

Resumen

La energía solar es la fuente de energía renovable más disponible. España, como el séptimo país del mundo en términos de generación de energía fotovoltaica, instala una gran cantidad de nuevos paneles solares cada año. Por lo tanto, es necesario utilizar técnicas de teledetección para monitorear la distribución y el crecimiento de los paneles solares fotovoltaicos. Mediante la aplicación de redes neuronales convolucionales se creará un modelo clasificador que identifique las placas solares en ortofotografías de 25cm de resolución espacial. Para generar un conjunto de muestras de entrenamiento y evaluación se digitalizarán manualmente todas las instalaciones de placas solares en el término municipal de L'Eliana (Valencia). Mediante análisis espacial se determinarán los fragmentos de las ortofotografías en los que existan placas solares. A partir de este conjunto de muestras se compararán diversos parámetros que influyen en la calidad del modelo clasificador para obtener el modelo óptimo. Se evaluará el modelo obtenido analizando sus fortalezas y debilidades. Dicho modelo se podrá aplicar en otros municipios y/o en años distintos al analizado para conocer la evolución de las instalaciones de placas solares.

Introducción

La capacidad de generación de energía solar fotovoltaica (PV) ha crecido un 41% anual desde 2009. Se proyecta que la capacidad de generación de energía solar fotovoltaica aumentará casi diez veces para 2040, apoyando sistemas energéticos que combaten el cambio climático y promueven el acceso universal a la energía. España, como séptimo país del mundo en generación de energía fotovoltaica, añade una gran cantidad de nuevos paneles solares cada año. Por ello, es crucial emplear técnicas de teledetección para monitorear su distribución y crecimiento. La mayoría de los estudios sobre la distribución de paneles solares utilizan imágenes satelitales (como Sentinel y Landsat) con aprendizaje profundo o automático para identificarlos globalmente [5][6]. Sin embargo, debido a la baja resolución espacial de estas imágenes (10 metros para Sentinel-2 y 30 metros para Landsat-8), se detectan principalmente grandes áreas solares en campos y áreas agrícolas. No son adecuadas para detectar pequeños paneles solares domésticos en techos de edificios. En España, el servicio PNOA proporciona ortofotografías de alta resolución actualizadas regularmente, con una resolución espacial mínima de 25 cm, ideal para identificar paneles solares domésticos en áreas urbanas. El ICV en la Comunidad Valenciana ofrece ortofotografías similares, actualizadas anualmente. En este trabajo de fin de máster (TFM), se utilizarán ortofotografías del ICV para identificar paneles solares en áreas urbanas mediante dos modelos de redes neuronales convolucionales (CNN). Se realizará una vectorización manual de paneles solares en L'Eliana (Valencia) para entrenar y evaluar los modelos. Se ajustarán y analizarán las muestras de entrenamiento repetidamente para obtener el mejor modelo, evaluando sus ventajas y desventajas. Se intentará aplicar el modelo a ortofotografías de otras áreas y años distintos a los de entrenamiento.

Metodología

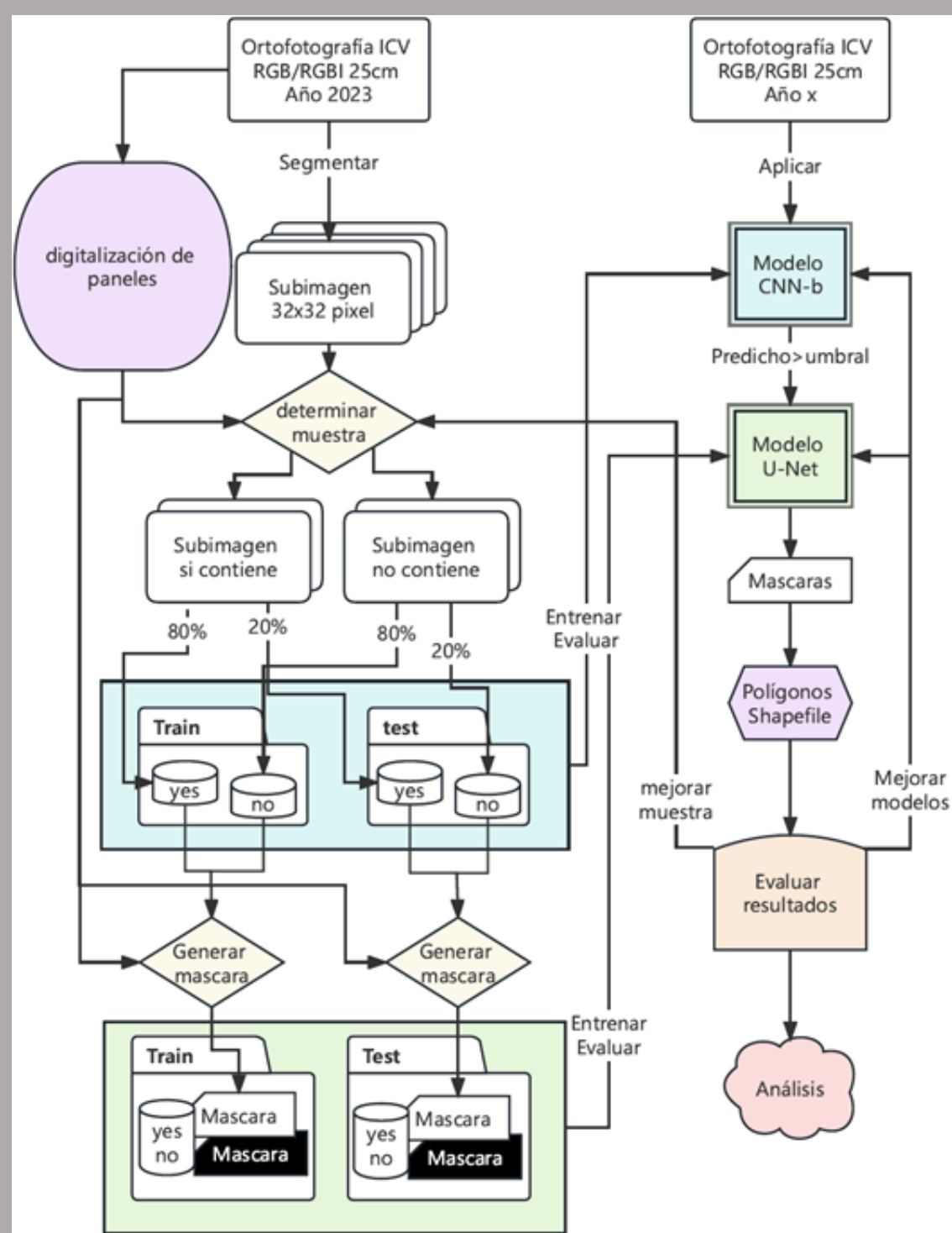


Diagrama de flujo de los procesos

A través del diagrama de flujo, se puede comprender que primero utilizamos las ortofotografías de 25 cm de resolución espacial de 2023 para digitalizar manualmente el área de estudio y dividirla en varias subimágenes. Luego, determinamos los conjuntos de entrenamiento y test para los modelos CNN-b y U-Net (también necesita generar máscaras para modelo U-Net). Después de entrenar los modelos, utilizamos estos dos modelos para aplicar el modelo CNN-b que hace un primer filtrado y luego el modelo U-Net que genera una máscara con la zona de las placas solares a las imágenes ortográficas de 25 cm de resolución espacial de cualquier año. Finalmente, generamos archivos shapefile a partir de las máscaras creadas por el modelo para el análisis de datos final. Según los resultados de estos datos, también optimizamos continuamente los dos modelos.

A partir del diagrama de flujo del modelo, podemos observar que primero determinamos la estructura básica de los modelos CNN-b y U-Net a través de una serie de intentos. Luego, evaluamos los resultados y realizamos modificaciones continuas y reentrenamientos del modelo. Al entrenar el modelo CNN-b, probamos dos enfoques: entrenar desde cero y mejorar un modelo preentrenado. Para ambos enfoques, probamos si usar o no el aumento de datos. Finalmente, descubrimos que el modelo entrenado desde cero con aumento de datos producía los mejores resultados. Luego, continuamos optimizando el modelo CNN-b modificando los parámetros de aumento de datos y eliminando algunas capas. Para el modelo U-Net, probamos el uso del aumento de datos y encontramos que la precisión era elevada. Luego, consideramos que la inclusión de muestras negativas (las muestras que no contiene las zonas de las placas solares) durante el entrenamiento debería tener un efecto positivo en el modelo, aunque no pudimos confirmarlo solo con el conjunto de validación. Por lo tanto, después de aplicar el modelo, realizamos ajustes y optimizaciones continuas basándonos en los resultados finales. En este proceso, añadimos la banda infrarroja y manualmente agregamos algunas muestras negativas que eran propensas a errores de identificación. Finalmente, descubrimos que el modelo U-Net entrenado con la inclusión de muestras negativas producía mejores resultados. Mientras ajustábamos las muestras negativas, también reentrenamos y optimizamos la capacidad de identificación del modelo CNN-b.

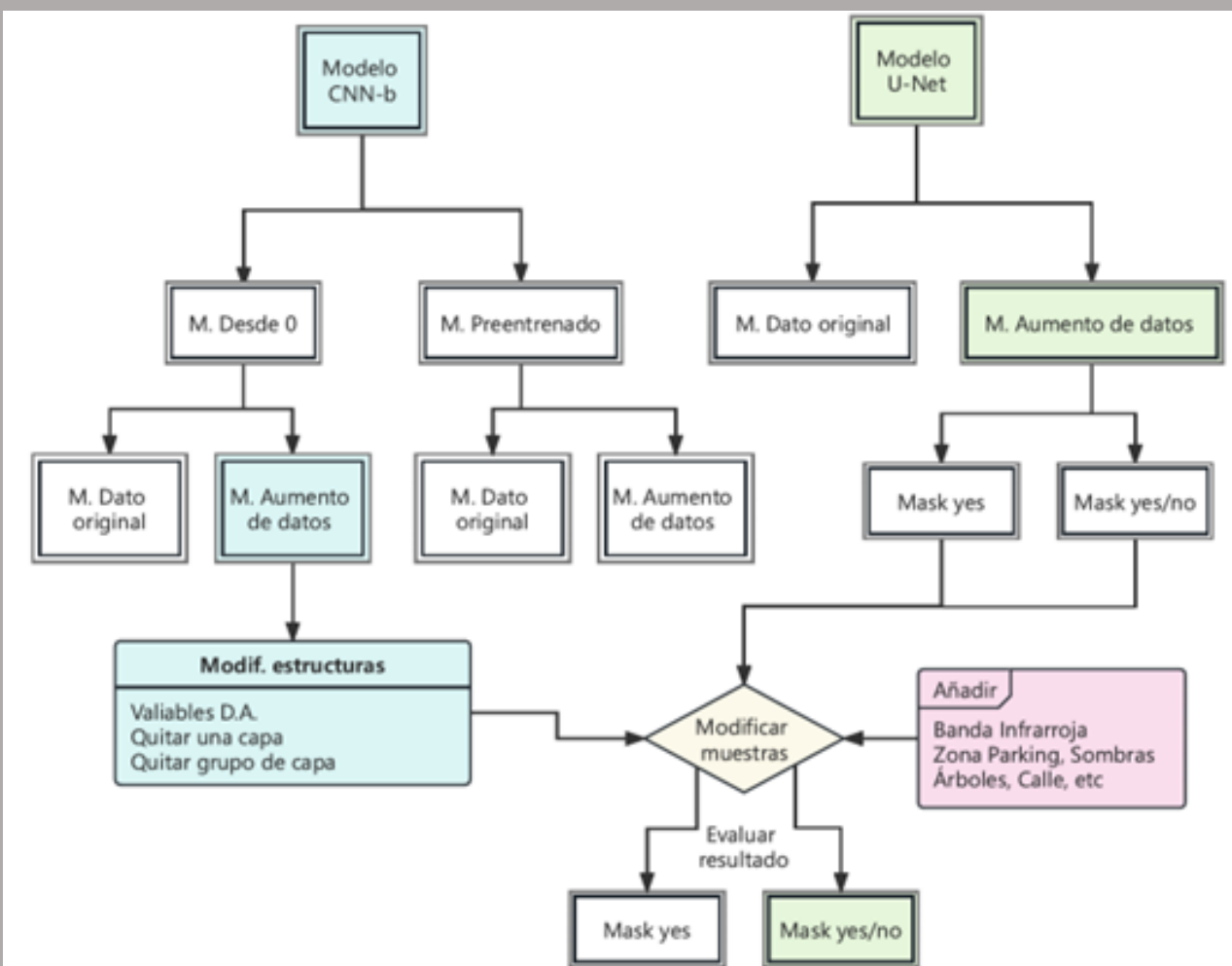


Diagrama de flujo de modelo

Resultados

En 2023, el código identificó 686 paneles solares existentes y 1709 nuevos, aunque puede haber errores debido a falsos positivos. La verificación manual mostró que la mayoría de los paneles solares de 2023 fueron correctamente identificados, con una precisión del 95.07% en los paneles existentes. Sin embargo, solo 261 de los 406 nuevos paneles solares detectados fueron correctos, resultando en una precisión del 64.28%. La mayoría de los errores se debieron a sombras de edificios y piscinas, difíciles de distinguir con una resolución de 25 cm. Estos resultados indican una baja tasa de falsos negativos pero una tasa relativamente alta de falsos positivos en el modelo, capaz de identificar la mayoría de los paneles solares pero con algunos errores en la clasificación. Estimamos que entre 2022 y 2023 se instalaron aproximadamente 1098 nuevos paneles solares en L'Eliana, con una tasa de crecimiento del 160.1%.



Conclusiones

En este trabajo, utilizamos las ortofotografías proporcionadas por el ICV para entrenar y aplicar un modelo CNN binario y un modelo U-Net. A través de ajustes y evaluaciones repetidas, finalmente obtuvimos los modelos más adecuados. Combinando ambos modelos, identificamos pequeños paneles solares en el área de L'Eliana, en la provincia de Valencia, y evaluamos su crecimiento. Nuestros dos modelos mostraron una precisión muy alta en la identificación de muestras negativas. Sin embargo, aún existe margen de mejora en la identificación de muestras positivas. Los modelos lograron identificar casi todos los paneles solares, pero también marcaron incorrectamente algunas áreas de sombra como paneles solares. Según el análisis realizado con nuestros modelos para los años 2022 y 2023 en L'Eliana, se instalaron 1098 nuevos paneles solares en 2023, lo que representa un aumento del 160.1% respecto a 2022. El uso del modelo U-Net en investigaciones relacionadas con la ciencia de la teledetección también es muy viable. U-Net puede generar eficazmente máscaras de las áreas necesarias para un análisis posterior, lo que ofrece una ventaja significativa en comparación con los modelos CNN tradicionales que solo pueden predecir por imagen. En conclusión, creemos que el uso del aprendizaje profundo para identificar objetivos pequeños es muy factible, pero aún presenta muchos desafíos. En particular, es difícil para el modelo distinguir entre áreas de sombra y áreas correctas. También hemos propuesto algunas posibles soluciones de optimización para futuras investigaciones.

Bibliografía

- [1] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.
- [2] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. Journal of Machine Learning Research, 15(1), 1929–1958.
- [3] Kruitwagen, L., Story, K. T., Friedrich, J., Byers, L., Skillman, S., & Hepburn, C. (2021). A global inventory of photovoltaic solar energy generating units. Nature, 598(7882), 604–610