

Metodología para la Clasificación de Caligrafías de Cordón de Soldadura MIG analizando el rendimiento de diferentes algoritmos de Procesamiento de Imágenes y Aprendizaje de Máquinas

Methodology for the Classification of MIG Welding Bead Calligraphy by analyzing the performance of different Image Processing and Machine Learning algorithms

Osiris Giuseppe García Zarate¹, Martín Valtierra Rodríguez¹, Juan Primo Benítez Rangel¹, Juan Pablo Amézquita Sánchez¹, José Luis González Córdoba^{1*}

¹ ENAP -Research Group, CA-Sistema Dinámicos y Control, Universidad Autónoma de Querétaro (UAQ), Campus San Juan del Río, Río Moctezuma 249, Col. San Cayetano, San Juan del Río 76807, México.

*Autor de correspondencia: jose.gonzalez.cordoba@uaq.mx

Resumen

El presente trabajo propone una metodología para la clasificación de caligrafías en cordones de soldadura utilizando técnicas de procesamiento de imágenes y aprendizaje de máquina. La metodología incluye las siguientes etapas: preprocesamiento de las imágenes mediante filtros adaptativos, extracción de características geométricas y de textura asociados a las clases de soldadura en condiciones limpias y sucias, así como la implementación de tres algoritmos de inteligencia artificial (Redes Neuronales Convolucionales - CNN, K-Nearest Neighbors - KNN y Máquinas de Soporte Vectorial - SVM) ejecutados desde una computadora personal cargada con imágenes de cordones de soldadura previamente preparadas, y la evaluación del desempeño de cada algoritmo mediante diferentes métricas de precisión y matrices de confusión. Los resultados obtenidos permiten tener un panorama general de los tres enfoques para la clasificación de patrones de soldadura, bajo distintas configuraciones de preprocesamiento y resolución de las imágenes. Cada algoritmo se evaluó bajo diferentes configuraciones de preprocesamiento y resoluciones de imagen para determinar su capacidad de distinguir características y patrones geométricos en los cordones de soldadura tipo “zigzag” y “lineal”, antes y después de limpieza. Los resultados obtenidos muestran que, aunque las redes neuronales convolucionales (CNN) alcanzan mayores precisiones con imágenes de alta resolución y preprocesamiento adecuado, los modelos KNN y SVM ofrecen una clasificación efectiva incluso con resoluciones intermedias y menor preprocesamiento. Esto indica que KNN y SVM son alternativas viables cuando se busca un balance entre costo computacional y precisión, permitiendo una implementación eficiente en sistemas con recursos limitados o para aplicaciones de respuesta rápida. En particular, el modelo SVM demostró una mayor estabilidad y capacidad de generalización bajo diferentes configuraciones de preprocesamiento y resolución, mientras que KNN, aunque más sensible a variaciones en la calidad de la imagen, logró un desempeño adecuado en condiciones intermedias. La metodología propuesta es capaz de clasificar los patrones de soldadura tipo “zigzag” y “lineal”, antes y después de la limpieza.

Palabras clave: Clasificación de imágenes, Redes Neuronales Convolucionales (CNN), K-Nearest Neighbors (KNN), Máquinas de Soporte Vectorial (SVM), Caligrafías de Soldadura, Aprendizaje de máquinas.

Introducción

La automatización del control de calidad en los procesos de soldadura ha ganado relevancia en los últimos años debido al creciente aumento en los requisitos de fusión para asegurar la calidad de las uniones soldadas. Según la Sociedad Americana Soldadura (AWS, American Welding Society), se estima que aproximadamente el 60% de las fallas estructurales en componentes metálicos están relacionadas con uniones de soldadura defectuosas, lo cual impacta directamente en la seguridad y durabilidad de las piezas fabricadas, así como de los usuarios o personal que interactúan con las piezas manufacturadas (American Welding Society, 2015).



Asimismo, el mercado global de automatización de soldadura alcanzó un valor de 5.2 mil millones de dólares en 2022, proyectándose un crecimiento anual del 8% durante los próximos cinco años, impulsado por la necesidad de mejorar la eficiencia y minimizar errores en la producción (Market Research Future, 2023). En este contexto, la forma y calidad del cordón de soldadura juegan un papel crucial en la seguridad y desempeño de las estructuras, especialmente en sectores como la automoción y la manufactura. Debido a esto, se requiere cada vez más de sistemas que permitan una evaluación precisa y en tiempo real de características específicas del proceso de soldadura para asegurar la integridad estructural y operativa de las piezas producidas.

En la literatura, existen múltiples enfoques orientados a la inspección automática de soldadura, utilizando técnicas de visión por computadora y aprendizaje automático, los cuales se han desarrollado de manera progresiva en la última década. Nguyen et al. (2016) propusieron una metodología basada en la extracción de características geométricas del cordón de soldadura utilizando transformadas de onda para detectar deformaciones, logrando un 89% de precisión en la detección de irregularidades. En un estudio posterior, Kim et al. (2018) utilizaron SVM para clasificar defectos internos en cordones de soldadura de arco sumergido, alcanzando un 91% de precisión en la identificación de discontinuidades. Más recientemente, Zhao y sus colaboradores, en 2019 implementaron técnicas de aprendizaje profundo combinadas con transformadas de Fourier para la segmentación del cordón, obteniendo una tasa de precisión del 92% en la clasificación de defectos en soldadura MIG. Así mismo Choi et al. (2020) llevaron a cabo una investigación sobre la identificación de patrones de soldadura en imágenes térmicas, utilizando redes neuronales recurrentes (RNN) para el monitoreo en tiempo real, con un 93% de precisión en la detección de anomalías. Finalmente, en el 2021 Yang y sus colaboradores emplearon CNN para la detección de defectos superficiales en patrones de soldadura TIG, logrando una precisión del 95.4% en la identificación de defectos como grietas y porosidades. Sin embargo, estos estudios se enfocan principalmente en la detección de defectos y la evaluación de la calidad visual de la unión de soldadura, dejando de lado características como la caligrafía del cordón, que describe el patrón de aplicación de la soldadura y, que está ligada íntimamente con la calidad de la unión de las piezas.

En este sentido, la caligrafía con la que se realiza el cordón de soldadura podría proporcionar información relevante sobre la técnica del soldador y el cumplimiento de condiciones de operación específicas. Además, la mayoría de los estudios reportados no ofrecen bases de datos públicas que permitan el desarrollo de nuevos algoritmos, limitando la replicabilidad y el avance en el área. Debido a esto, los métodos de soldadura actuales no se enfocan en la identificación de patrones de caligrafía ni en impacto que tienen en la calidad de la unión de los materiales, abriendo una oportunidad para el desarrollo de metodologías de IA que discriminen entre el tipo de caligrafías en función de parámetros como la penetración, la escoria, el chorreo, entre otros. Es por ello, que el presente trabajo propone una nueva metodología para la clasificación de patrones de caligrafía en cordones de soldadura basándose en visión por computadora e IA. Experimentalmente se desarrollaron los patrones de soldadura zigzag y lineal con ayuda de un Robot Industrial ABB IRB4600, evaluando el estado del cordón tanto en la condición de “recién soldado” como en la de “después de la limpieza” (nombradas en este trabajo como “limpia” y “sucia”). El objetivo de utilizar un robot industrial es asegurar un control sobre la repetitividad de las piezas soldadas en un entorno industrial real. El proceso de validación del estado del cordón de soldadura es planteado como un proceso intermedio en el ciclo de producción, permitiendo detectar automáticamente si se ha realizado la limpieza del cordón antes de avanzar a la siguiente etapa..

Marco Teórico.

Soldadura MIG

En la soldadura por Gas Inerte de Metal (MIG, por sus siglas en inglés), la alimentación continúa del material de aporte o micro alambre del electrodo y el gas protector se combinan para crear un arco eléctrico estable entre el alambre y el metal base, fundiendo el material del electrodo para formar la soldadura. Este arco se mantiene constante, y el gas protector expulsado de la tobera evita la contaminación del baño de fusión, lo que resulta esencial para mantener la integridad estructural de la soldadura. A diferencia de otros métodos de soldadura, como el uso de electrodos revestidos, la soldadura MIG no requiere cambios frecuentes de electrodos ni la eliminación de escoria, lo que la hace ideal para aplicaciones industriales (Weman & Lindén, 2006).

La energía calorífica necesaria para la soldadura MIG se describe mediante la ecuación:



$$Q = I^2 R t \quad (1)$$

Donde Q es el calor generado, I es la corriente de soldadura, R es la resistencia eléctrica y t el tiempo que fluye la corriente. Este calor derrite el material del electrodo y parte del material base, creando un charco de soldadura que, al solidificarse, forma una unión fuerte y duradera (Weman K. y Lindén G., 2006).

Los principales parámetros de ajuste en la soldadura MIG son: el diámetro del microalambre, el voltaje, la velocidad de alimentación de alambre y corriente, la velocidad de soldadura, la inductancia, el saliente de alambre, y la posición de la antorcha; además de los los patrones de tejido de la antorcha, la alimentación de hilo pulsado, entre otros (Madavi et al., 2021; Mishara. 2014; Weman & Lindén, 2006). La Figura 1 muestra los dos tipos de caligrafías utilizadas para el desarrollo del sistema, las imágenes fueron tomadas de dos cordones con parámetros adecuados para hilos consistentes y estéticos.

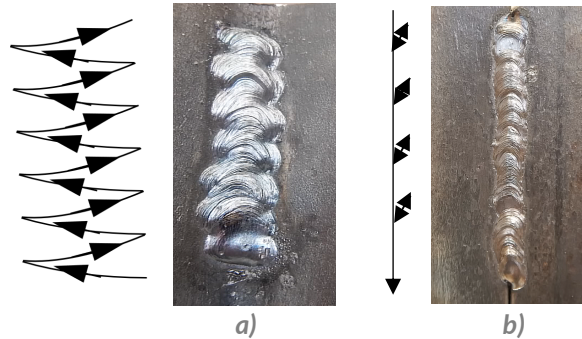


Figura 1. Caligrafías utilizadas en el proyecto actual
a) Caligrafía zigzag b) Caligrafía lineal (Autoría propia).

Técnicas de Procesamiento de Imágenes

Filtrado

El filtrado de imágenes implica la manipulación de la imagen original mediante la aplicación de una máscara (o kernel), que es una pequeña matriz de coeficientes utilizada para procesar cada píxel en función de sus vecinos (Tomasi & Manduchi, 1998). La máscara de transformación se desplaza por toda la imagen, calculando un valor ponderado de los píxeles vecinos según los coeficientes de la matriz. Este proceso, conocido como convolución, permite asignar el nuevo valor calculado al píxel central de la imagen resultante, generando así el efecto deseado de filtrado. Existen diferentes tipos de filtros que pueden ser utilizados para la mejora de las características de las imágenes, algunos de los cuales se presentan a continuación:

Filtro Gaussiano

"El filtro gaussiano suaviza la imagen aplicando una función de distribución normal, lo que permite calcular los valores ponderados de los píxeles cercanos. Este filtro es especialmente útil para reducir el ruido mientras se preservan mejor las características importantes de la imagen, como los bordes (Szeliski, R. (2010). El filtro gaussiano aplica una máscara cuyos coeficientes están determinados por la función gaussiana.

$$G(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \quad (2)$$

Donde $G(x, y)$ es el valor de la máscara en la posición (x, y) . Por otro lado, σ es la desviación estándar de la distribución gaussiana, que controla el grado de suavizado. Finalmente x y y son las coordenadas de la máscara con respecto del centro. La imagen filtrada se obtiene realizando la convolución de la máscara con la imagen original, este proceso es descrito en (3):

$$I'(i, j) = \sum_{k=-a}^a \sum_{l=-b}^b I(i+k, j+l) G(k, l) \quad (3)$$

Filtro Bilateral

El filtro bilateral es una técnica de suavizado avanzada que no solo tiene en cuenta la cercanía espacial de los píxeles, sino también la similitud de sus valores de intensidad. A diferencia de otros métodos de suavizado que tienden a difuminar los bordes, el filtro bilateral preserva los bordes nítidos al asignar pesos mayores a los píxeles con valores de intensidad similares al píxel central (Tomasi, C. & Manduchi, R., 1998). Matemáticamente, el filtro bilateral se define como:

$$I'(i,j) = \frac{\sum_{k=-a}^a \sum_{l=-b}^b I(i+k,j+l) G_s(k,l) G_r(I(i,j), I(i+k,j+l))}{\sum_{k=-a}^a \sum_{l=-b}^b G_s(k,l) G_r(I(i,j), I(i+k,j+l))} \quad (4)$$

Donde $I'(i,j)$ es el valor del píxel filtrado en la posición (i,j) , así como $I(i+k,j+l)$ es el valor de los píxeles vecinos, $G_s(k,l)$ es una función gaussiana que depende de la distancia espacial entre los píxeles (k,l) . Finalmente $G_r(I(i,j), I(i+k,j+l))$ es otra función gaussiana que depende de la diferencia de intensidad entre el píxel central y sus vecinos.

Filtro Anisotrópico

El filtrado anisotrópico, también conocido como difusión anisotrópica, es una técnica que suaviza las regiones homogéneas de la imagen mientras preserva los bordes mediante un proceso iterativo de difusión. A diferencia de los filtros isotrópicos como el gaussiano, la difusión anisotrópica ajusta la intensidad de los píxeles en función de la variación de intensidad local, controlada por un coeficiente de difusión que varía espacialmente (Perona, P. & Malik, J., 1990). Matemáticamente, la ecuación de difusión anisotrópica se expresa como:

$$\frac{\partial I}{\partial t} = \nabla(c(x,y,t)) \nabla I \quad (5)$$

Donde $\partial I / \partial t$ es la variación de la intensidad de la imagen a lo largo del tiempo t , así como ∇ es el operador gradiente y $c(x,y,t)$ corresponde al coeficiente de difusión, que controla la velocidad y dirección de la difusión en cada punto (x,y) . El coeficiente de difusión $c(x,y,t)$ se define como:

$$c(x,y,t) = \exp\left(-\frac{(\|\nabla I(x,y)\|)^2}{K^2}\right) \quad (6)$$

Donde $\|\nabla I(x,y)\|$ es el valor del gradiente de la imagen en la posición (x,y) , que mide la variación de intensidad. Por otro lado K es un parámetro que controla la sensibilidad del filtro a los bordes. Las técnicas de filtrado permiten preparar las imágenes de entrada para la extracción de características relevantes, facilitando la diferenciación de patrones geométricos complejos, como las caligrafías de soldadura, aumentando la precisión de los modelos de clasificación empleados en este estudio (Nguyen et al., 2016; Zhao et al., 2019).

Técnicas de Aprendizaje de Máquina para la Clasificación de Imágenes

Redes Neuronales Convolucionales (CNN)

Las CNN son un tipo de arquitectura de red especialmente diseñada para el procesamiento de imágenes, compuestas por capas convolucionales que aplican filtros para extraer características relevantes (LeCun, Bengio & Hinton, 2015), seguidas de capas de agrupamiento (pooling) que reducen la dimensionalidad de la imagen, manteniendo sus características más significativas (Krizhevsky, Sutskever & Hinton, 2012). La capacidad de las CNN para capturar patrones locales y estructuras jerárquicas las hace ideales para el análisis de imágenes complejas, como las de cordones de soldadura, donde las características geométricas y de textura son esenciales para la clasificación.

La CNN se configura ajustando hiperparámetros como la tasa de aprendizaje y el número de épocas. Además, se emplea la regularización mediante la desconexión aleatoria (Dropout). Un aspecto clave de las CNN es la capacidad de reutilizar los pesos a través de la convolución, lo que reduce significativamente el número de parámetros y evita el sobreajuste, incluso en redes con muchas capas (Simonyan & Zisserman, 2014). Además, en las capas densas para evitar el sobreajuste, cada imagen se normaliza y se ajusta a las dimensiones de entrada requeridas, y se organiza en conjuntos de entrenamiento, validación y prueba. La



evaluación del rendimiento del modelo se realiza utilizando métricas como la precisión, la sensibilidad (recall) y la matriz de confusión, permitiendo comparar el impacto de diferentes resoluciones y filtros de preprocesamiento para clasificar los patrones en las condiciones especificadas.

La Figura 2 muestra la arquitectura de la CNN. En este esquema se presenta cada una de las etapas, desde la entrada de la imagen hasta la capa de salida, destacando el proceso de convolución y agrupamiento, y la posterior clasificación de las imágenes de entrada.

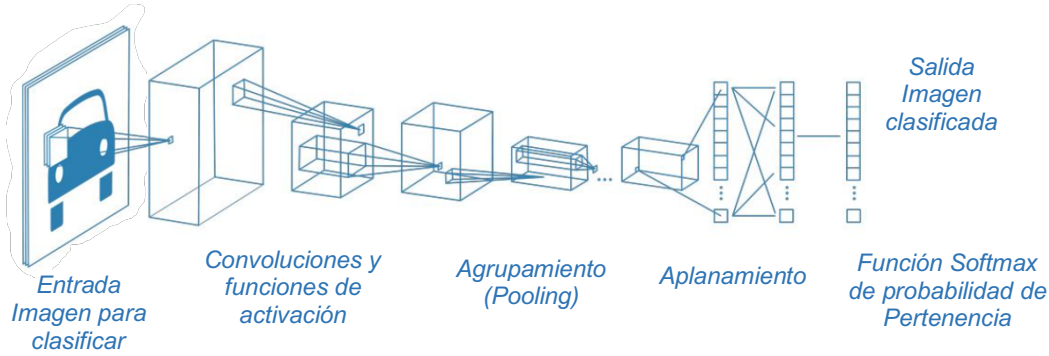


Figura 2. Diagrama general de red neuronal convolucional CNN adaptada de Matlab (2024).

Matemáticamente, la operación de una capa convolucional puede describirse como:

$$Z_{i,j,k} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \sum_{c=0}^{C-1} X_{i+m,j+n,c} W_{m,n,c,k} b_k \quad (7)$$

Donde $Z_{i,j,k}$ es el valor de la activación de salida en la posición (i, j) para el filtro k , así como $X_{i+m,j+n,c}$ es el valor de entrada de la imagen en la posición $(i + m, j + n)$ para el canal c , $W_{m,n,c,k}$ es el peso del filtro k en la posición (m, n) para el canal c , b_k es el sesgo (bias) asociado al filtro k , M y N son las dimensiones del filtro, y finalmente C es el número de canales de la imagen de entrada. La operación descrita por las CNN se realiza para cada filtro y en cada posición de la imagen de entrada, generando un mapa de características que resalta ciertas estructuras geométricas o de textura presentes en la imagen (LeCun, Bengio & Hinton, 2015). Este proceso permite que las capas convolucionales detecten patrones específicos como bordes, contornos y texturas a diferentes niveles de detalle, dependiendo del tamaño y número de filtros utilizados (Krizhevsky, Sutskever & Hinton, 2012).

Después de las capas convolucionales y de agrupamiento, las características extraídas se reorganizan en un vector de una dimensión (Aplanamiento) antes de pasar a capas densas para la clasificación final. La salida de una capa densa puede representarse como:

$$y_i = \sigma(\sum_{j=0}^{D-1} W_{j,i} P_j + b_i) \quad (8)$$

Donde y_i es la salida de la neurona i en la capa densa, así como $W_{j,i}$ es el peso que conecta la característica P_j con la neurona i , por su parte b_i es el sesgo de dicha neurona i , por otro lado D es el número total de características de entrada y finalmente $\sigma(\cdot)$ es la función de activación de la capa densa (softmax para la clasificación). Las capas densas son responsables de aprender relaciones más abstractas entre las características extraídas por las capas convolucionales, combinando de manera jerárquica la información capturada a lo largo de la red para generar la predicción final de la clase (Simonyan & Zisserman, 2014).

K-Nearest Neighbors (KNN)

El algoritmo k-Nearest Neighbors (KNN) es un método de clasificación no paramétrico que se basa en calcular la distancia entre puntos en el espacio de características. La clasificación se determina considerando los k vecinos más cercanos a una muestra de prueba, donde la clase dominante entre los vecinos determina la etiqueta asignada (Cover & Hart, 1967).

El rendimiento se mide mediante la precisión y la matriz de confusión, analizando la capacidad del modelo para separar correctamente las clases. Dado que KNN es un método basado en instancias, su desempeño puede ser afectado por la calidad y la dimensionalidad de las características (Weinberger & Saul, 2009), por lo que la combinación de filtros puede ayudar a resaltar estructuras relevantes para mejorar la clasificación.



El algoritmo KNN clasifica un punto de prueba x comparando sus características con los puntos de entrenamiento utilizando la métrica de distancia euclidiana, que se calcula como:

$$d(x, x_i) = \sqrt{\sum_{j=1}^D (x_j - x_{i,j})^2} \quad (9)$$

Donde D representa el número total de características, por su parte x_j y $x_{i,j}$ son las coordenadas de x y x_i en la dimensión j . Posterior a la clasificación del punto de prueba x y se identifican los k -vecinos más cercanos y se asigna la clase más frecuente entre ellos.

Máquinas de Soporte Vectorial (SVM)

Las SVM son clasificadores supervisados que buscan maximizar el margen entre clases mediante la identificación de un hiperplano que separa las muestras de distintas categorías (Cortes & Vapnik, 1995). El objetivo principal de las SVM es encontrar un hiperplano que maximice la distancia (margen) entre las muestras más cercanas de cada clase, conocidas como vectores de soporte, con el fin de minimizar el error de clasificación y mejorar la capacidad de generalización del modelo (Cristianini & Shawe-Taylor, 2000). Los parámetros del modelo, como la constante de regularización C , se ajustan para controlar el margen entre las clases y prevenir el sobreajuste.

$$\min \left(\frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^N \xi_i \right) \quad (10)$$

Sujeto a:

$$y_i(\omega x_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0 \quad (11)$$

Donde ω es el vector de pesos, así como C es la constante de regularización que controla el margen, y_i es la etiqueta de la clase para la muestra i , x_i es el vector de características de la muestra i , y finalmente ξ_i es la variable de holgura que permite la clasificación correcta de muestras con cierto margen de error. Las SVM se entrenan y evalúan en cada combinación de filtro y resolución para determinar la influencia del preprocesamiento en la clasificación de patrones.

El rendimiento se mide utilizando la precisión y las matrices de confusión para evaluar la capacidad del modelo de diferenciar las clases. A diferencia de KNN, que es sensible a la dimensionalidad, las SVM son más robustas ante variaciones en el espacio de características, lo que las hace adecuadas para condiciones donde las diferencias entre clases son sutiles y los patrones están distribuidos de manera no lineal.

Metodología Propuesta

La metodología desarrollada en el presente estudio se centra en la clasificación de imágenes de caligrafías de dos tipos de cordones de soldadura MIG, utilizando con un enfoque de procesamiento de imágenes y aprendizaje máquina. Se utilizan diversas configuraciones de preprocesamiento y tres modelos de clasificación distintos: CNN, KNN y SVM para evaluar la efectividad de cada técnica en la diferenciación de patrones geométricos y estados del cordón (recién soldado y después de limpieza) bajo condiciones industriales. A continuación, se describe cada etapa del proceso metodológico, desde la adquisición de datos hasta la evaluación del rendimiento de los algoritmos. En la Figura 3 se presenta un diagrama general de la metodología propuesta para la clasificación de las imágenes de los cordones de soldadura. El proceso incluye la adquisición de datos, la preparación y preprocesamiento de imágenes, la implementación de los modelos de clasificación (CNN, KNN y SVM) y finalmente la evaluación y comparación de resultados. Las imágenes analizadas se capturaron a partir de cordones de soldadura MIG generados en condiciones controladas. El conjunto de datos se compone de cuatro clases principales, cada una de ellas representando una combinación de geometría y estado de limpieza del cordón: 1. zigzag sucio (recién soldado), 2. zigzag limpio (después de la limpieza), 3. lineal sucio (recién soldado) y 4. lineal limpio (después de la limpieza).

Cada imagen se evaluó en diferentes tamaños, manteniendo una relación de aspecto constante de 1:1.6 para evitar distorsiones geométricas y asegurar una representación proporcional de los patrones. Con el fin de destacar características específicas y reducir el ruido, las imágenes fueron sometidas a un preprocesamiento mediante tres tipos de filtros, adicionalmente, se incluyeron imágenes sin filtro para evaluar el impacto de



estos preprocesamientos. Cada clase representa una combinación de geometría y estado de limpieza, simulando condiciones de operación reales.

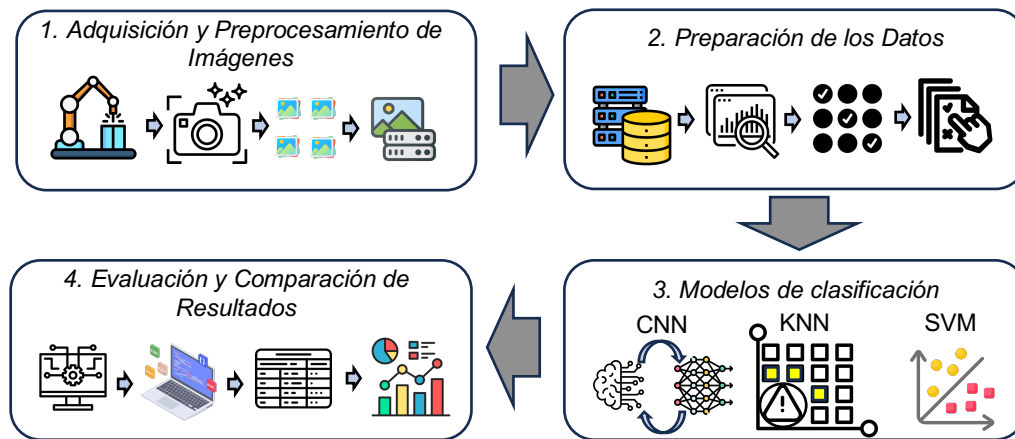


Figura 3. Diagrama de Metodología propuesta (Autoría propia).

Las imágenes son preprocesadas aplicando los filtros seleccionados (gaussiano, bilateral y anisotrópico), cada imagen, tanto en formato RGB o en escala de grises, se representa como una matriz de intensidad de píxeles, donde cada componente se organiza en un vector de características. Para las imágenes en RGB, se manejan tres vectores correspondientes a las componentes roja, verde y azul, mientras que para las imágenes en escala de grises se utiliza un único vector. Los filtros se aplican a estos vectores mediante la convolución descrita en las ecuaciones (3), (4) y (5), según el filtro seleccionado, actualizando los valores de cada vector acorde a las transformaciones correspondientes.

Posteriormente, los vectores resultantes de cada imagen filtrada se organizan en carpetas etiquetadas según la clase, el tipo de filtro aplicado, creando subconjuntos organizados de datos para cada configuración. Cada imagen fue etiquetada manualmente con un número de clase correspondiente a una de las cuatro categorías. Posteriormente, los datos se dividieron en tres conjuntos: entrenamiento (60%), validación (20%) y prueba (20%). Los tres enfoques para la clasificación fueron implementados, entrenados y evaluados bajo las distintas configuraciones de preprocesamiento y tamaño de resolución acorde a lo mencionado anteriormente.

El modelo CNN fue desarrollado con la biblioteca TensorFlow/Keras y consta de tres capas convolucionales, seguidas de capas de agrupamiento y capas densas. Las capas convolucionales utilizan 32, 64 y 128 filtros respectivamente, con función de activación ReLU y un tamaño de filtro de 3x3. Las capas de agrupamiento por su parte aplican un tamaño de 2x2 para reducir la dimensionalidad y preservar características significativas. Posteriormente la arquitectura se completa con una capa de Aplanado (Flatten), una capa densa con 64 neuronas con activación ReLU, y una capa de salida con activación softmax para la clasificación final en las cuatro categorías. El entrenamiento se llevó a cabo con una tasa de aprendizaje de 0.0001, durante 50 épocas y un tamaño de lote de 16. Para evaluar el rendimiento del modelo, se utilizaron las métricas de sensibilidad y la matriz de confusión, generando gráficos de evolución del entrenamiento y resultados numéricos en cada combinación de resolución y filtro de preprocesamiento.

El algoritmo KNN se implementó utilizando la distancia Euclidiana como métrica principal para medir la proximidad entre las muestras. Las imágenes fueron convertidas en vectores de características mediante aplanamiento y posteriormente estandarizadas utilizando la técnica StandardScaler para garantizar que cada característica contribuya de manera equitativa al cálculo de distancias, evitando que características con valores mayores dominen la métrica (Hastie, Tibshirani & Friedman, 2009). Se estableció $k = 3$ como el número óptimo de vecinos tras realizar validación cruzada, y para mejorar la eficiencia en grandes volúmenes de datos, se utilizó el algoritmo de búsqueda BallTree. La capacidad del modelo para diferenciar entre clases se evaluó utilizando las métricas de precisión y matriz de confusión en cada condición específica.

En el caso de la SVM se utilizó un kernel lineal, que proyecta los datos en un espacio de mayor dimensión para facilitar la separación de las clases en un dominio donde las fronteras no son lineales en el espacio original. La selección de una máscara lineal se justifica por la naturaleza de los patrones de soldadura, donde

las diferencias geométricas y de textura pueden ser capturadas eficientemente sin la necesidad de transformaciones complejas. Además de un parámetro de regularización C ajustado entre 0.1 y 10 para maximizar el margen entre las clases y evaluar su impacto en el rendimiento. Se realizó un análisis de sensibilidad, utilizando precisión y matriz de confusión como métricas de evaluación, para poder observar de determinar la influencia de los filtros de preprocesamiento en la capacidad de diferenciación de los patrones de soldadura.

Evaluación y Comparación de Resultados

La evaluación de los modelos se realizó utilizando métricas de precisión, Sensibilidad y F1-score, para analizar la capacidad de cada método para diferenciar entre las cuatro clases de patrones de soldadura bajo las diferentes resoluciones y filtros. Se generan tablas comparativas de resultados que muestran el desempeño de cada modelo en cada configuración, así como matrices de confusión que reflejan la distribución de predicciones correctas e incorrectas para cada clase. Para el caso de CNN, además se graficaron las curvas de precisión y pérdida a lo largo de las épocas de entrenamiento para ilustrar el comportamiento de cada modelo durante el ajuste.

Experimentación y Resultados

El proceso experimental comenzó con la preparación de piezas de soldadura bajo condiciones controladas, simulando el entorno real de un taller industrial utilizando un brazo robótico industrial ABB IRB4600-255/40 para garantizar la repetitividad y consistencia en las muestras producidas. El sistema cuenta con un control de precisión en el movimiento y en los parámetros de soldadura, como la corriente, el voltaje y la velocidad de avance del micro alambre, lo que permite generar patrones de cordón de soldadura con reproducibilidad y sin variaciones significativas entre muestras. Esto se realizó utilizando el inversor Millermatic 140, que ofrece un rango de corriente de salida de 30 a 140 A y una capacidad de ajuste automático de parámetros. Las piezas de prueba, fabricadas en acero al carbono, se procesaron bajo condiciones idénticas en términos de temperatura y material base para evitar influencias externas que pudieran distorsionar el resultado del cordón.

Las condiciones limpias se refieren a cordones tratados posteriormente mediante un proceso de limpieza superficial, mientras que las condiciones sucias simulan el estado fresco del cordón de soldadura inmediatamente después del proceso, sin ningún tratamiento adicional. La Figura 4 muestra las piezas de soldadura realizadas, con ejemplos de cordones antes y después del proceso de limpieza para las clases seleccionadas clase 1: zigzag sucia, clase 2: zigzag limpia, clase 3: lineal sucia y clase 4: lineal limpia. Cada muestra fue digitalizada utilizando una cámara con una resolución de 12 MP y un sensor RGB estándar, con una apertura de $f/1.6$, con luz de día. Cada imagen de los cordones de soldadura fue capturada y preprocesada en tres resoluciones distintas: 250x400, 350x560 y 450x720 píxeles, para evaluar el impacto de la resolución en la capacidad del modelo para extraer características relevantes.

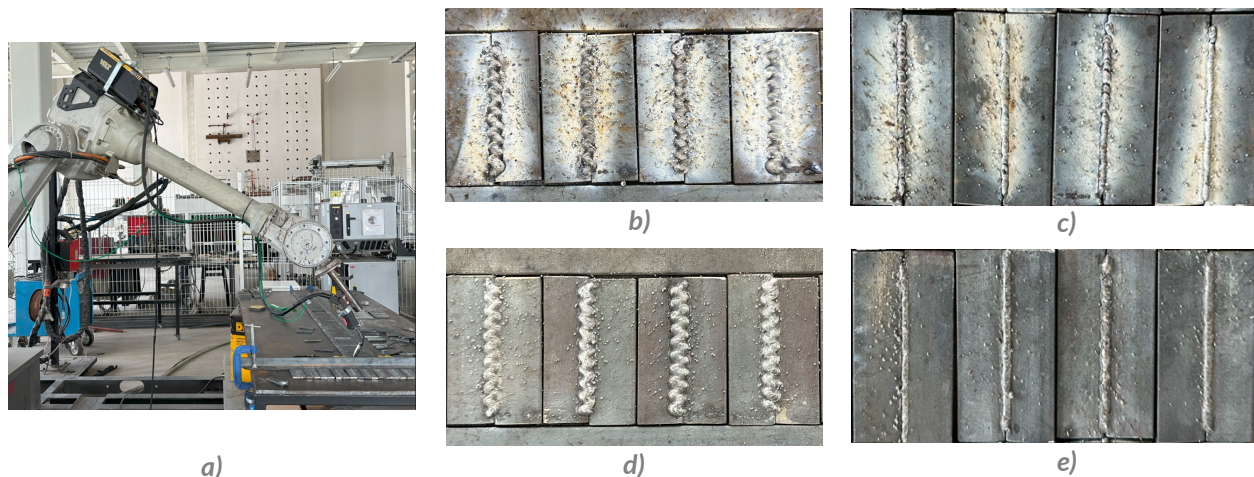


Figura 4. a) Proceso de manufactura de piezas muestra b) clase "zigzag sucia" c) clase "zigzag limpia" d) clase "lineal sucia" e) clase "lineal limpia" (Autoría propia).

Posteriormente, las imágenes procesaron y se les aplicaron los distintos filtros seleccionados (filtro anisotrópico que suaviza la imagen manteniendo bordes, filtro bilateral que elimina ruido preservando características del patrón y filtro gaussiano que reduce el ruido y resalta la geometría general) para comparar el efecto del preprocesamiento en la diferenciación de patrones. Se incluyó también un conjunto de imágenes sin filtros como referencia base. En la figura 5 se pueden observar las variaciones de las clases con cada uno de los filtros y sin filtros, se nota una calidad de imagen diferente en la mayoría de los casos, resaltando de forma diferente la Región de interés (ROI).

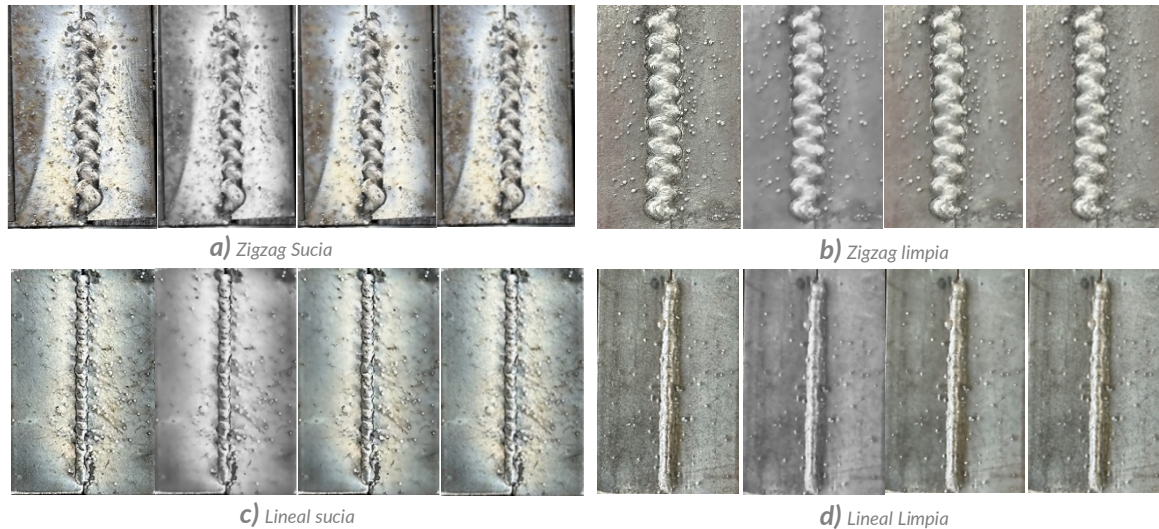


Figura 5. Aplicación de filtros, de izquierda a derecha: 1. Sin filtro, 2. Filtro Anisotrópico, 3. Filtro Bilateral 4., Filtro Gaussiano respectivamente para cada clase (Autoría propia).

Para cada combinación de resolución y filtro, se calcularon métricas de precisión, sensibilidad y Puntuación F1. Posteriormente, se seleccionaron dos configuraciones basándose en a) precisión promedio calculada sobre todas las combinaciones de filtro y resolución, como métrica de rendimiento global y b) mayor precisión observada individualmente, para identificar las mejores condiciones para clasificar patrones de soldadura. La Tabla 1 presenta un resumen del desempeño promedio de precisión, la sensibilidad y la puntuación F1 considerando las configuraciones posibles por modelo y resolución.

Tabla 1. Métricas promedio por modelo y resolución.

Modelo	Resolución	Precisión Promedio (%)	Sensibilidad Promedio (%)	Puntuación F1
CNN	250x400	64.96	65.90	62.31
	350x560	58.78	57.46	52.50
	450x720	69.87	67.21	65.96
KNN	250x400	68.78	62.81	60.65
	350x560	68.56	58.62	58.34
	450x720	67.84	58.78	58.28
SVM	250x400	78.25	78.68	78.06
	350x560	78.65	78.78	78.31

450x720 79.65 79.75 79.34

Por otra parte, la Tabla 2 muestra los valores de métricas obtenidos para la configuración específica que muestra el mejor desempeño de cada modelo.

Tabla 2. Configuración de mejor desempeño por modelo.

Modelo	Resolución	Filtro	Color o Gris	Precisión (%)	Sensibilidad (%)	Puntuación F1 (%)
CNN	450 x 720	Sin Filtro	Gris	84	81.25	81.5
KNN	350 x 560	Gaussiano	Gris	77	69	68
SVM	450 x 720	Bilateral	Color	83	82.75	82.5

Análisis de Resultados

En la Tabla 3, se observa que la precisión promedio del modelo SVM es la más alta, alcanzando un 78.85%, seguida de KNN con 68.39% y CNN con 64.54%. Esto sugiere que SVM y KNN presentan un mejor desempeño general frente a variaciones en las condiciones de resolución y preprocesamiento, mientras que CNN tiende a ser más sensible a estos cambios. En contraste, la Tabla 2 permite identificar las configuraciones de mejor desempeño para cada modelo, destacando el uso de resoluciones altas preprocesamiento adecuado para asegurar la precisión en las clases de soldadura.

Tabla 3. Valores promedio por modelo.

Modelo	Precisión Promedio (%)	Sensibilidad Promedio (%)	Puntuación F1
CNN	64.54	63.53	60.26
KNN	68.39	60.07	59.09
SVM	78.85	79.07	78.57

El análisis de los resultados muestra que, aunque el modelo CNN alcanzó la mayor precisión en una configuración con un 84%, el modelo SVM se considera superior en términos de desempeño global debido a su estabilidad en diferentes configuraciones de preprocesamiento y resolución. Al evaluar las métricas complementarias, como puntuación F1 y la sensibilidad, SVM obtiene un promedio de 81.2% en puntuación F1 y una sensibilidad promedio de 80.7% en comparación con los 75.5% y 72.3% respectivamente obtenidos por la CNN. Esto sugiere que SVM no solo tiene una precisión elevada, sino que también mantiene un equilibrio entre la identificación correcta de clases y la minimización de falsos positivos y falsos negativos (F1-Score), lo que lo convierte en un modelo más confiable para aplicaciones donde la generalización es clave.

Por otro lado, CNN muestra un buen rendimiento en configuraciones específicas, pero con una precisión promedio de 64.54%, y una puntuación F1 y sensibilidad más bajos, que reflejan una mayor variabilidad en sus resultados. Esto implica que CNN tiende a sobre ajustarse en condiciones con menor resolución o sin preprocesamiento adecuado, lo que lo hace menos robusto frente a cambios en las características de la imagen. Sin embargo, cuando se utilizan imágenes con alta resolución (450x720), la CNN logra capturar patrones complejos con alta precisión, mostrando que puede ser una excelente opción en condiciones experimentales estables.



La Figura 6 muestra las curvas de precisión durante las épocas de entrenamiento y validación para el modelo CNN con mejor desempeño, utilizando imágenes en escala de grises con resolución alta (450x720) sin filtro. Se observa cómo la precisión de entrenamiento (curva azul) incrementa de manera rápida en las primeras épocas, alcanzando un valor máximo cercano al 90%, mientras que la precisión de validación (curva naranja) fluctúa considerablemente y se estabiliza alrededor del 60-70%. Esta diferencia entre las curvas de entrenamiento y validación indica una ligera tendencia al sobreajuste, ya que la precisión de validación no logra seguir el mismo comportamiento estable de la precisión de entrenamiento.

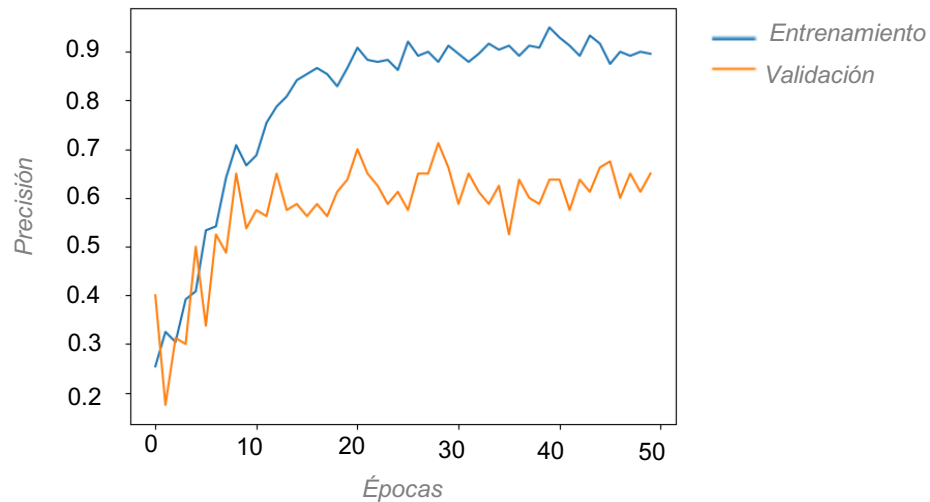


Figura 6. Gráficas de precisión para modelo CNN, Gris, alta resolución (450x720) sin filtro (Autoría propia).

Por otro lado, la figura 7 muestra la gráfica de pérdida del modelo de CNN con mejor desempeño, indicando una disminución constante en la curva de entrenamiento, alcanzando valores cercanos a 0.1 al final de las 50 épocas, mientras que la pérdida de validación se mantiene con variaciones más pronunciadas, sugiriendo que el modelo CNN está aprendiendo de manera efectiva, pero con cierta dificultad para generalizar a los datos de validación. Esto indica que, aunque el modelo puede aprender patrones complejos en el conjunto de entrenamiento, podría beneficiarse de un ajuste adicional para mejorar su capacidad de generalización.

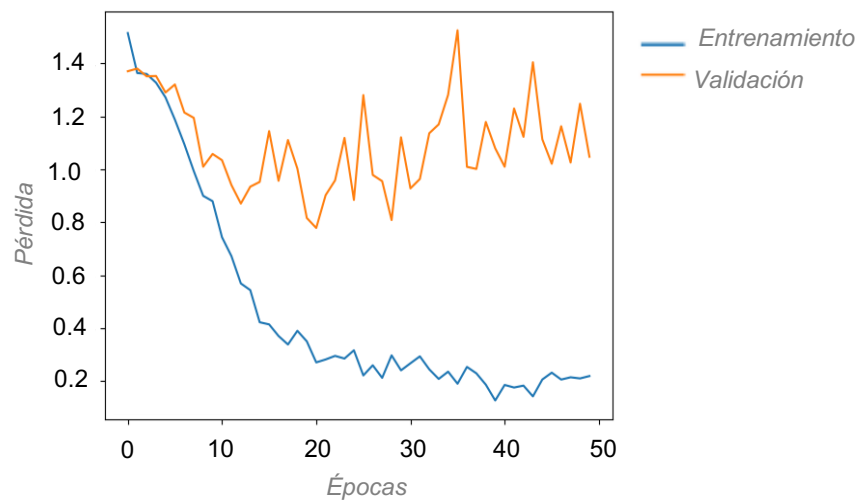


Figura 7. Gráfica de Pérdida para modelo CNN, Gris, alta resolución (450x720) sin filtro (Autoría propia).

La matriz de confusión mostrada en la Figura 8, presenta un buen desempeño para identificar la clase "lineal limpia", con 21 aciertos de 21 muestras, alcanzando una precisión del 100% en esta categoría. Para la clase "zigzag sucia", el modelo logra clasificar correctamente 18 de 22 muestras, con un error moderado de

clasificación hacia la clase "zigzag limpia" (2 instancias) y hacia las clases "lineal limpia" y "lineal sucia" (1 instancia cada una).

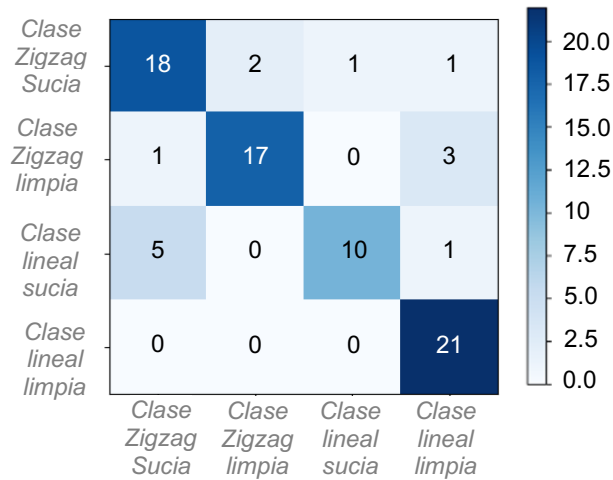


Figura 8. Matriz de Confusión para modelo CNN, Gris, alta resolución (450x720) sin filtro (Autoría propia).

En contraste, la clase "lineal sucia" muestra una mayor cantidad de confusiones, con 5 instancias clasificadas incorrectamente como "zigzag sucia", indicando que el modelo tiene dificultades para distinguir entre estas dos clases en condiciones de suciedad. Para la clase "zigzag limpia", se observa un nivel aceptable de aciertos (17/21 muestras), pero con 3 confusiones hacia la clase "lineal limpia". Este comportamiento puede deberse a las similitudes geométricas y de textura compartidas entre las clases "zigzag" y "lineal" en situaciones donde la suciedad afecta las características visuales, lo que reduce la capacidad de diferenciación del modelo. En general, la matriz sugiere que el modelo CNN es capaz de identificar correctamente las clases con patrones bien definidos como "lineal limpia", pero presenta confusiones entre clases con características más ambiguas, especialmente en condiciones de suciedad.

Por su parte, el modelo KNN mostró un buen rendimiento utilizando imágenes en escala de grises, resolución media (350x560) y preprocesamiento con un filtro Gaussiano. La Figura 9 presenta las matrices de confusión asociadas a esta configuración (mejor desempeño) y a la configuración de peor desempeño con resolución alta (450x720), a color y sin filtro. En la matriz de mejor desempeño (9-a), se observa un alto nivel de precisión en la clase "zigzag limpia" con 30 aciertos de 34 muestras, y resultado adecuado en la clase "zigzag sucia" con 25 de 30 aciertos. No obstante, se detectan algunas confusiones al clasificar la clase "lineal sucia" (6 instancias clasificadas incorrectamente como "zigzag limpia") y la clase "lineal limpia" (16 instancias clasificadas incorrectamente como "zigzag limpia"), indicando que el modelo tiende a asociar estos patrones lineales con la clase "zigzag".

En la configuración de peor desempeño (Figura 9-b), se observa un deterioro significativo en la capacidad de diferenciación, especialmente en las clases "lineal sucia" y "lineal limpia". La clase "lineal limpia" muestra 12 errores, clasificándose incorrectamente como "lineal sucia", mientras que la clase "zigzag limpia" tiene 15 errores hacia "zigzag sucia", lo que indica la alta sensibilidad del modelo a las variaciones en color y resolución. Esto indica que KNN depende en gran medida de la resolución y del preprocesamiento aplicado para diferenciar patrones geométricos, y que ciertas configuraciones pueden llevar a una pérdida drástica de precisión.

En cuanto a la SVM, la mejor combinación de parámetros se obtuvo utilizando imágenes a color, con una resolución alta de 450x720 y preprocesamiento con filtro bilateral. En la figura 10-a, se observa una alta tasa de aciertos. La clase "lineal limpia" fue identificada correctamente en 25 de 37 muestras, y la clase "zigzag sucia" con 27 aciertos de 28 muestras, lo que demuestra la capacidad del modelo para separar patrones geoméricamente distintos. Sin embargo, algunas confusiones aparecen en la clasificación de la clase "lineal sucia", con 4 instancias clasificadas incorrectamente como "zigzag sucia" y 12 instancias de "lineal limpia" clasificadas como "zigzag limpia". Estas confusiones sugieren que, aunque el filtro bilateral ayuda a resaltar características geométricas clave, aún existen ciertas similitudes entre las clases "zigzag" y "lineal" que el modelo no logra diferenciar completamente.

En contraste, la Figura 10-b muestra la matriz de confusión correspondiente a la configuración de peor desempeño, obtenida con una resolución media (350x560), sin preprocesamiento. En esta configuración, la clase "zigzag limpia" muestra un mayor nivel de confusión, con 21 aciertos y 5 errores hacia las clases "lineal limpia" y "lineal sucia", lo que indica que la falta de un preprocesamiento adecuado afecta significativamente la capacidad de la SVM para distinguir estas clases. De manera similar, la clase "lineal sucia" muestra confusiones más pronunciadas, con 4 instancias clasificadas erróneamente como "zigzag limpia".

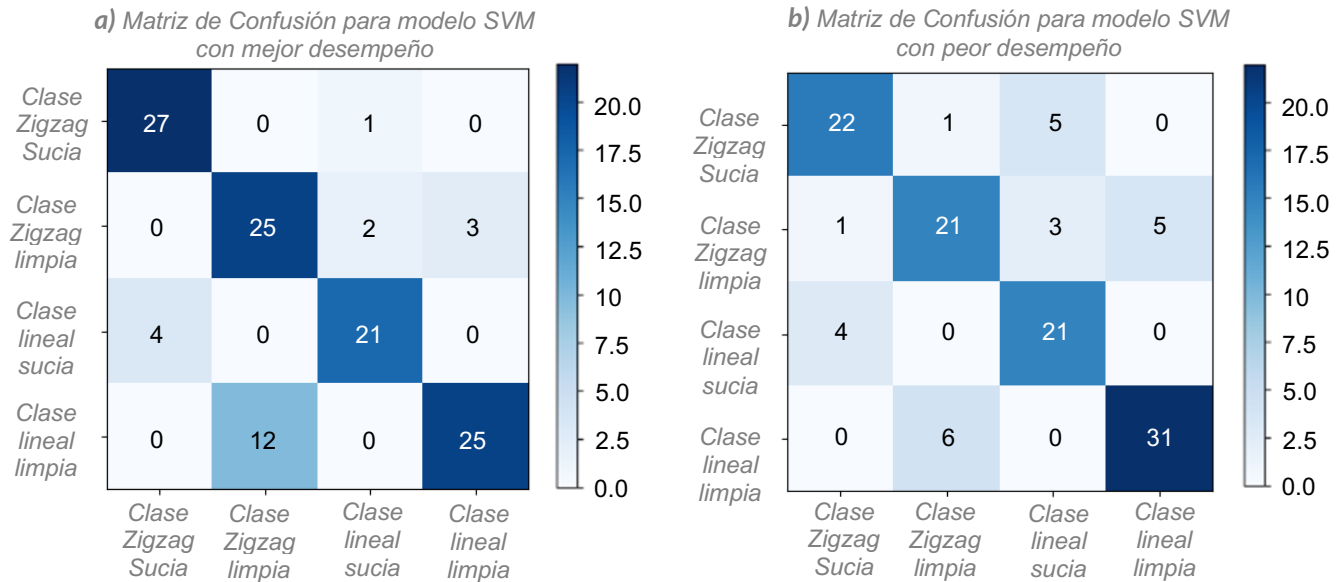


Figura 10. Matrices de Confusión a) SVM resolución 450x720, filtro bilateral, a color (mejor desempeño) b) SVM resolución 350x560, sin filtro, a color (peor desempeño) (Autoría propia).

Conclusiones

El desempeño de los modelos evaluados sugiere que la precisión de clasificación está influenciada tanto por la resolución de la imagen como por el tipo de preprocesamiento aplicado, aunque no siempre es necesario utilizar la resolución más alta para obtener buenos resultados. Se observó que, si bien las configuraciones de alta resolución combinadas con un preprocesamiento adecuado facilitan la extracción de características geométricas y de textura, resoluciones intermedias también pueden ser efectivas, especialmente cuando se busca un equilibrio entre precisión y eficiencia computacional.

Entre los tres algoritmos evaluados, SVM mostró un desempeño más consistente y robusto, alcanzando una precisión máxima del 83% con imágenes en color y resolución intermedia (450x720 píxeles) utilizando el filtro bilateral. Aunque CNN logró la mejor precisión puntual (84%), mostró una mayor variabilidad en sus resultados con cambios de configuración, obteniendo una precisión promedio del 64.54%. Esto sugiere que SVM es más confiable para generalizar bajo diferentes condiciones experimentales, mientras que CNN es más sensible a las variaciones de resolución y preprocesamiento.

Por otro lado, KNN fue el modelo más sensible a los cambios en la configuración. Su mejor desempeño se observó con imágenes en escala de grises, resolución media (350x560) y filtro gaussiano, alcanzando una precisión del 77%. Sin embargo, su precisión promedio fue del 68.39%, reflejando una mayor susceptibilidad a las variaciones en resolución y preprocesamiento. Este comportamiento indica que KNN depende en gran medida de un preprocesamiento específico para maximizar su capacidad de discriminación, pero pierde efectividad en condiciones de alta variabilidad.

En cuanto a la capacidad de generalización, CNN es capaz de capturar patrones visuales complejos, aunque tiende a sobreajustarse en condiciones de alta variabilidad, especialmente con resoluciones menores y sin filtros de preservación de características. Por otro lado, SVM mantiene un buen equilibrio entre precisión y estabilidad, destacando en la clasificación de clases con patrones claramente diferenciados, como "lineal limpia" y "lineal sucia", con un margen de error inferior al 5%.



KNN, en cambio, mostró una mayor propensión a confundir clases con características geométricas y de textura compartidas, especialmente en configuraciones sin preprocesamiento o con imágenes a color. En estas condiciones, el rendimiento disminuyó, generando errores de hasta el 16% en clases como "lineal limpia".

En general, los resultados indican que las redes neuronales convolucionales son particularmente efectivas para aplicaciones que requieren capturar patrones visuales complejos y sutiles en condiciones controladas y con preprocesamiento. SVM representa un punto intermedio, ofreciendo robustez en condiciones diversas y con mayor estabilidad bajo diferentes configuraciones de resolución. KNN puede ser adecuado en escenarios con menor variabilidad y características claramente definidas.

Finalmente, se sugiere que el uso de arquitecturas más profundas o la incorporación de técnicas como la transferencia de aprendizaje podrían mejorar la capacidad de CNN para generalizar en condiciones de mayor variabilidad. Además, la generación de datos sintéticos mediante redes generativas adversarias (GANs) podría aumentar la diversidad del conjunto de entrenamiento, reduciendo el riesgo de sobreajuste y mejorando la resiliencia de los modelos ante nuevas clases de patrones de soldadura.



Bibliografía/Referencias

- Altman, N. S. (1992). An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician*, 46(3), 175-185.
- American Welding Society. (2015). Recommended Practices for Automotive Weld Quality-Resistance Spot Welding. Retrieved from AWS Publications.
- Beyer, K., Goldstein, J., Ramakrishnan, R., & Shaft, U. (1999). When is "nearest neighbor" meaningful?. In *International conference on database theory*. Springer.
- Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory* (pp. 144-152).
- Burges, C. J. C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2), 121-167.
- Choi, S., Lee, K., & Kim, J. (2020). Identificación de patrones de soldadura en imágenes térmicas utilizando redes neuronales recurrentes (RNN) para el monitoreo en tiempo real. *Journal of Applied Thermal Engineering*, 93, 12-19.
- Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1), 21-27.
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273-297.
- Cristianini, N., & Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press.
- Fix, E., & Hodges, J. L. (1951). Discriminatory analysis, nonparametric discrimination: Consistency properties. Report Number 4, USAF School of Aviation Medicine.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer.
- Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065).
- Kim, J., Yang, H., & Park, S. (2018). Clasificación de defectos internos en soldaduras de arco sumergido utilizando máquinas de soporte vectorial (SVM). *International Journal of Welding Science*, 78(4), 45-53.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.
- Madavi, S., & Mishra, R. (2014). Parámetros y control en la soldadura MIG. *Journal of Welding Engineering*, 65(7), 88-99.
- Matlab. (2024). Convolutional Neural Networks (CNN). Recuperado de <https://la.mathworks.com/discovery/convolutional-neural-network.html>.
- Market Research Future. (2023). Global Welding Automation Market Report. Recuperado de <https://marketresearchfuture.com>.
- Nguyen, T., Do, H., & Nguyen, Q. (2016). Extracción de características geométricas para la detección de deformaciones en cordones de soldadura mediante transformadas de onda. *Journal of Computational Vision*, 45(3), 34-47.
- Perona, P., & Malik, J. (1990). Difusión anisotrópica para la preservación de bordes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7), 629-639.
- Schölkopf, B., & Smola, A. J. (2002). *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press.
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint*



arXiv:1409.1556.

Szeliski, R. (2010). Computer Vision: Algorithms and Applications. Springer.

Tomas, C., & Manduchi, R. (1998). Bilateral filtering for gray and color images. In Proceedings of the IEEE International Conference on Computer Vision. IEEE.

Weinberger, K. Q., & Saul, L. K. (2009). Distance metric learning for large margin nearest neighbor classification. Journal of Machine Learning Research, 10(2).

Weman, K., & Lindén, G. (2006). MIG/MAG Welding Guide. Woodhead Publishing.

Yang, Y., Zhao, L., & Tang, X. (2021). Detección de defectos superficiales en soldaduras TIG utilizando redes neuronales convolucionales. Journal of Manufacturing Processes, 53, 29-35.

Zhang, Z. (2016). A comprehensive evaluation of KNN for image classification.

Zhao, H., Luo, S., & Zhang, L. (2019). Clasificación de defectos en soldadura MIG utilizando aprendizaje profundo y transformadas de Fourier. Journal of Intelligent Manufacturing, 34(2), 234-246.

