

Vea discusiones, estadísticas y perfiles de autores para esta publicación en: <https://www.researchgate.net/publication/351382166>

Evaluación de modelos basados en inteligencia artificial para clasificar células fotovoltaicas defectuosas

Artículo en Ciencias Aplicadas · Mayo 2021

DOI: 10.3390 / app11094226

CITACIONES

0

LEE

32

6 autores, incluyendo:



Víctor Alonso Gómez

Universidad de Valladolid

48 PUBLICACIONES 336 CITACIONES

[VER EL PERFIL](#)



Luis Hernández-Callejo

Universidad de Valladolid

106 PUBLICACIONES 1,777 CITACIONES

[VER EL PERFIL](#)

Algunos de los autores de esta publicación también están trabajando en estos proyectos relacionados:





Microrred [Ver Proyecto](#)



Sistema de gestión energética de edificios no residenciales basado en modelos de predicción de demanda para el desplazamiento de cargas [Ver Proyecto](#)

Artículo

Evaluación de modelos basados en inteligencia artificial para clasificar células fotovoltaicas defectuosas

Álvaro Pérez-Romero¹, Víctor Felipe Mateo-Romero², Sara Gallardo-Saavedra³, Víctor Alonso-Gómez³, María del Carmen Alonso-García⁴  y Luis Hernández-Callejo^{3,*} 

- ¹ Departamento de Matemáticas, Estadística e Informática, Universidad de Cantabria, Av. de los Castros, s/n, 39005 Santander, España; alvaro.perezr@alumnos.unican.es
 - ² Departamento de Inteligencia Artificial, Universidad Politécnica de Madrid, Pº Juan XXIII 11, 28031 Madrid, España; h.mateo@alumnos.upm.es
 - ³ Departamento de Ingeniería Agrónoma y Departamento de Física Forestal y Aplicada, Universidad de Valladolid, Campus Universitario Duques de Soria, 42004 Soria, España; sara.gallardo@uva.es (SG-S.); victor.alonso.gomez@uva.es (VA-G.)
 - ⁴ Unidad de Energía Solar Fotovoltaica, Centro de Investigaciones Energéticas, Departamento de Energía, Medioambientales y Tecnológicas (CIEMAT), 28040 Madrid, España; carmen.alonso@ciemat.es
- * Correspondencia: luis.hernandez.callejo@uva.es; Tel.: +34-975129418

Abstracto: Fotovoltaica solar (PV) la energía ha experimentado un importante crecimiento y perspectiva durante la última década debido al constante desarrollo de la tecnología y su alta confiabilidad, junto con una drástica reducción de costos. Este hecho ha favorecido tanto su implementación a gran escala como la Generación Distribuida a pequeña escala (DG). Los sistemas integrados en los sistemas de distribución local se consideran una de las claves para un futuro entorno construido sostenible en Smart Cities (CAROLINA DEL SUR). Es necesario el funcionamiento y mantenimiento avanzados (O&M) de las plantas solares fotovoltaicas. Los datos potentes y precisos generalmente se obtienen in situ mediante curvas de corriente-voltaje (IV) o electroluminiscencia (EL) imágenes, con nuevos equipos y metodologías propuestos recientemente. En este trabajo, los autores presentan una comparación entre cinco AI-modelos basados en clasificar PV células solares según su estado, utilizando EL imágenes en el PV nivel de celda solar, mientras que las curvas de celda IV se utilizan en la fase de entrenamiento para poder clasificar las celdas en función de su eficiencia de producción. Esta clasificación automática de células defectuosas facilita enormemente la identificación de defectos para PV operadores de planta, disminuyendo la mano de obra humana y optimizando la ubicación del defecto. Además, este trabajo presenta una metodología para la selección de variables importantes para el entrenamiento de un clasificador de células defectuosas.

Palabras clave: defecto de la célula fotovoltaica; clasificador; inteligencia artificial



Citación: Pérez-Romero, Á.; Mateo-Romero, HF; Gallardo-Saavedra, S.; Alonso-Gómez, V.; Alonso-García, MdC; Hernández-Callejo, L. Evaluación de artificiales Modelos basados en inteligencia para clasificar células fotovoltaicas defectuosas. *Appl. Sci.* **2021**, *11*, 4226. <https://doi.org/10.3390/app11094226>

Editor académico: Francesco Calise

Recibido: 15 de abril de 2021

Aceptado: 30 de abril de 2021

Publicado: 6 de mayo de 2021

Nota del editor: MDPI se mantiene neutral con respecto a los reclamos jurisdiccionales en mapas publicados y afiliaciones institucionales.



Derechos de autor: © 2021 por los autores. Licenciario MDPI, Basilea, Suiza. Este artículo es un artículo de acceso abierto distribuido bajo los términos y condiciones de la licencia Creative Commons Attribution (CC BY) (<https://creativecommons.org/licenses/by/4.0/>).

1. Introducción

Durante la última década, la instalación a nivel mundial de plantas de generación renovable ha aumentado considerablemente. Entre las renovables, la fotovoltaica (PV) Las plantas solares han sido las más interesantes de los últimos años, y parece que serán las más instaladas en los próximos años [1,2]. Durante 2019, último año analizado en el Informe de situación global [3], Se instalaron 201 GW de capacidad de energía renovable en el mundo; 115 GW de energía solar PV capacidad, correspondiente a más del 57% del total de adiciones renovables. La capacidad instalada acumulada de energía solar fotovoltaica se elevó a 633,7 GW a finales de 2019.

La razón del espectacular crecimiento y la perspectiva de esta fuente de energía radica en el constante desarrollo de la tecnología y su alta confiabilidad. Esto ha hecho posible una drástica reducción de costos, lo que ha favorecido tanto su implementación a gran escala como la Generación Distribuida a pequeña escala (DG). Muchos países ya han comenzado a revisar sus políticas climáticas y energéticas. La innovación en el suministro de energía sostenible es, por tanto, fundamental para proporcionar fuentes de energía limpias y fiables y mejorar la calidad de vida en este planeta. Para lograr este objetivo, se ha lanzado la idea de edificios energéticos inteligentes o edificios energéticamente neutros. El principal objetivo de una regulación energética de un edificio es mantener

confort térmico interno, así como minimizar el consumo de energía [4]. *PV* sistemas y vehículos eléctricos (*Vehículos eléctricos*) integrados en los sistemas de distribución locales se consideran dos de las claves para un futuro entorno construido sostenible en Smart Cities (*CAROLINA DEL SUR*) [5]. El término "ciudad de la energía inteligente" ha surgido en paralelo con estos desarrollos, en particular desde al menos el cambio de la década de 2010 en relación con los componentes relevantes para la energía de las ciudades inteligentes [6]. En este escenario, asegurar la producción de energía es un factor clave para garantizar la rentabilidad de la planta, y esto ha obligado al diseño de operaciones y mantenimiento cada vez más inteligentes y avanzados (*O&M*) estrategias.

los *O&M* de *PV* Las plantas solares son fundamentales para los actores de la industria. El desarrollo de nuevos equipos y metodologías para su aplicación en *PV* plantas solares son necesarias y, en este sentido, la investigación y la industria en esta área están evolucionando rápidamente [7]. Algunos investigadores han trabajado con datos reales de *PV* plantas solares, y han obtenido importantes resultados mostrando defectos en las mismas [8]. los *PV* El inversor es un elemento crítico en *PV* plantas solares, pero también lo es *PV* modulo solar. *APV* El módulo solar se compone de *PV* células solares, y estas células pueden tener diferentes problemas o defectos, como se muestra en [9]. Aunque el conocimiento en el *PV* El nivel de celda solar es interesante para los operadores y mantenedores de *PV* plantas solares, las medidas deben tomarse en el *PV* modulo solar.

La detección de fallas en *PV* Los módulos solares están sujetos a la obtención de datos de campo, ya que actualmente la monitorización no alcanza el nivel de módulo en *PV* plantas. Se pueden obtener datos potentes y precisos mediante curvas de corriente-voltaje (*IV*), tanto de la *PV* celda solar (cuando se puede acceder a los contactos de la celda) o en el *PV* nivel del módulo solar [9,10]. Hoy en día, los trazadores intravenosos presentan algunos inconvenientes, como ser solo para el nivel de cadena, funcionar fuera de línea o ser costosos. Sin embargo, los instrumentos avanzados desarrollados recientemente permiten automatizar la adquisición de curvas del módulo *IV* en línea sin requerir la *PV* planta que se cerrará [11]. Además, también es posible realizar curvas *IV* oscuras de módulos en línea en *PV* plantas sin necesidad de desconectarlas del hilo. Para ello, una combinación de placas electrónicas en el *PV* Se requieren módulos y un inversor bidireccional, como se presenta en [12].

Otra forma de obtener datos de campo es tomando imágenes. Clásicamente, la imagen termográfica (*IRT*) se ha utilizado para la detección temprana de puntos calientes [13], y en los últimos años este *IRTE* El método se ha llevado a cabo mediante vuelo con drones [14,15]. Una de las técnicas de mantenimiento más prometedoras es el estudio de imágenes de electroluminiscencia (*EL*) como complemento de *IRT* análisis. Sin embargo, su elevado coste ha impedido su uso regular hasta la fecha. Los autores en [dieciséis] propuso una metodología de mantenimiento para realizar inspecciones *EL* in situ de la forma más eficiente posible utilizando un inversor bidireccional.

Todas estas técnicas de inspección permiten localizar e identificar los defectos presentes en *PV* módulos. Sin embargo, hoy en día las plantas solares son cada vez más grandes y el tratamiento manual de todos los datos obtenidos mediante las técnicas anteriores presentadas puede resultar una tarea muy costosa y que requiere mucho tiempo. Inteligencia artificial (*AI*) ya se está aplicando en *PV* plantas solares. *AI* La aplicación se ha centrado durante mucho tiempo en problemas de predicción de la producción de energía. Los autores en [17] desarrolló una solución que proporciona producción de electricidad basada en datos históricos y actuales de radiación solar disponibles en tiempo real. Otros autores han presentado un estudio de taxonomía, mostrando un proceso para dividir y clasificar los diferentes métodos de pronóstico, y los autores también presentaron las tendencias en *AI* aplicado a la previsión de generación en energía solar *PV* plantas [18]. El uso de redes neuronales artificiales (*ANA*) ha tenido éxito en la última década; algunos autores usan *ANN* junto con variables climáticas para pronosticar la generación en *PV* plantas solares [19], mientras que otros usan Support Vector Machine (*SVM*) junto con una optimización de los parámetros internos del modelo [20].

ANN también se han utilizado para otras tareas, como la detección de problemas en la producción de energía, como es el caso de [21], donde los autores utilizan la función de base radial (*RBF*) para detectar este tipo de fallas en la producción. Se busca un objetivo similar en [22], donde un *SVM*-Modelo basado en la descripción de un método de diagnóstico de fallas que utiliza la relación lineal entre la radiación solar y los gráficos de generación de energía. Esta investigación estudia los siguientes tipos de fallas: fallas del inversor, errores de comunicación, fallas del sensor, errores de la caja de conexiones e incendios de la caja de conexiones. El modelo clasifica la cadena y el inversor.

fracasos. Sin embargo, en realidad *PV* plantas, cada inversor puede cubrir miles de módulos y, por lo tanto, se puede perder información importante sobre fallas durante la clasificación.

Por tanto, es posible afirmar que el uso de *AI* es común en *PV* plantas solares. La investigación ha estudiado su aplicación en cuestiones de predicción de la producción de energía o para la detección de problemas en la producción de energía. Sin embargo, se ha destacado cómo la detección de defectos utilizando información a nivel de inversor puede ser imprecisa. En este trabajo, los autores presentan una comparación entre cinco *AI*-modelos basados en clasificar *PV* células solares según su estado en base a imágenes EL. Los cinco modelos bien conocidos utilizados para la clasificación han sido: *k* vecinos más cercanos (*KNN*), *SVM*, Bosques aleatorios (*RF*), Perceptrón multicapa (*MLP*) y redes neuronales convolucionales (*CNN*). Esta clasificación automática de células defectuosas facilita enormemente la identificación de defectos de forma precisa para *PV* operadores de planta, disminuyendo la mano de obra humana y optimizando la ubicación del defecto. Para ello, los autores utilizaron un ad-hoc *PV* módulo solar fabricado de forma especial desde *PV* las células solares tienen sus contactos posteriores accesibles, lo que permite su caracterización total [9]. Con este módulo fabricado, fue posible obtener cada curva IV de celda, además de las imágenes EL, por lo que fue factible etiquetar cada celda (grupo 1: buena, grupo 2: regular y grupo 3: mala) en función de su producción. eficiencia. Esto permitió una clasificación precisa para el entrenamiento de modelos. El estudio presenta un método novedoso para el etiquetado de células basado en su eficiencia de producción, y esto fue posible gracias a la personalización *PV* módulo solar, que diferencia claramente esta investigación. Las clasificaciones discutidas en este documento son una extensión del trabajo anterior "*Clasificador de defectos de células fotovoltaicas: una comparación de modelos*" Presentado en el III Congreso Iberoamericano de Ciudades Inteligentes de Smart Cities (ICSC-CITIES 2020) [23]. El documento está estructurado de la siguiente manera: Sección 2 presenta los materiales y la metodología utilizada, Sección 3 muestra los resultados, y la sección 4 contiene conclusiones y propuestas de trabajo futuro.

2. Materiales y métodos

Esta sección tiene como objetivo explicar los materiales utilizados, así como la metodología seguida para validar el clasificador.

2.1. Materiales

Se utilizó un módulo policristalino de 60 celdas compuesto por celdas con y sin defectos. Las vistas frontal y posterior del módulo se presentan en la Figura 1a, b respectivamente. El módulo se fabricó ad-hoc con todas las celdas accesibles desde la parte posterior del módulo. En cuanto al etiquetado de la celda, se han utilizado números del 0 al 59 para identificar la celda, como se detalla en la Figura 2. Además, las cuatro celdas de las esquinas se han etiquetado tanto en la Figura 1a, b para facilitar la comprensión.

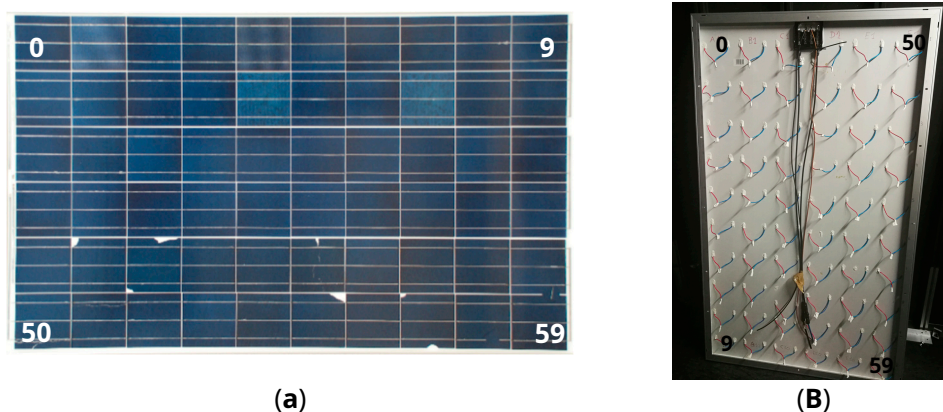


Figura 1. (a) Vista frontal del módulo. (b) Vista posterior del módulo.

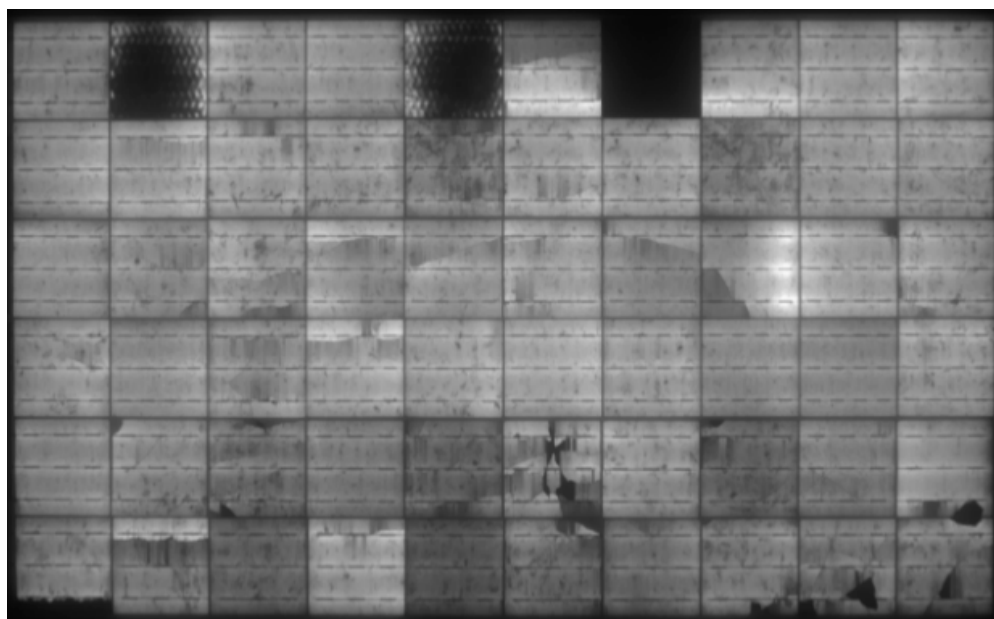


Figura 2. Imagen EL y numeración de células para el entrenamiento de modelos.

La primera cadena (primera y segunda columnas en la vista posterior) contiene defectos de fabricación, la cadena central (tercera y cuarta columnas) contiene fallas de soldadura, mientras que la tercera cadena (quinta y sexta columnas en la vista posterior) contiene deficiencias de rotura. Los defectos de baja eficiencia (celdas 1 y 4) y celdas de eficiencia media (celdas 14 y 17) se debieron a problemas de fabricación, con una eficiencia del 9% y 16,4% aproximadamente, pero sí se correspondieron con celdas rotas o en cortocircuito. La celda de cortocircuito (celda 6) se ha generado extendiendo las lengüetas de conexión de la celda más allá de la ubicación normal, provocando un cortocircuito en la celda. Para simular los defectos de soldadura defectuosa, los buses de la parte posterior de algunas celdas se dejaron sin soldar, ya sea un bus (celda 22) o dos buses (celda 34). No se han dejado tres buses sin soldar en ningún caso ya que hubiera significado que esta celda no estaría conectada en serie como el resto. En la celda 38, todas las pestañas se perdieron (sin soldar), aunque hicieron contacto permitiendo la producción del módulo. La celda con solo 1 cm soldada (celda 27) se usó para simular una mala soldadura, en la que solo se soldaba 1 cm del bus en lugar de los típicos 15 cm que se soldaban. La tercera cadena contiene algunas celdas agrietadas sin disminución del área de celda (celda 50 y 51), con disminución del área de celda (celda 41, 42, 55 y 57) o una combinación de ambas en la misma celda (celda 45). Cuando se coloca un trozo de celda rota encima de otra celda (celdas 49, 58 y 59), se genera un sombreado parcial, simulando el aspecto importante de los excrementos permanentes de aves. Este tipo de defectos se analizaron tal y como aparecen habitualmente en módulos comerciales en funcionamiento, ya sea durante la fabricación, el transporte o la operación. Sin embargo, los módulos comerciales no son accesibles a nivel de celda. Por eso se fabricó un módulo ad hoc para la caracterización de defectos.

Las características nominales de un módulo estándar de este tipo son: potencia nominal (P) 250 W, eficiencia 15,35%, corriente máxima en el punto de potencia (I_{mpp}) 8,45 A, tensión máxima en el punto de potencia (V_{mpp}) 29,53 V, corriente de cortocircuito (I_{sc}) 8,91 A y voltaje en circuito abierto (V_{oc}) 37,6 V. Teniendo 60 celdas en serie, se consideraron los valores nominales de una celda sana: potencia nominal 4,17 W, I_{mpp} 8,45 A, V_{mpp} 0,49 V, I_{sc} 8,91 A y V_{oc} 0,63 V.

La medición de la curva IV (a nivel de celda) y las imágenes EL se realizaron en las instalaciones del CIEMAT (Madrid, España). En resumen, las instalaciones utilizadas fueron las siguientes:

- Las mediciones en interiores se han realizado en el sistema comercial Pasan SunSim 3 CM, que consistió en un simulador solar de pulsos de luz clase AAA según norma IEC 60904-9, que puede realizar mediciones de curva IV en condiciones de Prueba Estándar;

- los *EL* e interior *IRT* Las pruebas se realizaron simultáneamente con el *EL* en esta cámara. El módulo se alimentó con una fuente de alimentación Delta SM 70-22. Un multímetro Fluke 189 conectado a los terminales del módulo que le permite registrar el voltaje exacto del módulo. *EL* y *IR* Las imágenes se capturaron con una cámara PCO 1300 y una cámara FLIR SC 640, respectivamente.

De esta forma, la información del *EL* imagen y curva IV de cada *PV* Se obtuvo la celda solar. En la siguiente figura 2, un *EL* Se muestra la imagen del módulo de medición y la numeración de las celdas.

Esta información podría servir para validar los diferentes modelos. En la fase de entrenamiento, las células utilizadas fueron etiquetadas según su potencia (medida a través de la curva IV) y con los siguientes criterios:

- Grupo 1: Poder $\geq 95\%$;
- Grupo 2: $80\% \leq \text{Poder} < 95\%$;
- Grupo 3: Potencia $< 80\%$.

Según las curvas IV medidas, la clasificación de las celdas en los tres grupos propuestos fue la siguiente:

- Grupo 1: 0, 2, 3, 5, 7, 8, 9, 10, 11, 12, 13, 15, 16, 18, 19, 20, 21, 22, 23, 24, 25, 28, 30, 31, 33 y 53;
- Grupo 2: 14, 17, 26, 27, 29, 32, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 46, 47, 48, 51, 52, 54, 55, 56, 57 y 58;
- Grupo 3: 1, 4, 6, 45, 49, 50 y 59.

Para la fase de prueba, los clasificadores solo utilizaron el *EL* imágenes, que eran perfectamente aplicables en el campo en común *PV* módulos. Las curvas IV solo se utilizaron en el entrenamiento del modelo para clasificar las células en función de su eficiencia de producción, y esto fue posible gracias a las *PV* módulo solar, que diferencia claramente esta investigación de trabajos anteriores.

2.2. Métodos

En esta sección se presenta la metodología seguida para la clasificación de células defectuosas. En primer lugar, el pretratamiento del *EL* se explicaron las imágenes. En segundo lugar, los mapas autoorganizados (*SOM*) se utilizaron para observar las similitudes detectadas entre células y grupos. En tercer lugar, se propusieron diferentes variables y se aplicó el método combinando *KNN* y *SVM*, se seleccionaron las variables con mejor desempeño. Finalmente, se desarrollaron y probaron cuatro clasificadores usando esas variables: *KNN*, *SVM*, *RF*, y *MLP*. Además, un quinto clasificador independiente, utilizando *CNN*, fue propuesto. Información más explícita sobre los modelos conocidos utilizados (*KNN*, *SVM*, *RF*, *MLP*, y *CNN*) se puede encontrar en las siguientes referencias destacadas [24–28].

Según el pretratamiento de *EL* imágenes para alimentar los clasificadores, ya que el *EL* La imagen se tomó de todo el módulo, como se muestra en la Figura 2, lo primero fue recortar las celdas para la evaluación individual. Para esto, un $115 \times$ Se estableció una ventana de recorte de 115 píxeles, que ofrecía el mejor ajuste superponiéndola en cada celda. Era necesario hacer zoom al máximo para ajustar la ventana de forma óptima, consiguiendo así que todas las imágenes recortadas estuvieran alineadas y evitando así incluir los márgenes negros que separan (por la estructura natural del panel) las celdas. Posteriormente, con el objetivo de mejorar los bordes, las imágenes se cargaron en Python, reduciendo cada lado en 1 píxel, dando finalmente como resultado 60 imágenes de tamaño 113×113 .

Dado el método de clasificación propuesto en este trabajo, en el cual las celdas fueron agrupadas según la potencia medida individualmente en cada una (grupos 1, 2 o 3, como se detalló anteriormente), no fue posible apreciar rasgos visibles (ver Figura 2) que caracterizó los elementos de los diferentes grupos, a excepción de los puntos negros del grupo 3.

Para demostrar la exactitud de esta clasificación, los autores propusieron utilizar *SOM* para poder observar las similitudes detectadas entre células. *SOM* es un tipo de red neuronal no supervisada, y su propósito es reducir la dimensionalidad de los datos preservando sus propiedades topológicas [29]. Se compone de solo 2 capas, la entrada y la

capas de salida. La capa de entrada tiene un número *norte* de neuronas equivalente a la dimensión de los datos. La capa de salida representa una matriz bidimensional de neuronas, cada una con un *norte*-dimensión de peso. Cada vez que el *SOM* se ejecuta el algoritmo del mapa, los pesos se inicializan aleatoriamente (una vez inicializados, se organizan y distribuyen en el mapa en función de su proximidad) y compiten entre sí cada vez que se ingresan datos en la capa de entrada. La neurona cuyo peso se asemeja más a la información de entrada es la ganadora, lo que provoca una actualización en el valor de su peso, así como en el de sus vecinas. Por lo tanto, el algoritmo consiste en iterar suficientes veces y elegir cada vez datos aleatorios en la muestra de entrenamiento para actualizar progresivamente el mapa hasta que se amolde a la estructura de la información inicial. En este estudio, un *SOM* se propuso una red de 20 filas y 20 columnas, utilizando la distancia euclidiana como distancia estándar, y los resultados se muestran en la sección de resultados.

Por tanto, era necesario decidir qué variables eran las representativas de los 60 miembros. Cada imagen (en escala de grises) era una matriz de 113 filas por 113 columnas, donde cada coordenada o píxel tenía valores de 0 (negro) a 255 (blanco). Dado que cada matriz tenía 12,769 coordenadas o píxeles que toman 256 valores posibles, entonces cada imagen también podría verse como una matriz aplanada, es decir, un vector $C_i = (p_{i,0}, \dots, p_{i,12768}) \in \mathbb{R}^{12769}$, $0 \leq i \leq 59$, donde $p_{i,0}, \dots, p_{i,112}$ son los valores de los píxeles que corresponden a la primera fila de la *i*-ésima matriz, $p_{i,113}, \dots, p_{i,226}$ son los píxeles corresponden a la segunda fila de la *i*-ésima matriz, y así sucesivamente. A partir de esta información se calcularon 28 variables candidatas, como se puede observar en la Figura 3.

$$C_i = (p_{i,0}, \dots, p_{i,12768}) \in \mathbb{R}^{12769}, p_{i,j} \in \{0, 1, \dots, 255\}, 0 \leq i \leq 59, 0 \leq j \leq 12768$$

- 1st to 9th variable $\longrightarrow \text{card}(\{p_{i,j} : 25(k-1) \leq p_{i,j} < 25k, 0 \leq j \leq 12768\}), 1 \leq k \leq 9$
- 10th variable $\longrightarrow \text{card}(\{p_{i,j} : 225 \leq p_{i,j} < 256, 0 \leq j \leq 12768\})$
- 11th to 19th variable $\longrightarrow P_{10 \cdot (k-10)}(C_i), 11 \leq k \leq 19$
- 20th variable $\longrightarrow P_{99}(C_i) - P_1(C_i)$
- 21st variable $\longrightarrow \sum_{j=0}^{12768} p_{i,j}$
- 22nd variable $\longrightarrow \text{mean}(C_i)$
- 23rd variable $\longrightarrow \text{var}(C_i)$
- 24th variable $\longrightarrow \text{mode}(C_i)$
- 25th variable $\longrightarrow \text{energy}(C_i)$
- 26th variable $\longrightarrow \text{entropy}(C_i)$
- 27th variable $\longrightarrow \text{kurtosis}(C_i)$
- 28th variable $\longrightarrow \text{skewness}(C_i)$

Figura 3. Seudocódigo del procedimiento de selección de las variables más significativas.

Dentro del conjunto de variables, los primeros 10 representan el número de píxeles de cualquier celda con valores comprendidos en bandas de longitud 25, excepto la décima variable que representa el número de píxeles entre 225 y 255. Las siguientes 9 variables representan los 9 percentiles desde el percentil décimo hasta el percentil nonagésimo. Los últimos 9 representan el rango, la suma global de los 12,769 píxeles, la media, la varianza, la moda, la energía, la entropía, la kurtosis y la asimetría estadística.

Antes de decidir las variables con las que iniciar el estudio, hubo que tener en cuenta algunos requisitos que influyen en su elección. Principalmente se buscó obtener los mejores resultados globales, pero también se utilizó para minimizar el error en la clasificación de los elementos relacionados con el grupo 3, es decir, evitar clasificar elementos del grupo 1 o 2 en el grupo 3 y, viceversa al revés. Dado que también podría existir la posibilidad de que existiera más de una variable que arrojara valores similares para celdas de grupos opuestos y por lo tanto causara ruido en la información, era necesario buscar y dejar de lado esas variables.

Dada la dificultad de saber qué variables ofrecen resultados óptimos, la estrategia elegida consistió en repetir el siguiente método sucesivamente, estandarizando previamente todos los datos (60 celdas explicadas en 28 variables) en media 0 y varianza 1. Para ello se eligieron 55 celdas aleatorias 20.000 veces.

En los siguientes párrafos se explica el método y en la Figura se muestra la metodología propuesta para la selección de las principales variables. 4.

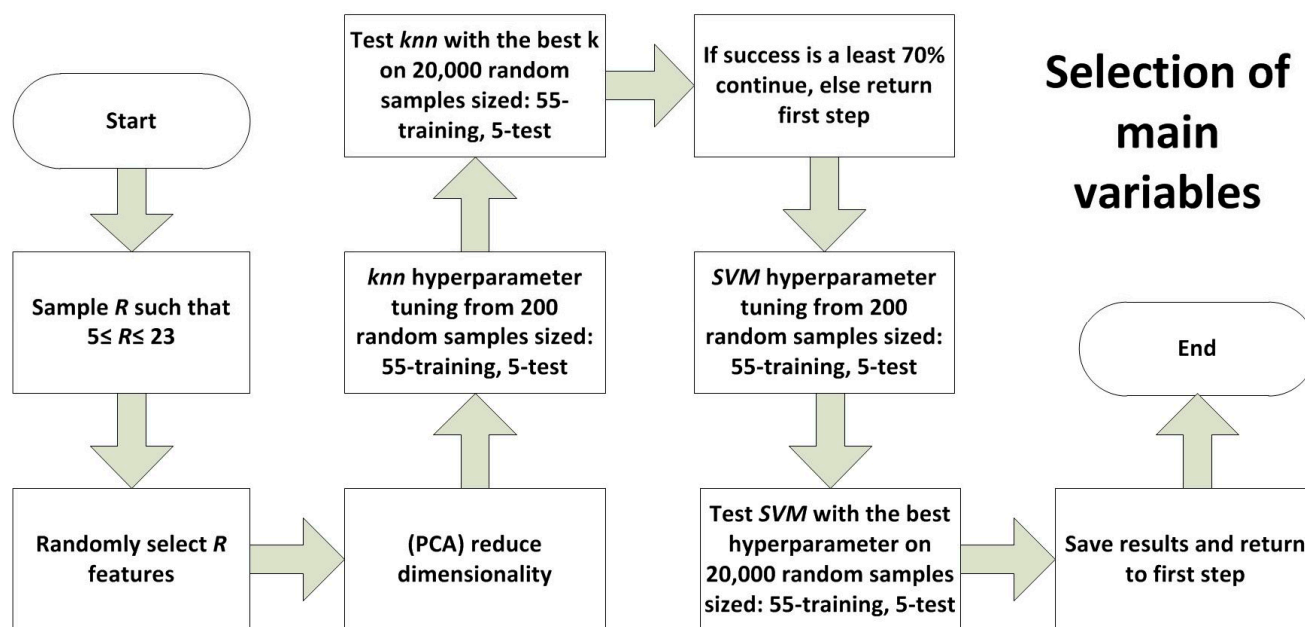


Figura 4. Metodología propuesta para la selección de principales variables.

1. Muestra R tal que $5 \leq R \leq 23$ y seleccionar características R al azar: Se obtiene un número R aleatorio entre 5 y 23. A continuación, se eligen variables R al azar entre las 28 posibles;
2. Análisis de componentes principales (PCA) Reducir dimensionalidad: Posteriormente se aplica el análisis de componentes principales, guardando las primeras variables que explican más del 99,5% de la varianza de los datos. Por tanto, a partir de ahora, se trabaja con 60 individuos explicados en como máximo R nuevas variables;
3. KNN ajuste de hiperparámetros a partir de 200 muestras aleatorias de tamaño 55, entrenamiento y 5 pruebas, prueba KNN con la mejor K en 20.000 muestras aleatorias de tamaño 55-entrenamiento y 5-test: como estábamos interesados en obtener una buena clasificación, el número óptimo de vecinos k Se buscó eligiendo entre 1 y 10, partiendo del que ofrezca mejores resultados al ser aplicado en 200 muestras aleatorias de tamaño 55-entrenamiento, 5-test. Una vez k se ha obtenido, el porcentaje de éxito con KNN ahora se estima a partir de 20.000 muestras aleatorias de tamaño 55, entrenamiento, 5 pruebas. Finalmente, se anota la proporción de malas clasificaciones relacionadas con el grupo 3. Si el porcentaje de éxito con KNN es menos estricto que el 70%, el paso 1 se convierte en:
4. SVM ajuste de hiperparámetros de 200 muestras aleatorias de tamaño 55 entrenamiento, 5 pruebas: ahora, superando el 70% de éxito con KNN , SVM se aplica teniendo en cuenta los siguientes parámetros:
 - # Core: función encargada de transportar los datos a una dimensión superior donde se pueda lograr una mejor separación de los mismos. Para el experimento se tuvieron en cuenta las funciones sigmoidea, polinomial y gaussiana y el núcleo lineal;
 - # Parámetro de penalización C : es un indicador del error que se está dispuesto a tolerar. Los valores de C de 10, 50, 75 y 100 se tuvieron en cuenta para el experimento;

- # Gamma: indica hasta qué punto se tienen en cuenta los puntos al trazar el límite de separación. Los valores gamma de 1, 0,8, 0,6, 0,4, 0,1, 0,01 y 0,001 se tuvieron en cuenta para el experimento;
- # Grado: grado de la función en el núcleo del polinomio. Se tomaron en cuenta los grados 1, 2, 3 y 4 para el experimento;
- # A partir de estos parámetros, se realizó una búsqueda entre todas las combinaciones posibles para calcular cuál de ellas ofrecía mejores resultados aplicada a 200 muestras aleatorias de tamaño 55-training, 5-test. GridSearchCV se utilizó para realizar la tarea anterior.

5. Prueba *SVM* con los mejores hiperparámetros en 20.000 muestras aleatorias de tamaño 55-entrenamiento 5-test:
Una vez que se ha obtenido la combinación ideal, la eficacia de *SVM* se estima que se ejecuta el método supervisado en función de estos parámetros y se aplica a 20.000 muestras aleatorias de tamaño 55, entrenamiento, 5 pruebas. Se guardan las proporciones de aciertos y errores de clasificación relacionados con el grupo 3;
6. Guarde los resultados y vuelva al primer paso: Vuelva al paso 1.

Para una recogida de datos suficientemente amplia, fue necesario ejecutar el proceso anterior en Python durante unas 40 h para obtener 1800 iteraciones, de las cuales 250 correspondieron a aquellos casos en los que *KNN* y *SVM* se calcularon al mismo tiempo (ya que *KNN* logró más del 70% de éxito).

El dato disponible fue de 60 celdas, lo que supuso muy poca información con la que realizar el estudio. Esto influyó en la búsqueda de los mejores parámetros para *KNN* y *SVM*, ya que *Validación cruzada* no se aplicó (difícilmente se obtendrían los mejores parámetros). En su lugar, se aplicó una búsqueda basada en 200 muestras aleatorias de tamaño 55-train, 5-test.

Considerando los resultados obtenidos, se podría concluir que las variables 4, 5, 8, 10, 13, 17, 19, 20 y 24 ofrecieron buenos resultados globales así como bajas tasas de error al realizar clasificaciones relacionadas con el grupo 3. Por lo tanto, esas variables se han seleccionado como representativas y se redujo el número de ellas de 9 a 7 aplicando *PCA* (guardando las primeras variables que explican más del 99,5% de la varianza de los datos). Posteriormente se volvieron a realizar grupos de 55 celdas para entrenar cada modelo, y se validó con los 4 clasificadores restantes en cada caso. Los clasificadores a probar fueron los siguientes: *KNN*, *SVM*, *RF*, y *MLP*. Por otra parte, *CNN* también fue probado. Sin embargo, este clasificador no siguió la misma estrategia que los cuatro algoritmos mencionados anteriormente (utilizó directamente el 60, 113 × 113 matriz como datos de entrada). Estos modelos fueron elegidos y comparados por ser los más utilizados en clasificación [30–32].

A continuación se incluyen algunos detalles de la arquitectura de los modelos utilizados. *KNN* y *SVM* Las arquitecturas ya han sido descritas en los párrafos anteriores, en los que se detalló la selección de variables representativas. El número de vecinos considerados en *KNN* fue igual al número de ellos obtenido en el algoritmo anterior cuando se seleccionaron las variables representativas. Lo mismo ocurrió con los hiperparámetros obtenidos en *SVM*.

De acuerdo a *RF*, cada clasificador se construyó en base a 500 árboles. Adicionalmente, *ajuste de hiperparámetros* se aplicó, combinando los siguientes parámetros:

1. Profundidad máxima: representa el número máximo de niveles permitidos en cada árbol de decisión. Se tuvieron en cuenta los valores 20, 40, 60, 80 y 100;
2. Puntos mínimos por nodo: este es el número mínimo de datos permitidos en cada partición. Se tuvieron en cuenta los valores 1, 2, 3, 4 y 5.
3. Máximo de variables: indica el número máximo de variables (cho $\sqrt{\frac{n}{2}}$ sen al azar) que se tienen en cuenta al particionar un nodo. Generalmente $\sqrt{\frac{n}{2}}$, n eres tú $\sqrt{\frac{n}{2}}$ como $\sqrt{\frac{n}{2}}$ parámetro estándar, donde n es el número de variables totales, pero n - 1, n y n + 1 se tuvieron en cuenta.

En el caso de *MLP*, el modelo se construyó a partir de una capa de entrada formada por 7 neuronas (coincidiendo con la dimensionalidad de los datos), una primera capa oculta formada por 128 neuronas, una segunda capa oculta formada por 64 neuronas y una capa de salida formada de sólo

3 neuronas (coincidiendo con el número de clasificaciones). La red neuronal creada era densa, una red formada por neuronas que estaban conectadas a todas las neuronas posibles pertenecientes a capas contiguas. El activador utilizado en el proceso fue el rectificador o *ReLU* activador, excepto en la última capa donde el *softmax* se utilizó la función. Además, *ajuste de hiperparámetros* se aplicó, teniendo en cuenta los siguientes parámetros:

1. *Épocas*: indica la cantidad de veces que la red neuronal lee los datos de la muestra de entrenamiento para ajustarse a ellos (traducido en una actualización sucesiva de sus parámetros). Se tuvieron en cuenta los valores 25, 50, 75, 100, 150 y 200;
2. *Lotes*: indica la velocidad con la que se actualizan los parámetros de la red a medida que avanzan las épocas. Se tuvieron en cuenta los valores 15, 25, 50, 75, 100, 150 y 200.

Para la validación de los diferentes modelos y obtención de resultados, la metodología utilizada con cada uno de los 4 clasificadores se muestra en la Figura 5.

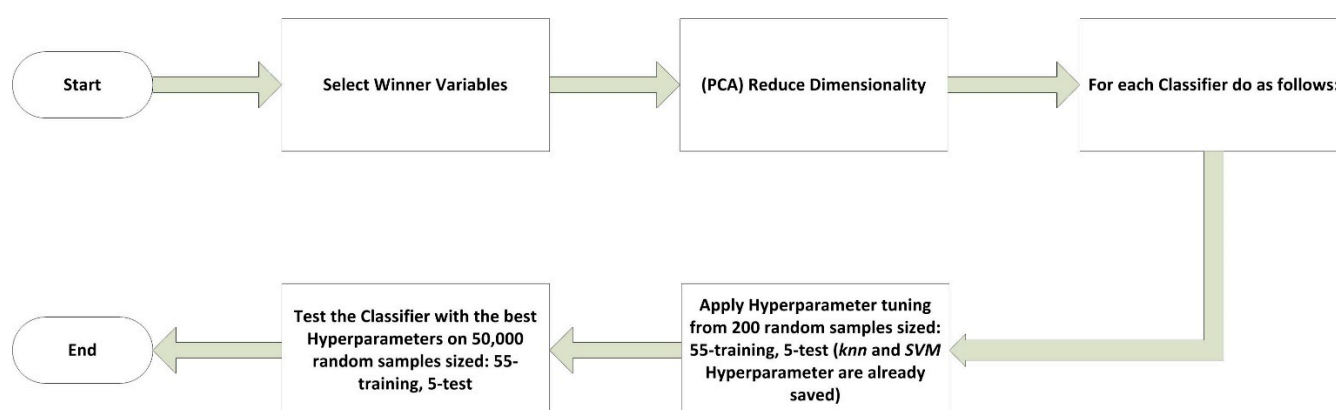


Figura 5. Metodología utilizada con cada uno de los 4 clasificadores.

En el caso de *CNN*, fue construido con un enfoque similar al *MLP*. Para encontrar suficientes patrones para resolver este problema, hemos utilizado una capa convolucional con 64 filtros y un tamaño de kernel de 3×3 . En cuanto a reducir la dimensionalidad, también utilizamos una capa maxpool. Finalmente, se introdujo una capa densa de 128. En esta arquitectura, no utilizamos la optimización dinámica de parámetros. Las redes necesitan un elevado número de épocas para entrenarse (unas 1000). Usamos el Optimizador de Nadam [33] ya que es el mejor en las pruebas que hemos ejecutado. También aprovechamos la parada anticipada, deteniendo el entrenamiento cuando no hemos obtenido mejores resultados en un número determinado de épocas.

Para la validación de los diferentes modelos y obtención de resultados, la metodología utilizada con cada uno de los cinco clasificadores fue la siguiente:

1. Para *KNN* y *SVM*:

una. Pruebe el clasificador con los mejores hiperparámetros (calculados previamente) en 50.000 muestras aleatorias de tamaño: 55 entrenamientos, 5 pruebas.

2. Para *RF* y *MLP*:

una. Aplique el ajuste de hiperparámetros de 200 muestras aleatorias de tamaño: 55 entrenamiento y 5 pruebas.

B. Pruebe el clasificador con los mejores hiperparámetros en 50.000 muestras aleatorias de tamaño: 55-training y 5-test.

3. Para *CNN*:

una. Pruebe el clasificador con los mejores hiperparámetros (establecidos manualmente) en 100 muestras aleatorias de tamaño: 55 entrenamiento y 5 pruebas.

3. Resultados y discusión

Esta sección proporciona una descripción concisa de los resultados experimentales, su interpretación, así como la discusión de los resultados y cómo se pueden interpretar.

3.1. Justificación de la potencia nominal inicial correcta

Como ya fue mencionado, *SOM* se utilizó para tener una idea de si la clasificación de potencia era correcta.

A continuación, figura 6 muestra cuatro mapas obtenidos considerando los datos con las variables representativas obtenidas previamente. Cada píxel representaba una de las 400 posibles neuronas de salida. Los píxeles cercanos con valores oscuros reflejan la proximidad entre sí, mientras que los píxeles cercanos con el contraste entre claro y oscuro representan la distancia. Los elementos del grupo 1 se representan en rojo, los elementos del grupo 2 en verde y los elementos del grupo 3 en azul.

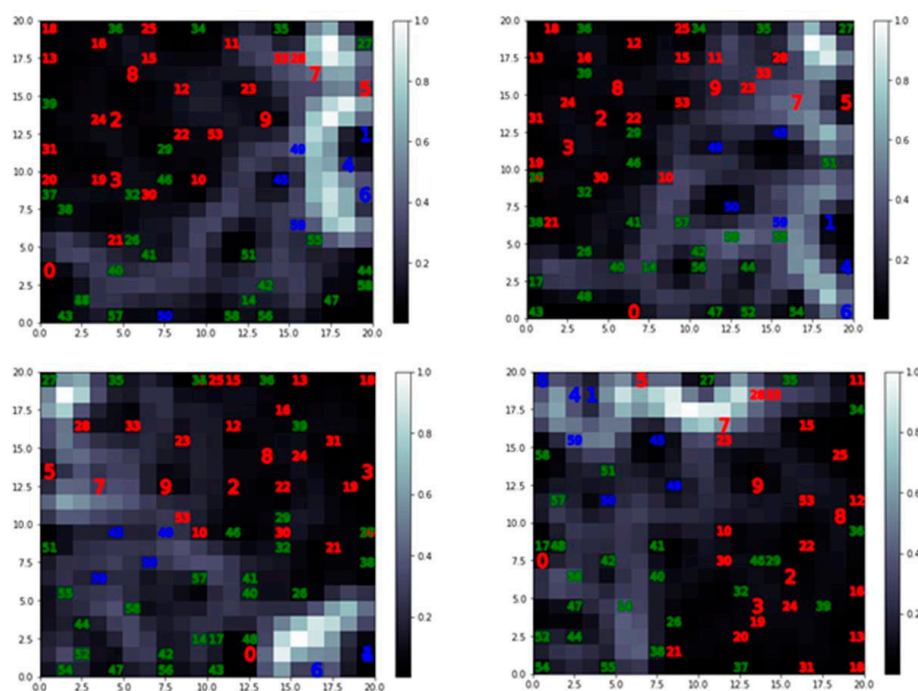


Figura 6. Resultado al aplicar *SOM* a los datos disponibles [23].

Observando los resultados, se pudo determinar la existencia de una determinada agrupación entre los elementos de un mismo color, y por tanto de un mismo grupo. Se pudo concluir que la clasificación basada en potencia era correcta. Dentro del grupo 3, se dedujo por el borde blanco que las celdas que más se distinguen del resto de los grupos fueron las número 1, 4 y 6. Además, comparando el grupo 1 y 2, se pudo concluir que era más fácil cometer un error al clasificar las celdas del grupo 2 (verde) en el grupo 1 (rojo) que de lo contrario. Esto se debió a que algunos puntos verdes se mezclaron dentro de la masa principal de puntos rojos, lo que no sucedió en el grupo verde, donde sus elementos apenas tenían elementos rojos en su interior, a excepción del elemento 0. De manera similar, al clasificar elementos del grupo 3, se pudo cometer cierto error e identificarlos como elementos del grupo 2, o viceversa, por su mayor cercanía (en comparación con el grupo 1). Esto podría comprobarse observando los resultados de las clasificaciones, que se mostrarán más adelante.

3.2. Clasificación de variables

Como ya se mencionó, la detección de las variables más importantes para el entrenamiento de los modelos es crucial. Para ello, fue necesario ejecutar el proceso anterior en Python durante 40 h para obtener 1800 iteraciones, de las cuales 250 iteraciones corresponden a aquellos casos en los que *KNN* y *SVM* se calcularon al mismo tiempo.

Siguiente, tabla 1 muestra los mejores resultados obtenidos según diferentes criterios, como el éxito en la clasificación con *KNN* y *SVM* (columnas 1 y 2), y la proporción de

malas clasificaciones causadas entre los grupos 1 y 2 con respecto al grupo 3, utilizando *KNN* y *SVM* respectivamente (columnas 4 y 5). Por lo tanto, el valor 0.9331 en la fila 5 y la columna 5 se interpretó asociando 93.331% al porcentaje de malas clasificaciones (aplicando *SVM* con las variables indicadas) que no estaban relacionadas con el grupo 3, al que solo correspondía el 6,669%. Por tanto, una alta proporción representó un error menor en la clasificación de elementos relacionados con el grupo 3. La tercera columna es la suma de la cuarta y quinta columnas.

Tabla 1. Relación entre el éxito en la clasificación y las variables utilizadas.

	<i>KNN</i> Éxito	<i>SVM</i> Éxito	<i>KNN + SVM</i> Grupo de éxito 3	<i>KNN</i> Éxito Grupo 3	<i>SVM</i> Éxito Grupo 3	Variables
Top 1 <i>KNN</i> éxito	0,745	0,756	1.425	0,734	0,691	2, 4, 5, 7, 10, 13, 16, 17 y 22
Top 1 <i>SVM</i> éxito	0,7077	0,7628	1.5973	0,6975	0,8998	4, 5, 8, 10, 13, 17, 19, 20 y 24
Top 1 <i>KNN + SVM</i> éxito grupo 3	0,7005	0,7187	1,7646	0.9093	0.8552	2, 4, 6, 8, 10, 20, 23 y 24
Top 1 <i>KNN</i> grupo de éxito 3	0,7095	0,7187	1,7646	0.9093	0.8552	2, 4, 6, 8, 10, 20, 23 y 24
Top 1 <i>SVM</i> grupo de éxito 3	0,7099	0,7401	1.6371	0,704	0.9331	2, 4, 5, 7, 8, 9, 10, 12, 15, 16, 17, 18, 19, 20 y 21

Además, se estudió la frecuencia de aparición de cada una de las 28 variables en todas las iteraciones (1800 en total) en base a dos criterios:

- Éxito con *KNN* superior al 68,5% (percentil setenta y cinco);
- La proporción de malas clasificaciones no relacionadas con el grupo 3 superior al 79,4% (percentil ochenta y cinco).

A partir del primer criterio, se dedujo que las variables candidatas buenas eran 2, 4, 5 y 7, mientras que las variables candidatas malas eran 1, 3, 6, 26, 27 y 28. A partir del segundo criterio, se concluyó que las variables candidatas buenas las variables fueron 2, 4, 6, 23 y 28 mientras que las malas candidatas fueron 1, 7, 25 y 27. Comparando los resultados obtenidos, se pudo concluir que existía una relación entre la importancia de una variable y su frecuencia de aparición siguiendo los dos criterios mencionados.

Se podría considerar el mismo tipo de criterios para *SVM*. Sin embargo, era posible que, dado el bajo número de iteraciones (250), los resultados no fueran del todo fiables.

3.3. Convergencia y resultados de los modelos

Un aspecto importante al trabajar con IA es la convergencia del modelo. A continuación, figura 7 muestra el comportamiento del hit obtenido con cada uno de los clasificadores, en función del número de iteraciones realizadas. En el caso de *MLP*, No se alcanzaron 50.000 iteraciones debido al alto costo computacional, aunque no hubo pérdida de eficiencia, como bien se observó en todos los modelos ya que hubo cierta convergencia en un menor número de iteraciones.

En el *CNN* La compilación del modelo parecía un problema crítico. La cantidad de tiempo necesaria para terminar una iteración fue extremadamente alta. Esto fue vital para decidir el número de iteraciones. Los autores finalmente decidieron utilizar 100 iteraciones de 1000 épocas. Con este número de iteraciones se obtuvo cierta convergencia.

De los resultados que se muestran en la Figura 7, el modelo más exitoso de los primeros cuartos clasificadores (*KNN*, *SVM*, *RF*, y *MLP*) era *SVM*, seguido de cerca por *MLP* y *RF*. El peor resultado lo obtuvo *KNN*; sin embargo, la tasa de éxito fue del 70,61% y se pudo considerar de alto valor. Por otro lado, se pudo ver que el quinto clasificador, *CNN*, logró un éxito superior al resto, superando el 80%.

Mesa 2 muestra los resultados (porcentaje de éxito) de los cinco modelos utilizados una vez clasificadas las celdas utilizadas para la fase de validación. También se muestra el tiempo (horas) requerido. Con respecto al tiempo, la primera columna de tiempos muestra el tiempo necesario para ubicar los parámetros ideales (*ajuste de hiperparámetros*), mientras que la segunda columna de tiempos indica el tiempo necesario para el entrenamiento del clasificador.

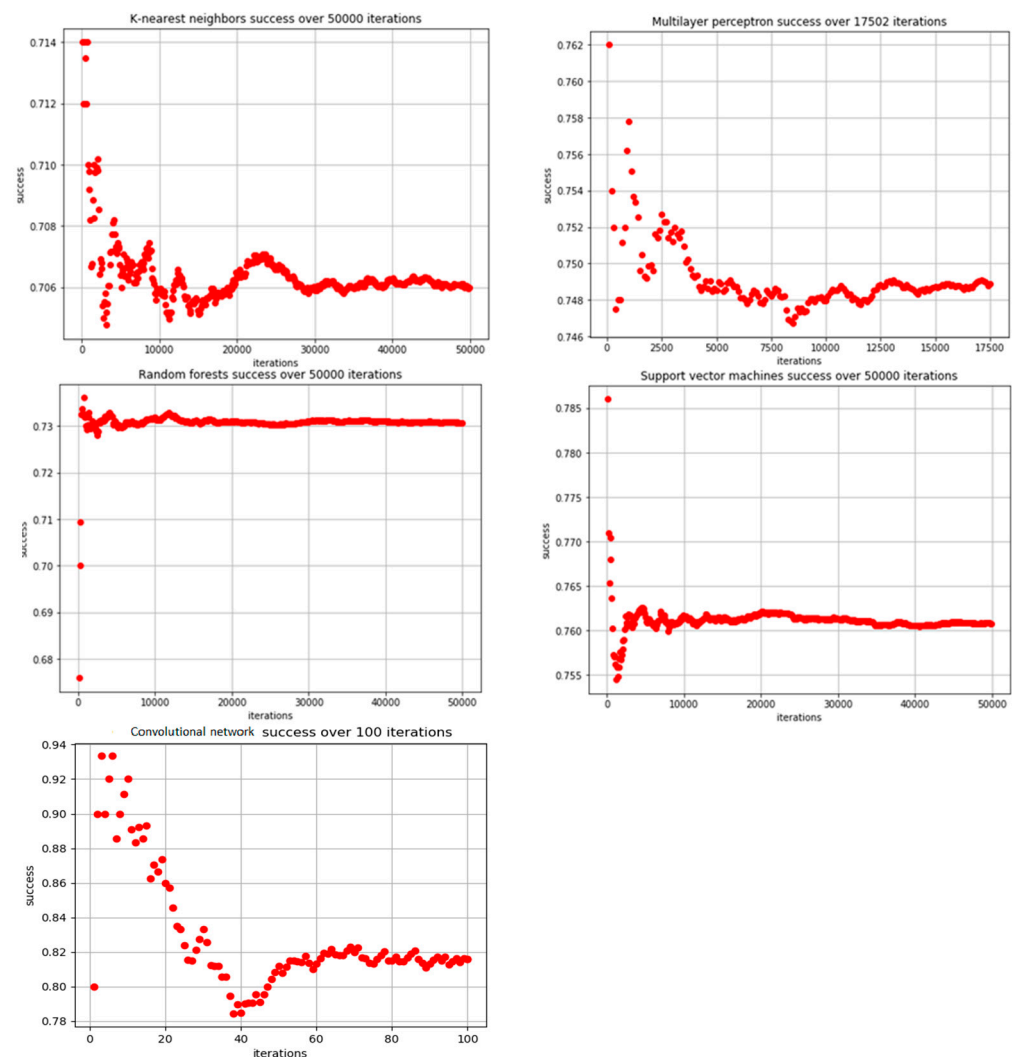


Figura 7. Comparación del éxito de la clasificación frente al número de iteraciones de los cinco modelos [23] y elaboración de los autores.

Tabla 2. Resultados de clasificación en los cinco modelos.

	Éxito de clasificación	Tiempo dedicado Ajuste de hiperparámetros (h)	Tiempo dedicado a las pruebas 50.000 (17.502 con <i>MLP</i> 100 <i>CNN</i>) Muestras (h)
<i>KNN</i>	0,7061	0,0017	0.0267
<i>SVM</i>	0,7607	0.0223	0.0196
<i>RF</i>	0,7308	2.1552	8.8071
<i>MLP</i>	0,7488	0.9633	15.5442
<i>CNN</i>	0.8160		100

En cuanto al tiempo empleado en localizar los principales parámetros, el modelo más rápido fue *KNN*, con *SVM* siguiente y con un tiempo cercano a 1 h, *MLP*. *RF* presentó el peor momento para localizar estos parámetros, requiriendo más de 2 h.

En cuanto al tiempo dedicado a la formación, *SVM* fue el más rápido, seguido de cerca por *KNN*. En el otro extremo, *RF* requirió casi 9 h, seguido de *MLP* con 15,54 h, y *CNN* con 100 h.

Observando los resultados presentados en la Tabla 2, se podría concluir que *CNN* presentó el mayor porcentaje de éxito, con un 81,6%, pero también fue extremadamente lento (alrededor de 100 h).

SVM fue el modelo que presentó la segunda mayor eficiencia (76,07%), y fue el modelo con tiempos razonablemente menores (búsqueda de parámetros y entrenamiento).

Uno de los principales objetivos de la clasificación era detectar células defectuosas (grupo 3). En este sentido, Table 3 muestra los resultados de la clasificación de celdas en el grupo 3. De la misma forma, la Tabla 3 muestra los resultados de la clasificación errónea entre grupos.

Tabla 3. Resultados de clasificación del grupo 3 y resultados de clasificación errónea entre grupos.

	Éxito en Grupo 3	1 Missclassified como 2	1 Missclassified como 3	2 Missclassified como 1	2 Missclassified como 3	3 Missclassified como 1	3 Missclassified como 2
<i>KNN</i>	0,7002	0,1017	0	0,5985	0,0117	0,0579	0,2302
<i>SVM</i>	0,9055	0,1377	0,0077	0,7678	0,0142	0,0044	0,0752
<i>RF</i>	0,7746	0,2525	0	0,5221	0,0142	0,0173	0,1939
<i>MLP</i>	0,8224	0,2166	0,0018	0,6058	0,0687	0,0073	0,0998
<i>CNN</i>	0,6956	0,2282	0,0217	0,3586	0,0865	0,0543	0,25

Centrándonos solo en la clasificación del grupo 3, *SVM* obtiene resultados de éxito del 90,55%, mientras que *MLP* obtiene el 82,24%, *RF* obtiene el 77,46%, *KNN* 70,02%, y el peor resultado es para *CNN* con 69,56%.

También se pudo observar que apenas existía confusión entre celdas del grupo 1 con el grupo 3 o celdas del grupo 2 con el grupo 3. En el primer caso, *KNN* y *RF* no presentó confusión (0%), mientras que *MLP*, *SVM*, y *CNN* presentaron 0,18%, 0,77% y 2,17%, respectivamente. En el segundo caso, *KNN*, *SVM*, y *RF* presentó un valor por debajo del 1,5%, mientras que *MLP* presentó un valor de 6,87% y *CNN* el valor más alto, 8,65%.

Aparece una mayor confusión entre los grupos 1 y 2. Como puede verse, la clasificación errónea de las células del grupo 1 como grupo 2 varió entre el 10,17% para *KNN* y 25,25% para *RF*. En el caso de clasificación errónea de las células del grupo 2 al grupo 1 fue cuando el *CNN* funcionó mejor que los otros algoritmos, con un 35,86% para *CNN*, alcanzando el 76,78% en el caso de *SVM*. Esta inexactitud se debió a la similitud entre algunas celdas, como se puede ver en la Figura 6.

4. Conclusiones

La obra presenta diferentes *PV* clasificadores de defectos celulares, utilizando cinco modelos de clasificadores diferentes, *KNN*, *SVM*, *RF*, *MLP*, y *CNN*. La clasificación se realizó en base a imágenes EL y curvas IV, todas ellas a nivel solar. *PV* nivel celular. Para todos los casos se obtuvo una buena clasificación y se analizaron las diferencias entre los modelos propuestos. *CNN* presentó el mayor porcentaje de éxito, con 81,6%, pero también fue extremadamente lento (alrededor de 100 h). *SVM* fue el modelo que presentó la segunda mayor eficiencia (76,07%), y fue el modelo con un tiempo de cálculo razonablemente corto.

La mayor aplicación de los clasificadores es la energía solar defectuosa. *PV* clasificación de células. Además, este grupo de celdas es el de mayor interés ya que es el grupo que contiene celdas con producción eléctrica casi nula. El trabajo también presenta un método para seleccionar variables de interés, que servirá para entrenar los diferentes modelos de los clasificadores. Este proceso es fundamental ya que el uso de variables sin relevancia puede provocar ruido en el entrenamiento y la consecuente obtención de malos resultados en la clasificación.

El estudio se ha centrado en *PV* células solares de un solo *PV* módulo solar. Como trabajo futuro, se propone ampliar el conjunto de datos y evaluar cómo esta ampliación mejora los resultados en cada uno de los cinco modelos. También trabajaremos en la recopilación de datos de celdas individuales (aisladas, que no forman parte de un módulo) para expandir el conjunto de datos en el futuro. También vamos a explorar la opción de generar datos sintéticos con *GAN* redes. Los autores también trabajarán en la recopilación de datos de células fotovoltaicas individuales aisladas (que no forman parte de un módulo) para ampliar el conjunto de datos en el futuro. También vamos a explorar la opción de generar datos sintéticos con *GAN* redes. Los autores también ampliarán este trabajo, aplicando la clasificación a *PV* células solares a otros módulos. Pruebas *IRT* clasificadores basados en imágenes y *IRT* y *EL*. Los clasificadores de imágenes juntos también se proponen como trabajos futuros. Este trabajo es de interés ya que se sabe que ambas técnicas son complementarias en ciertos aspectos. Otra aplicación de *AI* en el cual

los autores están trabajando en la estimación de la curva IV de *EL* imágenes y *IRT*, al nivel del solar *PV* módulo.

Contribuciones de autor: Conceptualización, LH-C., SG-S. Y VA-G.; metodología, UNA.P.-R., LH-C., SG-S. Y VA-G.; software, UNA.P.-R. y HFM-R.; validación, LH-C., SG-S., VA-G., y MdCA-G.; análisis formal, LH-C. y MdCA-G.; investigación, LH-C., SG-S., VA-G. y MdCA-G.; recursos, UNA.P.-R., LH-C., SG-S., VA-G. Y MdCA-G.; curación de datos, UNA.P.-R., SG-S. y HFM-R.; redacción: preparación del borrador original, UNA.P.-R., LH-C. Y SG-S.; redacción — revisión y edición, LH-C. y MdCA-G.; visualización, LH-C. y MdCA-G.; supervisión, LH-C. y MdCA-G.; administración de proyectos, LH-C.; adquisición de financiación, LH-C., SG-S. y VA-G. Todos los autores han leído y aceptado la versión publicada del manuscrito.

Fondos: Esta investigación fue financiada por el Ministerio de Industria, Economía y Competitividad “número de concesión” RTC-2017-6712-3 “con nombre” Desarrollo de herramientas Optimizadas de operación y mantenimiento predictivo de Plantas fotovoltaicas — DOCTOR-PV “y por la Universidad de Valladolid 2018 contrato de investigación predoctoral de la solicitud de presupuesto 180113-541A.2.01e691.

Declaración de la Junta de Revisión Institucional: No aplica.

Declaración de consentimiento informado: No aplica.

Declaración de disponibilidad de datos: No se facilitan los datos.

Conflictos de interés: Los autores declaran no tener ningún conflicto de intereses.

Referencias

- Scholten, D.; Bazilian, M.; Overland, I.; Westphal, K. La geopolítica de las renovables: nuevo tablero, nuevo juego. *La política energética* **2020**, *138*, 111059. [CrossRef]
- Gugler, K.; Haxhimusa, A.; Liebensteiner, M.; Schindler, N. Oportunidades de inversión, incertidumbre y energías renovables en los mercados eléctricos europeos. *Energy Econ.* **2020**, *85*, 104575. [CrossRef]
- REN21. *Informe de estado global de energías renovables 2020*; Secretaría de REN21: París, Francia, 2020; ISBN 978-3-948393-00-7. Disponible en línea: <http://www.ren21.net/gsr-2020/> (consultado el 15 de enero de 2021).
- Behzadi, A.; Arabkoohsar, A. Estudio de viabilidad de un sistema de energía de edificio inteligente que comprende paneles solares FV / T y una unidad de almacenamiento de calor. *Energía* **2020**, *210*, 118528. [CrossRef]
- Fachrizal, R.; Ramadhani, UH; Munkhammar, J.; Widmin, J. Evaluación combinada de la capacidad de alojamiento de PV-EV para una red de distribución de BT residencial con carga de EV inteligente y reducción de PV. *Sostener. Red de redes de energía* **2021**, *26*, 100445. [CrossRef]
- Thornbush, M.; Golubchikov, O. Ciudades energéticas inteligentes: la evolución del nexo ciudad-energía-sostenibilidad. *Reinar. Dev.* **2021**, *100626*, 100626. [CrossRef]
- Hernandez-Callejo, L.; Gallardo-Saavedra, S.; Alonso-Gómez, V. Una revisión de los sistemas fotovoltaicos: Diseño, operación y mantenimiento. *Sol. Energía* **2019**, *188*, 426–440. [CrossRef]
- Gallardo-Saavedra, S.; Garzaandez-Callejo, L.; Duque-Pmirez, O. Análisis cuantitativo de tasas y modos de falla en plantas fotovoltaicas. *Energía* **2019**, *183*, 825–836. [CrossRef]
- Gallardo-Saavedra, S.; Garzaandez-Callejo, L.; Alonso-García, MDC; Santos, JD; Morales-Aragónmis, JI; Alonso-Gómez, V.; Mas tón-helechoandez, A.; Gonzalez-Rebollo, MA; MercadoInez-sacristaan, O. Caracterización no destructiva de defectos de células solares fotovoltaicas mediante electroluminiscencia, termografía infrarroja, curvas I – V y pruebas visuales: estudio experimental y comparación. *Energía* **2020**, *205*, 117930. [CrossRef]
- Blakesley, JC; Castro, FA; Koutsourakis, G.; Laudani, A.; Lozito, GM; Fulginei, FR Hacia la extracción no destructiva de la curva característica de la celda IV individual a partir de las mediciones del módulo fotovoltaico. *Sol. Energía* **2020**, *202*, 342–357. [CrossRef]
- Morales-Aragónmis, J.; Gallardo-Saavedra, S.; Alonso-Gómez, V.; SaSánchez-Pacheco, F.; Gonzalez, M.; MercadoInez, O.; Muñoz-GarcIsoy.; Alonso-GarcIsoy.; Garzaandez-Callejo, L. Electrónica de bajo costo para rastreo iv online a nivel de módulo fotovoltaico: Desarrollo de dos estrategias y comparación entre ellas. *Electrónica* **2021**, *10*, 671. [CrossRef]
- Morales-Aragónmis, J.; Alonso-GarcIsoy.; Gallardo-Saavedra, S.; Alonso-Gómez, V.; Balenzategui, J.; Redondo-Plaza, A.; Garzaandez-Callejo, L. Medición distribuida online de curvas iv oscuras en plantas fotovoltaicas. *Apl. Sci.* **2021**, *11*, 1924. [CrossRef]
- Jordan, DC; Silverman, TJ; Wohlgemuth, JH; Kurtz, SR; VanSant, KT Modos de degradación y falla fotovoltaica. *Prog. Fotovoltio. Res. Apl.* **2017**, *25*, 318–326. [CrossRef]
- Gallardo-Saavedra, S.; Garzaandez-Callejo, L.; Duque-Perez, O. Revisión tecnológica de la instrumentación utilizada en la inspección termográfica aérea de plantas fotovoltaicas. *Renovar. Sostener. Energía Rev.* **2018**, *93*, 566–579. [CrossRef]
- Gallardo-Saavedra, S.; Hernandez-Callejo, L.; Duque-Pérez, O.; Hernandez, L. Influencia de la resolución de imagen en inspecciones termográficas aéreas de plantas fotovoltaicas. *IEEE Trans. Ind. Informar.* **2018**, *14*, 5678–5686. [CrossRef]

16. BallestIn-Fuertes, J.; Muñoz-Cruzado-Alba, J.; Sanz-Orsorio, JF; Garzaandez-Callejo, L.; Alonso-Gómez, V.; Morales Aragonés, JI; Gallardo-Saavedra, S.; MercadoInez-Sacristan, O.; Mas tón-helechoandez, UNA. Novedosa técnica de mantenimiento de electroluminiscencia de plantas fotovoltaicas a escala de servicios públicos mediante un controlador de inversor de potencia bidireccional. *Apl. Sci.* **2020**, *10*, 3084. [\[CrossRef\]](#)
17. Gligor, A.; Dumitru, C.-D.; Grif, H.-S. Solución de inteligencia artificial para la gestión de una unidad de producción de energía fotovoltaica. *Procedia Manuf.* **2018**, *22*, 626–633. [\[CrossRef\]](#)
18. Wang, H.; Liu, Y.; Zhou, B.; Li, C.; Cao, G.; Voropai, N.; Barakhtenko, E. Investigación taxonómica de la inteligencia artificial para la predicción determinista de la energía solar. *Energy Convers. Manag.* **2020**, *214*, 112909. [\[CrossRef\]](#)
19. Kayri, I.; Gencoglu, MT Predicción de la producción de energía de un panel fotovoltaico a través de redes neuronales artificiales utilizando indicadores atmosféricos. *Computación neuronal. Apl.* **2017**, *31*, 3573–3586. [\[CrossRef\]](#)
20. Li, L.-L.; Wen, S.-Y.; Tseng, M.-L.; Chiu, ASF Predicción de matriz fotovoltaica en el método de potencia de salida a corto plazo en un sistema de generación de energía centralizada. *Ana. Oper. Res.* **2018**, *290*, 243–263. [\[CrossRef\]](#)
21. Hussain, M.; Dhimish, M.; Titarenko, S.; Mather, P. Algoritmo de detección de fallas fotovoltaicas basado en redes neuronales artificiales que integra dos parámetros de entrada bidireccionales. *Renovar. Energía* **2020**, *155*, 1272–1292. [\[CrossRef\]](#)
22. Cho, K.-H.; Jo, H.-C.; Kim, E.-S.; Park, H.-A.; Park, JH Método de diagnóstico de fallas del generador fotovoltaico utilizando una máquina de vectores de soporte. *J. Electr. Ing. Technol.* **2020**, *15*, 1669–1680. [\[CrossRef\]](#)
23. Pmirez-romero, A.; Garzaandez-Callejo, L.; Gallardo-Saavedra, S.; Alonso-Gómez, V.; Alonso-García, MdC; Mateo-Romero, clasificador de defectos de células fotovoltaicas de alta frecuencia: una comparación de modelos. En *Actas de la III Conferencia Iberoamericana de Ciudades Inteligentes*, San Josmi, Costa Rica, 9 a 11 de noviembre de 2020; págs. 257–273.
24. Bishop, CM Reconocimiento de patrones y aprendizaje automático. En *Ciencias de la información y estadística*; Springer: Nueva York, NY, EE. UU., 2006; págs. 21–24. ISBN 9780387310732.
25. Kohonen, T. Formación autoorganizada de mapas de características topológicamente correctos. *Biol. Cybern.* **mil novecientos ochenta y dos**, *43*, 59–69. [\[CrossRef\]](#)
26. Kohonen, T. *Autoorganización y memoria asociativa*, 2ª ed.; Springer: Berlín / Heidelberg, Alemania, 1988; ISBN 978-3540183143.
27. Rumelhart, DE; Hinton, GE; Williams, RJ Learning representations by back-propagating errors. *Naturaleza* **1986**, *323*, 533–536. [\[CrossRef\]](#)
28. Silverman, BW; Jones, MC; Fix, E. Una contribución importante al análisis discriminante no paramétrico y la estimación de densidad: comentario sobre fix y hodge (1951). *En t. Stat. Rvdo.* **1989**, *57*, 233. [\[CrossRef\]](#)
29. Kohonen, T. *Mapas autoorganizados*; Serie Springer en Ciencias de la Información: Berlín / Heidelberg, Alemania, 2001; Volumen 30, ISBN 978-3-540-67921-9.
30. Kothari, S.; Oh, H. Redes neuronales para el reconocimiento de patrones. *Adv. Computación.* **1993**, *37*, 119–166. [\[CrossRef\]](#)
31. Clancey, WJ Clasificación heurística. *Artif. Intell.* **1985**, *27*, 289–350. [\[CrossRef\]](#)
32. Rodrigues, MA Invariantes para el reconocimiento y clasificación de patrones. En *Aprendizaje conjunto*; World Scientific: Singapur, 2000; Volumen 42.
33. Dozat, T. *Incorporación de Nesterov Momentum en Adam*; Trabajo de ICLR: París, Francia, 2016; págs. 1–4.