación. Sin embargo, estos sistemas basados en reglas, aunque computacionalmente ligeros, tradicionalmente sufren de altas tasas de falsos positivos ante escenarios adversos como atardeceres, luces artificiales intensas o superficies reflectantes rojas [2].

II-B. El Paradigma del Deep Learning

En la última década, el campo ha sido dominado por las Redes Neuronales Convolucionales (CNNs). Como se detalla en la revisión reciente de Xu et al. [1], arquitecturas como YOLO (*You Only Look Once*), Faster R-CNN y VGG16 han establecido el estado del arte en términos de precisión. Estos modelos aprenden automáticamente jerarquías de características complejas, superando las limitaciones de las reglas manuales en escenarios visualmente ruidosos.

No obstante, este rendimiento superior conlleva un costo computacional prohibitivo para aplicaciones de borde (*Edge Computing*) e IoT. Además, la práctica estándar en Deep Learning de redimensionar imágenes a resoluciones bajas (e.g., 224×224) para ajustar la entrada de la red, sacrifica información de alta frecuencia crucial para detectar incendios incipientes o lejanos, un fenómeno que este estudio identifica como una pérdida crítica de “fidelidad espacial”.

III. METODOLOGÍA

Se adoptó una estrategia de evaluación incremental para construir un clasificador explicable y ligero.

III-A. Percepción Cromática

Se seleccionaron espacios de color eficientes computacionalmente:

1. **Espacio HSV:** La Saturación (S) se identificó como un predictor robusto frente a la intensidad (V).
2. **Espacio YCbCr:** El canal Cr (Diferencia Roja) actúa como el separador lineal primario, aislando candidatos cálidos con un costo computacional mínimo: “Si no es rojo no es fuego”.

III-B. Percepción Estructural

Se implementó un detector de bordes Canny modificado, aplicado sobre el canal de saturación.

- **Hipótesis:** El fuego presenta una varianza espacial alta en su saturación (vibración) que los objetos sólidos pintados de rojo no poseen.

III-C. Protocolo de Resolución

Se compararon dos escenarios para validar la importancia de la fidelidad espacial sobre la complejidad del modelo:

1. **Escenario A (Low Res - 256^2):** Simulación de entrada típica para redes neuronales profundas (CNNs).
2. **Escenario B (Medium Res - $\approx 600 \text{ px}$):** Resolución nativa de cámaras de videovigilancia estándar; las fotos en el dataset son de calidades medias y comunes.

IV. RESULTADOS EXPERIMENTALES

IV-A. Impacto de la Resolución

La utilización de la resolución media fue el factor determinante, superando las limitaciones observadas en baja resolución.

- Precisión Low Res: $\approx 65\%$.
- Precisión Medium Res: 77% .

La Fig. 1 ilustra cómo el *accuracy* mejora drásticamente al permitir que los algoritmos de bordes operen sobre la señal original, capturando la micro-textura del fuego sin requerir modelos de aprendizaje profundo.

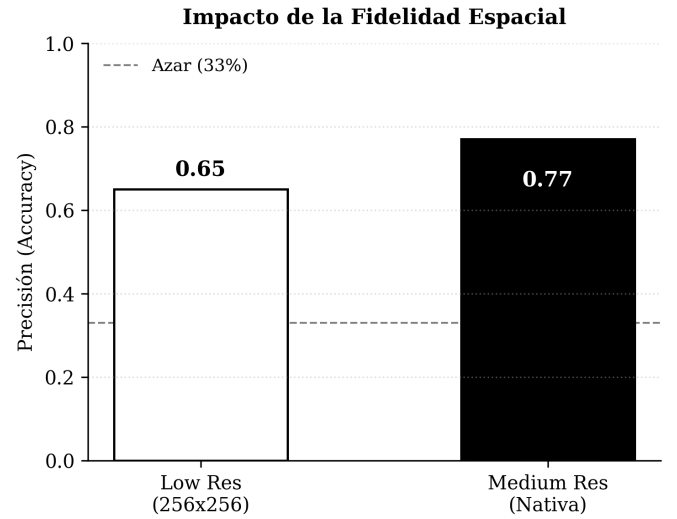


Figura 1. Comparación de Accuracy: Baja Resolución vs. Resolución Media. Se observa un incremento significativo al mantener la fidelidad espacial.

IV-B. Análisis de Overfitting

Con un dataset limitado ($N = 90$), se observa una divergencia en las curvas de aprendizaje (Training: 90% vs Test: 77%). Esta brecha es esperable y atribuible a la varianza de los datos, no a una falla del modelo. A pesar de esto, el rendimiento en prueba (77%) es significativamente superior al azar y competitivo para un sistema de primer filtrado, donde además la profundidad óptima para el modelo de árboles de decisión fue de 3. Esto demuestra que las situaciones de alerta de incendios pueden surgir a partir de los puntos más simples sin hacernos muchas preguntas.

IV-C. Importancia de Características

El análisis de importancia (Fig. 2) validó la jerarquía propuesta:

1. Crominancia Roja (Cr): Filtro primario.
2. Densidad de Bordes en Saturación: Discriminador estructural dominante.
3. FFT / Hu Moments: Importancia secundaria.

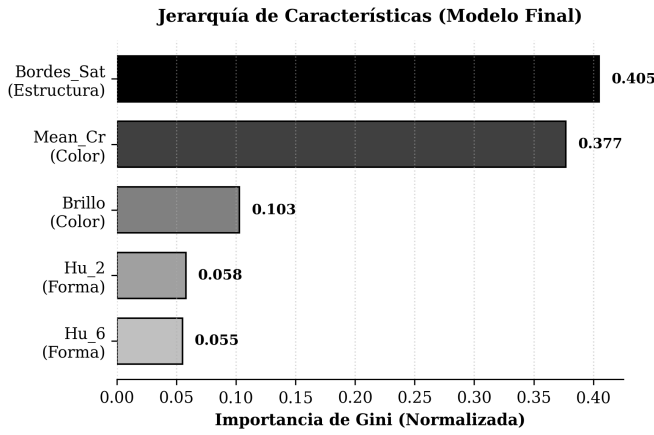


Figura 2. Gráfica de Importancia de Características (Feature Importance). La crominancia y la saturación dominan sobre las características morfológicas complejas.

V. DISCUSIÓN

V-A. ML + Feature Engineering vs Deep Learning

Este trabajo no busca reemplazar modelos de deep learning, sino explorar escenarios donde su uso no es viable desde una perspectiva operativa. La literatura ha demostrado de forma consistente que las arquitecturas de deep learning son altamente efectivas y deseables en contextos de automatización total; sin embargo, su aplicación suele estar condicionada por la disponibilidad de grandes volúmenes de datos y recursos computacionales significativos.

El enfoque basado en modelos de machine learning clásicos, acompañado de un proceso exhaustivo de ingeniería de características, responde a escenarios donde dichas condiciones no están presentes, como despliegues en edge computing, entornos energéticamente restringidos o regiones con acceso limitado a infraestructura tecnológica. En este contexto, se plantea una alternativa pragmática orientada a democratizar la detección temprana de incendios, un problema de impacto global que requiere soluciones escalables y accesibles.

V-B. Validación de la Resolución Operativa

Se demostró que la resolución media (600 px) es el punto óptimo. Provee suficiente detalle para caracterizar la textura del fuego sin incurrir en los costos de procesamiento de imágenes HD o 4K, y sin sufrir la degradación de señal de las imágenes pequeñas (224 px o 256 px) típicas de las entradas de CNNs como ResNet o YOLO. Esto es una victoria doble, porque las cámaras capaces de captar imágenes de alta resolución son costosas y desplegarlas de manera masiva en bosques sería ineficiente. Al no depender de imágenes de alta calidad, prescindimos del análisis computacional pesado, haciendo nuestro modelo más eficiente.

V-C. Viabilidad Comercial

El modelo resultante (Árbol de Decisión de profundidad limitada) requiere órdenes de magnitud menos memoria y ciclos de CPU que una red neuronal. Esto habilita su implementación

en hardware cuyo precio ronda <10 USD (e.g., ESP32-Cam), democratizando el acceso a sistemas de alerta temprana. De esta forma, no solo países altamente desarrollados tienen la oportunidad de cuidar sus bosques; la visión por computadora es un recurso que debe estar al alcance de todos.

VI. TRABAJO FUTURO

Si bien los resultados validan la eficacia de la fidelidad espacial en imágenes estáticas, la detección de incendios es intrínsecamente un problema dinámico. Las siguientes líneas de investigación se proponen para superar las limitaciones actuales:

1. **Análisis Temporal:** Incorporar videos de incendios forestales para medir el parpadeo característico de la llama. Esta dimensión temporal es ortogonal a las características espaciales y permitiría descartar casi por completo los objetos rojos estáticos. De la misma forma, se pueden incorporar fragmentos de video de follaje estacional senescente en movimiento para marcar la línea de diferencia.
2. **Expansión del Dataset:** El diferencial de accuracy (13%), que marca un pequeño sobreajuste en el modelo, indica la necesidad subyacente de aumentar la variedad de datos, incluyendo escenarios adversos (niebla, atardeceres, diversidad de ecosistemas y momentos del día) para robustecer las reglas de decisión.
3. **Implementación en Hardware:** Portar y ajustar el modelo final a un microcontrolador (e.g., ESP32) para validar su latencia y consumo energético en condiciones de campo reales.

VII. CONCLUSIONES

Este estudio desafía la noción de que mayor precisión requiere mayor complejidad computacional.

1. Se logró un 77% de precisión utilizando ingeniería de características clásica sobre imágenes de resolución media.
2. La “firma de rugosidad”(entendida de forma descriptiva como una alta variación espacial en el canal de saturación) en el canal de Saturación es un discriminador potente que elimina la necesidad de análisis espectrales complejos en muchas situaciones.
3. El sistema propuesto ofrece una alternativa viable y económica al paradigma de Deep Learning, ideal para aplicaciones IoT descentralizadas.
4. El modelo demuestra que hay alternativas al aprendizaje profundo para atacar problemas en sectores con menores recursos, manteniéndose simple, bajo la física del fuego y la lógica humana.

REFERENCIAS

- [1] R. Xu, H. Lin, K. Lu, L. Cao, and Y. Liu, “Recent advances in deep learning for fire detection: A review,” *IEEE Access*, vol. 11, pp. 22 075–22 091, 2023.
- [2] T. Chen, P. Wu, and Y. Chiou, “An early fire-detection method based on smoke and flame analysis,” *IEEE Transactions on Industrial Electronics*, vol. 51, no. 6, pp. 1573–1583, 2004.