

DIGITALIZACIÓN DEL MAPA DE LA PISCINA DE COMBUSTIBLE GASTADO MEDIANTE TÉCNICAS DE APRENDIZAJE AUTOMÁTICO



ÁNGEL RAMOS
Ingeniero
ENUSA



DAVID VERDEJO
Ingeniero
ENUSA



ÁNGEL RODRÍGUEZ
Ingeniero
ENUSA



ROBERTO RODRÍGUEZ
Ingeniero
ENUSA

INTRODUCCIÓN

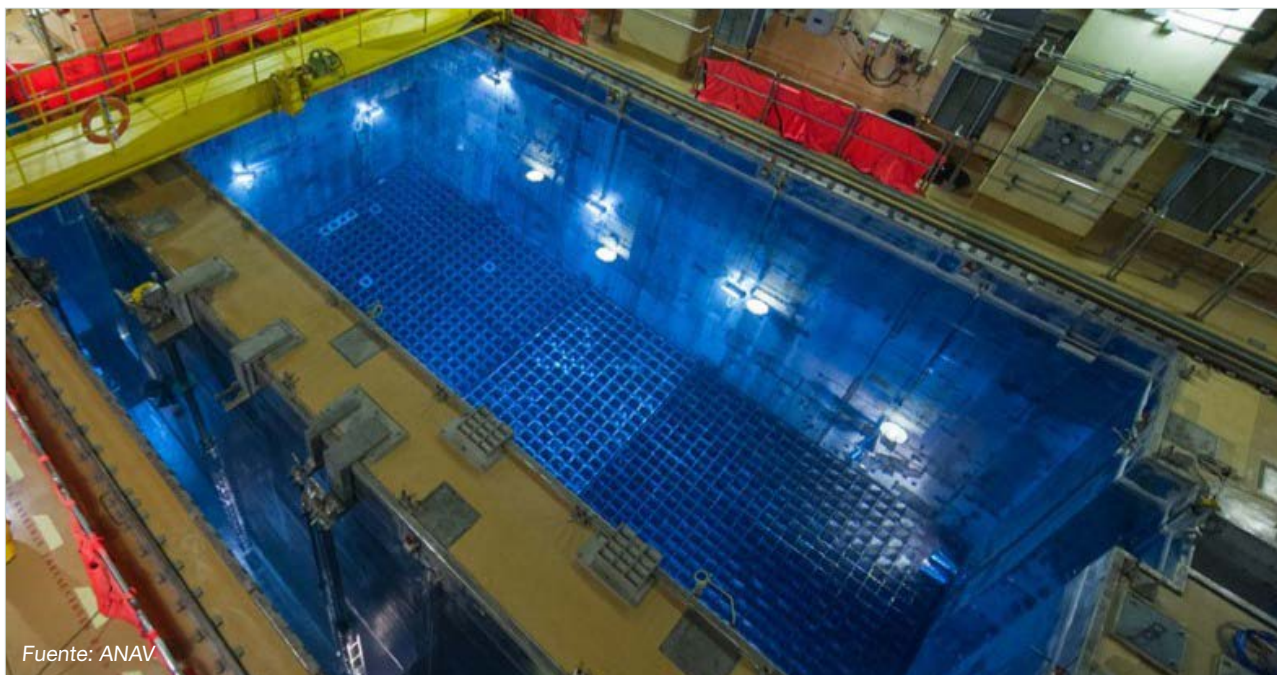
En los últimos años, la industria nuclear ha experimentado un aumento significativo en la aplicación de tecnologías de aprendizaje automático o *machine learning* (ML) e inteligencia artificial (IA) para mejorar la seguridad, la eficiencia y la fiabilidad en diversos aspectos de las operaciones nucleares [1, 2, 3]. Estos avances han facilitado la optimización del diseño de centrales nucleares, el mantenimiento predictivo, la monitorización de reactores y la evaluación de riesgos, entre otras aplicaciones [4, 5, 6]. En este panorama en rápida evolución, la industria nuclear española también ha reconocido la utilidad de estas tecnologías y está desarrollando soluciones innovadoras de ML e IA para diversos problemas y tareas en el sector nuclear [7].

Dentro de esta nueva línea de transformación digital, uno de los proyectos que ENUSA ha desarrollado recientemente es la automatización de la verificación del mapa de la piscina de combustible gastado PWR, existente en cada central nuclear, con el objetivo de simplificarla y reducir, en la medida de lo posible, potenciales errores y el tiempo de procesamiento en su realización. Aquí se presentan los fundamentos y requerimientos de esta tarea, y las soluciones de aprendizaje automático empleadas para llevarla a cabo.

MAPA DE LA PISCINA DE COMBUSTIBLE GASTADO

Tras cada recarga del reactor de una central nuclear, con la cual se inicia un nuevo ciclo de producción de energía, es necesario hacer una verificación del mapa de la piscina de combustible gastado (PCG) de la planta. Esta tarea forma parte de las instrucciones técnicas impuestas por el Regulador de Seguridad Nuclear a cada central nuclear. Su objetivo es mantener un inventario controlado de todo el combustible gastado y los residuos radiactivos de alta actividad que se almacenan en la piscina. Por consiguiente, ENUSA, tras la recarga y en un plazo máximo de seis meses, realiza una verificación del mapa de la PCG que conserva la planta en sus registros, mediante una inspección por vídeo del contenido de cada una de las celdas de los bastidores que conforman la PCG. Esta labor la lleva a cabo en la mayoría de las plantas PWR españolas.

En el proceso de elaboración del mapa de verificación de la PCG, un equipo técnico especializado se divide en dos grupos. Un grupo, desde el puente grúa de la piscina, maneja una cámara de vídeo subacuática fijada al extremo de una pértiga para visualizar cada una de las celdas de la piscina. La cámara se coloca justo encima de la celda, enfocando el conjunto de combustible, o los componentes que contiene, de forma que



Fuente: ANAV

sus códigos de identificación sean legibles. El resto del equipo se sitúa al borde de la piscina, con un inspector controlando el enfoque de la cámara y la iluminación, mientras va diciendo en voz alta, en alfabeto radiofónico, las coordenadas de la celda que se está visualizando, y los códigos de identificación de los componentes que se puedan ver en la imagen. Otros inspectores cercanos transcriben los códigos identificados a un archivo y verifican que el componente encontrado en cada celda coincide con lo esperado según el mapa de la piscina que la planta proporciona al inicio de la inspección, de acuerdo con sus registros. La señal de video y audio se graba en archivos, uno por cada fila de celdas en la piscina, para tener un registro de todo el proceso.

Además de la verificación continua de los códigos identificados que se hace durante la inspección en planta, personal técnico revisa el archivo generado, cotejándolo con los videos grabados, para confirmar que el contenido de cada celda se haya transcrito correctamente, y corregir, si es el caso, datos erróneos. A lo largo de este proceso el contenido visualizado se contrasta con el mapa de piscina de referencia que proporciona la planta, realizando así su verificación. Todos estos pasos de verificación buscan eliminar cualquier error humano en la transcripción de los códigos durante la inspección, ya que se trata de una tarea repetitiva que requiere atención constante y dura varios días. De esta forma se prepara el informe final de verificación del mapa de la PCG con todas las garantías de fiabilidad.

Dados los múltiples pasos de verificación que implica este procedimiento, se propuso simplificarlo al máximo, mediante la aplicación de técnicas de ML, para ahorrar tiempo sin comprometer la fiabilidad de la verificación. La idea era incorporar herramientas específicas al proceso que analizaran la imagen registrada por la cámara y, simultáneamente, procesaran el audio con la locución del inspector, para finalmente transcribir de forma automática los códigos a un mapa en formato Excel.

TÉCNICAS DE MACHINE-LEARNING EN LA ELABORACIÓN DEL MAPA

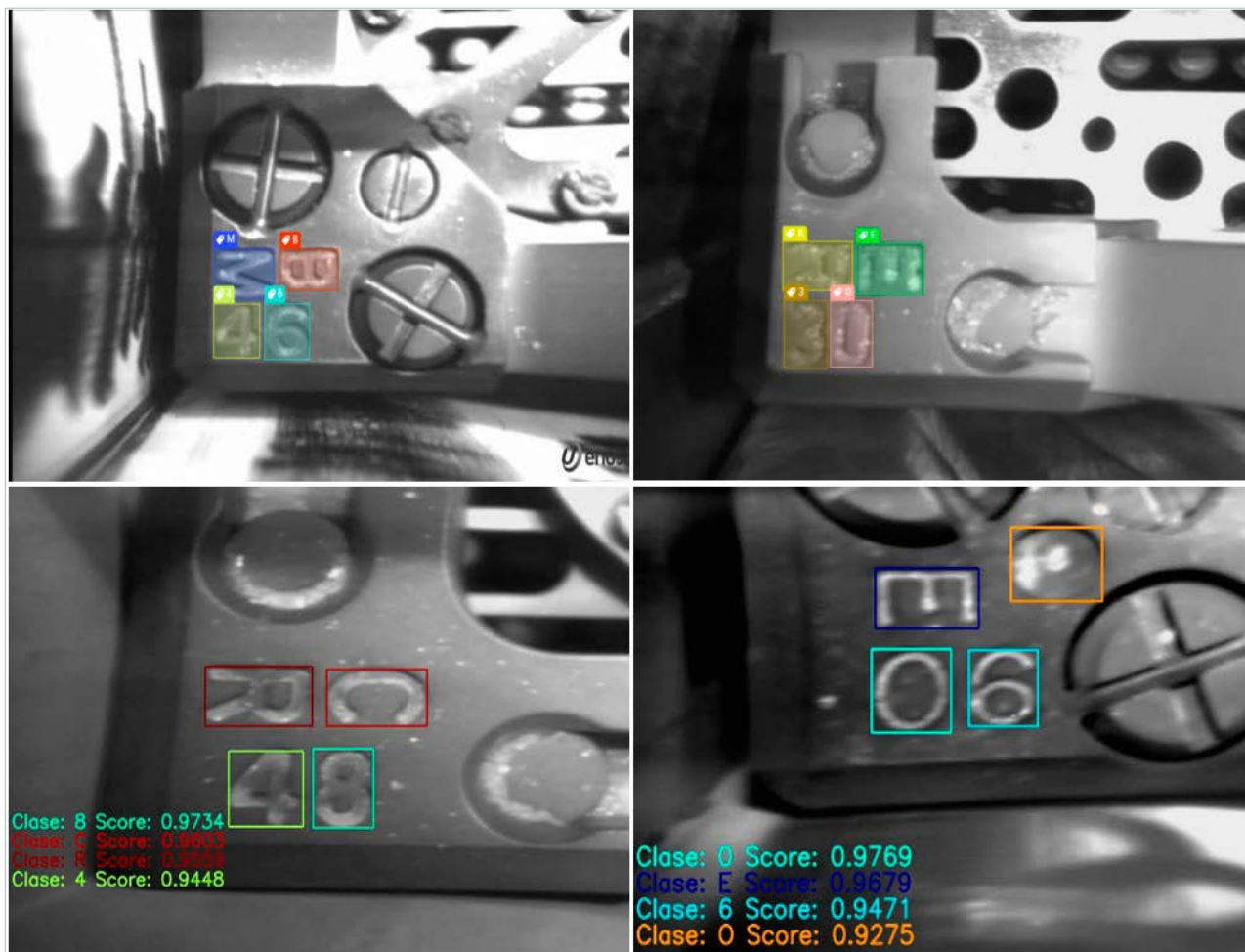
Las técnicas de ML que pueden utilizarse para automatizar la elaboración del mapa de la PCG se encuentran en dos áreas principales, con un amplio desarrollo industrial, incluyendo el ámbito nuclear: la visión artificial y el procesamiento secuencial de datos.

Las técnicas de ML de visión artificial son directamente aplicables al procesamiento de la señal de video grabada durante la inspección de la piscina. Estas técnicas se aplican a diversos problemas en la industria nuclear [8], como la detección de grietas y defectos en componentes de reactores o la inspección visual del combustible. Por otro lado, el reconocimiento y transcripción del audio grabado durante la elaboración del mapa se enmarcan en una de las principales aplicaciones de las técnicas de procesamiento secuencial de datos, concretamente el reconocimiento de voz. Este campo ha experimentado un gran progreso en los últimos años y cuenta con múltiples herramientas de libre acceso, con una alta fiabilidad en la transcripción de voz de diferentes hablantes y en distintos idiomas.

IDENTIFICACIÓN DE CÓDIGOS EN EL VÍDEO

Los códigos de identificación de los elementos combustible constan de tres o cuatro caracteres alfanuméricos, cada uno de los cuales se ha considerado como un objeto independiente que hay que detectar e identificar en la imagen de vídeo.

Para la detección de objetos en una imagen, las redes neuronales profundas (DNN en inglés), en concreto las que incorporan redes convolucionales [9, 10], es una de las técnicas de ML con mayor potencial, ya que proporciona un análisis de la imagen más eficiente que las técnicas de procesamiento tradicionales, al permitir el aprendizaje automático de las características visuales más representativas de cada problema particular. Las principales ventajas de un sistema basado en DNN son una



mayor fiabilidad en la detección y clasificación de objetos o patrones en una imagen, y una menor sensibilidad a las variaciones en sus condiciones de registro, como la luminosidad o el enfoque.

Para el problema actual, se eligió la DNN preentrenada YOLO [11]. Esta red está especialmente diseñada para la detección y localización de objetos en imágenes, y ha sido optimizada con cientos de miles de imágenes de objetos comunes etiquetados. Con este entrenamiento previo, que es computacionalmente muy costoso, se llega a una red con decenas de millones de parámetros optimizados. Una red así preentrenada permite abordar más fácilmente un problema particular como el aquí descrito. Para ello, la red se somete a un reentrenamiento que parte de los parámetros de la red preentrenados, utilizando imágenes con los objetos a detectar (como los códigos de los elementos combustibles) que han sido adecuadamente etiquetados. De esta manera no se requieren tantas imágenes de entrada ni tantos pasos de entrenamiento como en el primero, porque los parámetros iniciales de la red ya están más cerca de sus valores óptimos que si se entrenase desde cero.

Se extrajeron unas 2.400 imágenes de grabaciones procedentes de campañas del mapa de distintas piscinas de centrales españolas PWR, en las que aparecen códigos de identificación de elementos combustible. Aproximadamente

9000 caracteres de estos códigos fueron etiquetados, tarea consistente en clasificarlo y delimitar cada carácter con un recuadro que indica su posición dentro de la imagen. Estos datos se emplearon para el reentrenamiento de la red YOLO, y así ajustarla a este problema concreto. Se ha comprobado que el modelo obtenido tiene una gran precisión en la identificación y posicionamiento de caracteres en imágenes no empleadas en su entrenamiento, y es de gran generalidad, siendo válido para códigos de distintas piscinas. Este procedimiento permite mejorar la precisión del modelo con sucesivos entrenamientos realizados con un conjunto mayor de imágenes etiquetadas.

Hasta el momento, el modelo identifica únicamente los códigos de los elementos combustible, pues son los más numerosos. Para otros componentes que también tienen un código, como tapones o barras de control, es más laborioso obtener un amplio conjunto de datos para reentrenar la red al ser menos numerosos, y se ha decidido posponer su automatización a una etapa posterior.

TRANSCRIPCIÓN Y EXTRACCIÓN DE INFORMACIÓN DEL AUDIO

Los modelos de reconocimiento automático del habla (ASR en inglés) se utilizan para la tarea de transcribir el habla en texto. La precisión de estos modelos ha aumentado significativamente desde que se adoptó un modelo híbrido basado en DNN hace una década [12]. Más recientemente,

los modelos ASR han mejorado su rendimiento utilizando la arquitectura Transformer [13], que puede aprender características de datos sin etiquetar. De esta manera, se han desarrollado modelos ASR multilingües que pueden gestionar varios idiomas con la misma red, mediante un preentrenamiento con gran cantidad de texto sin etiquetar de varios idiomas, y posteriormente con un ajuste fino para un idioma y tarea específicos.

En 2022, OpenAI lanzó los modelos Whisper [14], que son modelos Transformer entrenados con datos débilmente supervisados, pero en mucha mayor cantidad que sistemas anteriores. Esta enorme cantidad de datos les permite trabajar eficazmente con conjuntos de datos existentes y nuevos, eliminando la necesidad de ajustes específicos y obteniendo resultados precisos. Los modelos Whisper, disponibles en diferentes tamaños, están entrenados para tareas de reconocimiento y transcripción de voz, y son muy robustos frente a variaciones de acentos y ruido de fondo. Aprovechando todos estos avances, los modelos Whisper se han elegido para el procesamiento de audio en la automatización del mapa, empleando herramientas desarrolladas para aplicarlos en transcribir la locución en tiempo real.

Para extraer información de la transcripción de audio (coordenadas de las celdas y códigos), se localizan palabras clave pronunciadas por el inspector, las cuales indican el momento en que se pronuncian los diferentes códigos. Asimismo, el uso del alfabeto radiofónico facilita la identificación de caracteres. Al contrario que en el procesamiento de vídeo, a partir de la transcripción de voz es posible la identificación de otros códigos distintos a los de los elementos combustibles, estando implementado de esta forma para los códigos de tapones.

RESULTADOS

Para el procesamiento del vídeo se ha logrado un modelo general de la red que obtiene precisiones (identificación correcta del código presente en la imagen) superiores al 95%. Hay que destacar la generalidad del modelo logrado, porque permite usarlo en inspecciones de distintas PCG, siendo siempre posible mejorar su eficiencia introduciendo más imágenes nuevas en el conjunto de reentrenamiento. El resultado de la predicción incluye tanto el carácter detectado como su ubicación dentro de la imagen, enmarcándolo con gran precisión, tanto en imágenes de diferente luminosidad como en las que hay distintos entornos. Conocer la ubicación de los caracteres dentro de la imagen es fundamental para recombinarlos y determinar el código completo.

En el procesamiento de la señal de vídeo en directo se extraen y analizan fotogramas varias veces por segundo, con tiempos de procesamiento del orden de decenas de milisegundos. El análisis de la señal en tiempo real presenta además la ventaja de que, para cada elemento combustible visualizado, se contrastan los resultados obtenidos por varios fotogramas que incluyen el mismo código, proporcionando finalmente mayor eficacia que la estimación del código realizada con una sola imagen.

En cuanto al procesamiento del audio, las pruebas realizadas con grabaciones de campañas de verificaciones del mapa

han demostrado que los modelos Whisper ofrecen una eficiencia de transcripción muy alta, y que no se requiere entrenamiento para este problema. Además, el mismo modelo es aplicable para transcribir locuciones de diferentes inspectores con un rendimiento similar. No obstante, el modo de transcripción en tiempo real es necesario en esta tarea, ya que el análisis del habla debe realizarse simultáneamente con la señal de vídeo. Esto significa que la señal de audio debe procesarse al mismo tiempo que se graba, y las transcripciones resultantes deben entregarse con una latencia mínima, a la vez que la señal de vídeo muestra el código.

Se han analizado diferentes tamaños del modelo Whisper, que varían en el número de parámetros entrenados, encontrándose que el tamaño del modelo influye en la precisión de la transcripción y en el tiempo de ejecución, por lo que es necesario encontrar un equilibrio entre ambos. Finalmente, se eligió una versión intermedia de Whisper, cuya latencia en la transcripción es de entre uno y dos segundos, que es adecuada para esta tarea. La precisión de la transcripción obtenida es lo suficientemente alta como para extraer la información necesaria para determinar con exactitud los códigos y las coordenadas pronunciadas por el operador.

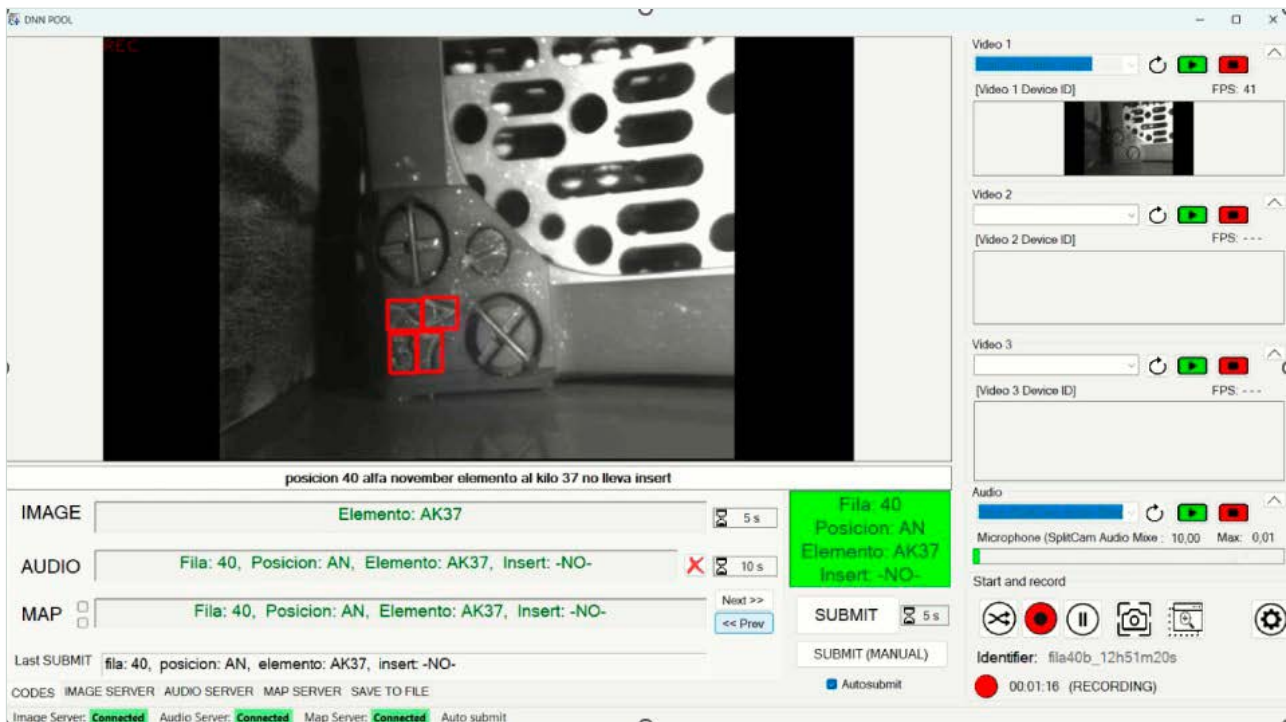
INTEGRACIÓN

La aplicación completa está formada por varias aplicaciones individuales que trabajan de manera conjunta para poder desarrollar cada una en el entorno y lenguaje más eficiente en cada caso y obtener el máximo rendimiento. Con esto se ha conseguido el reto de implementar las herramientas de ML de forma que procesen la información en tiempo real, sin aumentar los tiempos de la inspección.

Se ha desarrollado una aplicación de control, denominada DnnPool, y tres servidores de procesamiento (audio, vídeo y mapa). DnnPool es una aplicación de escritorio que se encarga de la gestión de los dispositivos de vídeo (hasta 3 canales) y un dispositivo de audio; la grabación en un archivo multimedia del canal visualizado en cada momento y el audio; y el envío de fotogramas del vídeo y fragmentos de audio a los respectivos servidores. La gestión de los servidores (lanzamiento, conexión, comunicación, comandos intermedios) se realiza también a través de la aplicación DnnPool para facilitar la tarea al operador. Los servidores son aplicaciones desarrolladas en Python por ser el lenguaje y entorno más adecuado para las herramientas de ML.

El análisis independiente del audio y el vídeo sirve como autoverificación al comparar los resultados de ambos procesos, los cuales reflejan la información que proporciona el registro de la cámara y el inspector durante el proceso. Así, la aplicación de control compara constantemente los resultados de audio, vídeo y el resultado esperado según el mapa proporcionado por la planta. Cuando todos los datos coinciden, se le indica al servidor de mapa que lo grabe en el archivo de mapa de resultados.

El desarrollo de esta aplicación ha incorporado mejoras adicionales al proceso, como facilitar y simplificar el montaje necesario para la elaboración del mapa, unificando todo en un solo ordenador y reduciendo el número de dispositivos necesarios.



CONCLUSIONES

Se han aplicado técnicas de aprendizaje automático para automatizar la tarea de verificación del mapa de la piscina de combustible gastado de centrales nucleares. La identificación automática de los códigos de los elementos combustible, en el canal de vídeo y en el de audio, se logra con una tasa de éxito de alrededor del 95 %, utilizando diferentes tipos de redes neuronales, en un procesamiento en tiempo real, con tiempos de ejecución lo suficientemente cortos como para permitir su integración en el proceso de elaboración de los mapas de piscina.

Se ha desarrollado una aplicación que integra estas herramientas y facilita el proceso en múltiples aspectos a los operarios que lo realizan in situ. El sistema completo ya se está empleando por ENUSA en las campañas de elaboración del mapa de las PCG, la primera vez de forma completa en enero de 2026, con buena aceptación y resultados. Se ha conseguido una mejora del proceso completo al hacerlo más ágil y sencillo, permitiendo ahorrar tiempo en las sucesivas verificaciones que requiere este trabajo.

REFERENCIAS

- [1] M. Benedict, and M. S. El-Genk, "Advances in AI and ML for the nuclear industry", Nuclear Engineering and Design, vol. 340, pp. 115–126, 2018.
- [2] M. S. El-Genk, and J.-M. P. Tournier, "Machine learning for enhanced safety and efficiency in nuclear power plants", Progress in Nuclear Energy, vol. 111, pp. 392–404, 2019.
- [3] Y. Zhang, R. Li, and Z. Xie, "A review of artificial intelligence in nuclear engineering", Journal of Nuclear Science and Technology, vol. 57(10), pp. 1177–1200, 2020.
- [4] K. G. Metzroth, A. M. Ibrahim, and M. S. El-Genk, "Machine learning applications in nuclear reactor design and analysis", Annals of Nuclear Energy, vol. 118, pp. 29–44, 2018.
- [5] H. A. Gohel, H. Upadhyay, L. Lagos, K. Cooper, and A. Sanzeteenea, "Predictive maintenance architecture development for nuclear infrastructure using machine learning", Nuclear Engineering and Technology, vol. 52(7), pp. 1436–1442, 2020.
- [6] L. Chen, Z. Wang, and G. Sun, "Real-time monitoring and fault diagnosis for nuclear power plants using artificial intelligence: A survey", IEEE Access, vol. 8, pp. 37045–37059, 2020.
- [7] A. Ramos, A. Carrasco, J. Fontanet, L. E. Herranz, D. Ramos, M. Díaz, J. M. Zazo, O. Cabellos, and J. Moraleda, "Artificial intelligence and machine learning applications in the Spanish nuclear field", Nuclear Engineering and Design, vol. 417, 112842, 2024.
- [8] C. Tang, C. Yu, Y. Gao, J. Chen, J. Yang, J. Lang, C. Liu, L. Zhong, Z. He, and J. Lv, "Deep Learning in Nuclear Industry: A Survey", Big Data Mining and Analytics, vol. 5(2), pp. 140–160, 2022. <https://doi.org/10.26599/BDMA.2021.9020027>.
- [9] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning", Nature, vol. 521, pp. 436–444, 2015.
- [10] W. Rawat, y Z. Wang, "Deep convolutional neural networks for image classification: A comprehensive review", Neural Comput., vol. 29(9), pp. 2352–2449, 2017.
- [11] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, pp. 779–788. <https://>

doi.org/10.1109/CVPR.2016.91

[12] G. Hinton, L. Deng, D. Yu, G. Dahl, A. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. Sainath, and B. Kingsbury, "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups", IEEE Signal Processing Magazine, 29, pp. 82–97. 2012. <https://doi.org/10.1109/msp.2012.2205597>.

[13] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, and I. Polosukhin, "Attention is all you need". Advances in neural information processing systems, 30, 2017.

[14] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, "Robust Speech Recognition via Large-Scale Weak Supervision", 2022, arXiv:2212.04356v1. ■