DE LA NOCHE AL INFINITO 2.0: ¿NOS INVENTAMOS EL CIELO O LO REVELAMOS CON IA?

Imma Garrofé

Me gusta fotografiar el cielo, pero también intento entender qué ocurre cuando proceso las imágenes. Por eso he recuperado los datos que obtuve en 2024 y los he vuelto a procesar en 2026, con el objetivo de estudiar cómo han cambiado los resultados con las actualizaciones de las herramientas de inteligencia artificial. Además, he realizado un experimento: he añadido a una imagen de Namibia un gradiente y ruido similares a los de un cielo contaminado, para comprobar si el programa consigue limpiarlos sin alterar la información.

A finales de agosto apareció una nueva actualización de BlurXTerminator, así que también he procesado de nuevo algunas imágenes para evaluar sus efectos. Quería comprobar qué mejoras son reales, qué limitaciones tienen estas herramientas y hasta qué punto el procesado puede llegar a mostrar detalles que no estaban realmente en el cielo.

Tres cielos, el mismo equipo

En el trabajo original comparé imágenes tomadas desde tres lugares: Mollerussa, Àger y Namibia. Mollerussa, cerca de Lleida, presenta una contaminación lumínica elevada, equivalente a un cielo Bortle 5–6. Àger ofrece un cielo mucho más oscuro, de alrededor de Bortle 3, y Namibia, donde observé con compañeros de la asociación en el campamento Tivoli Astrofarm, proporcionó las mejores condiciones.

Utilicé el mismo equipo en las tres ubicaciones. Como tenía que transportarlo hasta Namibia, elegí un sistema modular Askar V de 270 milímetros de focal, al que llamo «el Lego» porque se monta y desmonta con facilidad. También utilicé dos filtros duales Askar Color Magic. De esta manera, pude comparar las imágenes manteniendo el equipo y la focal constantes.

La primera conclusión ya era clara en 2024: el cielo manda. Namibia, con poca contaminación lumínica, ofrecía las mejores imágenes; Mollerussa, las peores. Las imágenes de Àger quedaron por debajo de lo esperado porque durante aquellas noches hizo mucho calor y la transparencia atmosférica fue mala.



Figura 1. La misma imagen desde Mollerussa (izquierda), Àger (medio) y Namibia (derecha)

Qué entiendo por inteligencia artificial

La inteligencia artificial es el campo general que incluye el aprendizaje automático o *machine learning*. En este último, un modelo aprende patrones a partir de ejemplos; las redes neuronales son una de las técnicas que se utilizan para ello.

Por ejemplo, si detecto un punto débil en una imagen, ¿es una estrella tenue o es ruido? Podría intentar distinguirlo mediante reglas, teniendo en cuenta el brillo, la forma o la presencia del punto en varias tomas. Un modelo de aprendizaje automático, en cambio, aprende a combinar patrones a partir de ejemplos, y una red neuronal puede aprender qué características buscar en pequeños grupos de píxeles.

En este estudio me centro específicamente en las herramientas BlurXTerminator y NoiseXTerminator de PixInsight. Entre 2024 y 2026 cambió el modelo utilizado por BlurXTerminator, de AI4 a ML4. También hubo cambios en el modelo de reducción de ruido. Por tanto, no comparo métodos completamente distintos, sino versiones y modelos actualizados de las mismas herramientas.

Datos y método de trabajo

Partí de unas ocho horas de datos obtenidos en cada ubicación y trabajé con cuatro canales: RGB, hidrógeno alfa (Hα), oxígeno III (OIII) y azufre II (SII). Los canales Hα, OIII y SII proceden de separar la información de una imagen en color; no son imágenes monocromas independientes. Por eso comparé cada canal consigo mismo, en lugar de comparar directamente unos canales con otros.

El conjunto de datos dio lugar a doce combinaciones de ubicación y canal, que evalué mediante distintas mediciones. Es una limitación importante: se trata de doce combinaciones, no de doce repeticiones independientes del mismo experimento.

Realicé dos tipos de análisis: uno estadístico y otro visual. Para las mediciones utilicé PixInsight en fase lineal, antes de estirar el histograma, cuando los valores de los píxeles siguen siendo proporcionales a la señal registrada. Apliqué el mismo procedimiento que en 2024: eliminación automática del gradiente de fondo, calibración de color solo en RGB y, después, BlurXTerminator y NoiseXTerminator. La automatización del gradiente me permitió evitar una variable adicional que habría introducido diferencias si lo hubiera ajustado manualmente.

Para las comparaciones visuales procesé también las imágenes de banda estrecha, estiré el histograma y realicé los últimos retoques en Photoshop. Sin embargo, utilicé estas versiones únicamente para la valoración visual; las cifras proceden de las imágenes lineales.

Qué medí en las imágenes

En las estrellas estudié el número de detecciones, el ruido de fondo, la anchura estelar y la excentricidad, que indica hasta qué punto las estrellas son redondas. El número de detecciones debe interpretarse con cautela: que el programa identifique más puntos no demuestra por sí solo que todos correspondan a estrellas reales.

En la nebulosa y el fondo elegí dos rectángulos fijos: uno situado sobre la nebulosa y otro sobre el cielo de fondo. Con ellos medí el contraste, es decir, cuánto destaca la nebulosa respecto al fondo; la textura, que describe sus variaciones sobre ese fondo; y la rugosidad, que mide las variaciones en relación con la propia señal.

Para estas medidas utilicé la mediana en lugar de la media, porque una zona con muchas estrellas podría alterar notablemente el promedio. También utilicé la desviación absoluta mediana. Descarté la relación señal-ruido como indicador en este estudio porque los valores obtenidos eran prácticamente iguales y, por tanto, no aportaban información útil para esta comparación.

Comparación entre 2024 y 2026

Al comparar visualmente las versiones, el cambio se aprecia sobre todo en Mollerussa. En Àger también se observan diferencias, mientras que en Namibia apenas se notan: la calidad inicial del cielo ya era muy alta y el margen de mejora era menor.

En las gráficas distinguí la imagen original, el resultado del procesado de 2024 y el de 2026. En términos generales, el procesado mejoró las métricas respecto a los datos originales tanto en 2024 como en 2026. Sin embargo, las diferencias entre las dos versiones fueron pequeñas en la mayoría de los casos. Los cambios más claros aparecieron en algunos archivos que partían de más ruido.

El número de estrellas detectadas aumentó de forma notable en ciertos casos. En los datos de Àger, donde las condiciones atmosféricas habían sido desfavorables, el incremento fue de alrededor del 24 %; en Mollerussa llegó aproximadamente al 80 %. En cambio, el ruido de fondo, la anchura de las estrellas y su excentricidad variaron mucho menos.

En las mediciones de nebulosa también encontré mejoras en algunos canales. En hidrógeno alfa y azufre II de Mollerussa, el contraste y la textura aumentaron cerca de un 50 %, mientras que la rugosidad apenas varió, alrededor de un 0,1 %. En RGB de Àger, el contraste aumentó aproximadamente un 54 % respecto a 2024. Estos cambios se apreciaban también al examinar las imágenes, aunque no todos eran igual de evidentes a simple vista.

El experimento: contaminar una imagen de Namibia

Quise comprobar si el algoritmo podía eliminar la contaminación y el ruido de una imagen de buena calidad hasta dejarla equivalente a la original. Para ello tomé una imagen de Namibia y le añadí el gradiente de fondo de una imagen de Mollerussa. En una primera prueba, sumé el gradiente sin retirar antes el de Namibia y el resultado perdió un 88 % de las detecciones estelares. Comprendí que el procedimiento no era correcto y lo modifiqué.

En la segunda prueba retiré primero el gradiente de Namibia y añadí el de Mollerussa. Como el fondo resultante era demasiado suave y no reproducía las fluctuaciones de la captura real, incorporé ruido gaussiano hasta aproximar el nivel de ruido de Mollerussa, con una diferencia de alrededor del 0,1 % en la medición utilizada.

Era consciente de que se trataba de una simulación: no podía reproducir exactamente el *seeing*, las condiciones atmosféricas ni toda la estructura del ruido real. Mi intención no era crear una imagen idéntica a una toma auténtica de Mollerussa, sino realizar una prueba controlada para observar cómo respondía el algoritmo.



Figura 2. La misma imagen limpia y degradada

En el canal RGB, la imagen limpia tenía unas 36.000 detecciones. Al degradarla, el número bajó a unas 32.000, mientras que el ruido aumentó, como cabía esperar. Después del procesado con las herramientas de IA, las detecciones subieron a unas 44.000 en la imagen limpia y a unas 47.000 en la degradada. La diferencia de unas 3.300 detecciones en la imagen degradada no podía interpretarse como estrellas reales: yo había añadido ruido, no estrellas ni información nueva.

El comportamiento no fue idéntico en todos los canales. En algunos, el fondo quedó ligeramente desplazado. Las mediciones de contraste y textura resultaron poco útiles para evaluar esta prueba, porque dependían del ruido que yo mismo había añadido. Al fijarme en la rugosidad, la imagen contaminada se mantenía en general bastante cerca de la limpia, aunque no era idéntica píxel a píxel.

La conclusión de este experimento es que el programa puede reducir de forma eficaz el ruido de una imagen degradada, pero también puede detectar como estrellas algunos elementos que en realidad proceden del ruido. El resultado visual puede parecer mejor sin que eso signifique que se haya recuperado información astronómica auténtica.

Qué aporta la actualización más reciente

A finales de agosto probé una nueva actualización de BlurXTerminator con los dos archivos de mejor calidad y los dos de peor calidad. La anchura de las estrellas disminuyó aproximadamente un 15 %, de modo que parecían más puntiagudas, pero no se hicieron más redondas. Además, el ruido de fondo aumentó.

Esto muestra que una actualización no mejora todos los aspectos a la vez. Puede afinar las estrellas, pero hacerlo a costa de incrementar el ruido, dependiendo de la imagen. Tampoco

puedo afirmar que la reducción de la anchura implique una mayor resolución real: una imagen puede parecer más definida en la pantalla sin que eso demuestre que contenga más información.

Conclusiones: ¿inventamos el cielo o lo revelamos?

Mis datos indican que las herramientas de IA pueden mejorar determinadas métricas y que el efecto de las actualizaciones suele ser más evidente en los archivos que parten de más ruido. Sin embargo, los resultados dependen de cada imagen y ninguna versión mejora todos los parámetros al mismo tiempo.

En el experimento de contaminación aparecieron unas 3.300 detecciones adicionales que no podían corresponder a estrellas reales, porque no había añadido estrellas ni información, sino ruido. Es posible que el algoritmo confundiera parte de ese ruido con señales estelares. Por eso, una mejora numérica necesita siempre una interpretación y una comprobación cuidadosas.

Me gustaría contrastar en el futuro las detecciones con datos de Gaia, para comprobar cuáles corresponden a estrellas catalogadas. No he podido hacerlo todavía, y el número reducido de combinaciones estudiadas también limita las conclusiones.

Así que vuelvo a la pregunta del título: ¿nos inventamos el cielo o lo revelamos? La IA puede ayudarnos mucho a extraer y presentar detalles, pero no garantiza por sí sola que todo lo que aparece en una imagen sea real. Si el procesado no elimina correctamente el ruido, pueden surgir estructuras que no estaban en el cielo. Y, sobre todo, ningún algoritmo sustituye a un buen cielo: por mucho que procese una imagen de Mollerussa, no puedo convertir sus condiciones en las de Namibia o Àger. Debemos seguir protegiendo la calidad del cielo nocturno.

Preguntas

¿El ruido gaussiano añadido reproduce el ruido real de una imagen astronómica?

No exactamente. Consideré utilizar ruido de Poisson, pero en mis pruebas se disparaba en las zonas brillantes. Elegí el gaussiano porque me permitía aproximarme mejor al nivel de ruido de Mollerussa. Aun así, el ruido sintético no puede reproducir el *seeing*, las condiciones atmosféricas ni toda la estructura de una captura real. Es una simulación útil para experimentar, pero no demuestra que ambos ruidos sean iguales.

¿Por qué pueden aparecer detecciones que no corresponden a estrellas?

Los modelos de aprendizaje automático se entrenan con muchos ejemplos y aprenden a distinguir patrones que se parecen a estrellas o a ruido. En una imagen con mucho ruido, el algoritmo puede confundir algunas fluctuaciones con señales estelares. Por eso, detectar más puntos no equivale necesariamente a recuperar más estrellas reales.

¿Procesarías de esta manera tus propias imágenes?

No necesariamente. Para mejorar una imagen concreta probaría diferentes ajustes y valoraría qué funciona mejor en cada caso. En este trabajo, en cambio, repetí el mismo procedimiento en todas las imágenes porque quería comparar estadísticamente los resultados, aunque ese enfoque no sea el que utilizaría para procesar una fotografía individual.