



Detección de componentes electrónicos mediante el método de la bolsa de palabras visuales (BoVW)

Electronic component detection using the bag of visual words (BoVW) method

Teth Azrael Cortes Aguilar*  y Adriana Tovar Arriaga 

Departamento de Sistemas Computacionales y Electrónica, Tecnológico Nacional de México: ITJMMPyH
Camino Arenero 1101, Col. El Bajío, 45017. Zapopan, Jalisco, México

*teth.cortes@zapopan.tecmm.edu.mx

PALABRAS CLAVE: RESUMEN

Componentes electrónicos, bolsa de palabras visuales, aprendizaje supervisado, visión por computadora

El problema de la gestión de residuos electrónicos es uno de los más apremiantes en la actualidad. Sin embargo, su gestión adecuada representa una oportunidad lucrativa debido a los diversos metales preciosos y tierras raras que contienen. La tarea de clasificar componentes electrónicos reciclados suele ser realizada por humanos, pero un sistema de visión por computadora y un método de inteligencia artificial como la bolsa de palabras visuales (BoVW) podrían mejorarla. Este artículo presenta el desarrollo de una interfaz gráfica de usuario que consta de dos etapas. En la primera etapa, se crea el conjunto de datos y se entrena el modelo BoVW. En la segunda etapa, se prueba el modelo BoVW con nuevas imágenes adquiridas en tiempo real. La eficiencia del modelo BoVW mejora al aumentar el número de datos utilizados en la etapa de entrenamiento. Se lograron métricas de exactitud en la clasificación del 97%, 94%, 93%, 92% y 89% con circuitos integrados DIP 8, TO-220, display, resistencias y capacitores respectivamente. Nuestros hallazgos indican que el método BoVW puede lograr predicciones precisas en la etapa de prueba con un número menor de imágenes en la etapa de entrenamiento que otros algoritmos de aprendizaje automático.

KEYWORDS: ABSTRACT

Electronic component, bag of visual words, supervised learning, computer vision

The problem of managing electronic waste is one of the most pressing in modern society. However, the appropriate management of electronic waste represents a lucrative opportunity due to the various precious metal and rare earth elements it contains. The task of sorting recycled electronic components is usually done by humans, but a computer vision system and an artificial intelligence method such as the Bag of Visual Word (BoVW) could improve it. This paper presents the development of a graphical user interface that consists of two stages. In the first stage, the dataset is created and the BoVW model is training. In the second stage, the BoVW model is tested with new images acquired in real time. The efficiency of the BoVW model improves by increasing the number of data used in the training stage. Classification accuracies of 97%, 94%, 93%, 92% and 89% were achieved with integrated circuit DIP 8, TO-220, display, resistor and capacitor respectively. Our findings indicate that the BoVW method can achieve accurate predictions in the testing stage with a smaller number of images in the training stage than other machine learning algorithms.

• Recibido: 8 de junio de 2024 • Aceptado: 10 de septiembre de 2024 • Publicado en línea: 1 de febrero de 2025

1. INTRODUCCIÓN

La elevada demanda entre los consumidores de productos tecnológicos, ha desencadenado volúmenes sin precedentes de aparatos electrónicos, que al final de su vida útil suponen un enorme desafío para el medio ambiente y la salud humana, debido a la presencia de sustancias altamente tóxicas. La gestión de los residuos electrónicos (*e-waste*) es uno de los problemas más acuciantes para la sociedad actual. Por otro lado, la gestión adecuada de los desechos electrónicos representa una oportunidad muy lucrativa debido a los diversos elementos de tierras raras y metales preciosos que contienen [1]. Aunque son los consumidores quienes determinan el destino de los residuos electrónicos, son las empresas dedicadas al reciclaje las que logran reducir la emisión de contaminantes y recuperar elementos valiosos.

Una aplicación común en la manufactura de ensamblajes electrónicos es la detección de componentes mediante visión por computadora para la inspección de la integridad y calidad de las tarjetas de circuito impreso. Sin embargo, la mayoría de los componentes usados a menudo se descartan en lugar de reutilizarse, generando residuos electrónicos. No obstante, los componentes ensamblados en estas tarjetas de circuito impreso podrían retirarse mediante procesos térmicos o mecánicos [2]. Además, teniendo en cuenta que clasificar componentes es una tarea rutinaria, realizada manualmente por el ser humano, esto podría cambiar y mejorarse utilizando un sistema de clasificación automatizado.

El reconocimiento de objetos es una de las técnicas de visión por computadora más utilizada en la industria electrónica para identificar componentes durante los procesos de fabricación de ensamblajes [3] y en procesos de reciclaje [2]. Los distintos métodos de inteligencia artificial desarrollados en la actualidad, tienen el

potencial de mejorar la eficiencia de las técnicas de reconocimiento de objetos. Recientemente los métodos de aprendizaje automático (*Machine Learning*) y aprendizaje profundo (*Deep Learning*) son temas de investigación que buscan desarrollar procesos más exactos y con un menor costo computacional [4]. En los métodos de aprendizaje automático se entrena un modelo a partir de un conjunto de datos etiquetados (*dataset*). El desempeño del modelo entrenado, se puede expresar mediante una matriz de confusión, que permite visualizar la cantidad de aciertos y de errores, comparando los datos actuales o reales contra los datos predichos para cada clase. En la evaluación de modelos de aprendizaje automático entrenados, también se usa frecuentemente otro criterio de desempeño denominada tasa de *accuracy*, que se calcula como el número de muestras clasificadas correctamente entre el número total de muestras [5]. En la ecuación (1) los aciertos del modelo entrenado se cuentan como Verdaderos Positivos VP y Verdaderos Negativos VN mientras que los errores se cuentan como Falsos Positivos FP y Falsos Negativos FN. El total de muestras es la suma de los aciertos y los errores.

$$accuracy = \frac{VP+VN}{VP+VN+FP+FN} \quad (1)$$

En el presente artículo, se describe un sistema basado en visión por computadora para reconocer y clasificar componentes electrónicos mediante el método de aprendizaje automático denominado bolsa de palabras visuales BoVW (*Bag of Visual Words*). La aplicación de una técnica de iluminación adecuada, así como de un filtro digital para la detección de bordes y la selección de imágenes para entrenamiento sin componentes parciales en el plano focal, contribuyó a mejorar la exactitud de las predicciones.

El primer objetivo de investigación del presente trabajo, consistió en demostrar que el método BoVW es capaz de reconocer

componentes electrónicos por su tipo de empaquetado o encapsulado. El segundo objetivo fue demostrar que BoVW a pesar de ser un método más simple que otros métodos más sofisticados de aprendizaje profundo, logra métricas aceptables de *accuracy* con una menor cantidad de imágenes en la etapa de entrenamiento y en consecuencia con un menor costo computacional.

Una contribución de la presente investigación es la metodología implementada. Las imágenes de entrenamiento y prueba se adquirieron directamente de fotogramas seleccionados de una señal de video, generada por una cámara conectada a una computadora. Se programó en MATLAB una interfaz gráfica de usuario GUI (*Graphical User Interface*) que consta de dos ventanas. En la primera ventana, se etiquetan las imágenes que conforman el *dataset* para aprendizaje supervisado y se entrena el modelo BoVW. En la segunda ventana, se hacen predicciones para imágenes inéditas distintas a las de la etapa de entrenamiento. El contenido del presente artículo está dividido de la siguiente manera. En la sección 2 se describe el método de bolsa de palabras visuales utilizado para reconocer componentes electrónicos por su tipo de empaquetado. En la sección 3 se presenta una revisión del estado del arte en la aplicación de algoritmos de aprendizaje automático para la identificación y clasificación de componentes electrónicos. En la sección 4 de metodología, se describen los elementos de hardware del sistema de adquisición de imágenes, así como el procedimiento para el entrenamiento y prueba del modelo BoVW, también se describen las características de las imágenes que conforman el *dataset* de entrenamiento. En la sección 5 se presentan los resultados de la etapa de validación, comparando la mejora en la exactitud del modelo BoVW cuando se incrementa la cantidad de datos de entrenamiento. En la sección 6 de discusión, presentamos una

interpretación de nuestros resultados a partir de la comparación del conocimiento previo y la contribución de nuestro proyecto. En la sección 7 se presentan las conclusiones. Finalmente, en la sección 8 se exponen las propuestas de trabajo futuro para mejorar la predicción de componentes con empaquetado tipo DIP (*Dual in line package*).

2. BOLSA DE PALABRAS VISUALES (BOVW)

Una de las herramientas más utilizadas en la visión artificial para clasificar imágenes es el método de bolsa de palabras visuales BoVW [6]. No importa que la imagen se rote o se cambie de tamaño, esta se representa en función de sus características principales o puntos de interés, que se denominan descriptores de aspectos claves. En el método BoVW, primero se extraen los descriptores de la imagen, luego se agrupan mediante algoritmos como k-means, donde los centros de las agrupaciones se utilizan para construir los vocabularios del diccionario visual. Después se obtienen histogramas en función de la frecuencia de las palabras del vocabulario. Los histogramas conforman la bolsa de palabras visuales, que da lugar al nombre del método BoVW. Posteriormente el histograma se convierte en un vector y se asocia a una clase, como se muestra en la Figura 1. Para la clasificación de las imágenes se debe entrenar un algoritmo de aprendizaje supervisado como máquinas de vectores de soporte SVM (*Support Vector Machine*) [7]. Sin embargo, la detección de puntos de interés puede estar afectada por distorsiones en la imagen debido a sombras, texturas en el fondo y objetos irrelevantes presentes en la escena. En consecuencia, algunos de los puntos de interés detectados no aportan información útil para la tarea de clasificación y únicamente agregan ruido a los descriptores del vocabulario, afectando el rendimiento del método BoVW. Algunos algoritmos buscan evitar el ruido aplicando algún criterio para seleccionar solo los puntos relevantes, sin embargo, muchos de estos puntos de interés no informativos son

imposibles de eliminar mediante el procesamiento de las imágenes.

Algunas acciones para solucionar este problema, consisten en delimitar la región de cobertura, restringir el plano focal de la cámara para observar un solo objeto,

remover de la escena objetos extraños, evitar la formación de sombras con una iluminación adecuada, seleccionar superficies para el fondo de la imagen sin texturas o rugosidades, así como superficies poco reflejantes que no generen zonas brillantes.

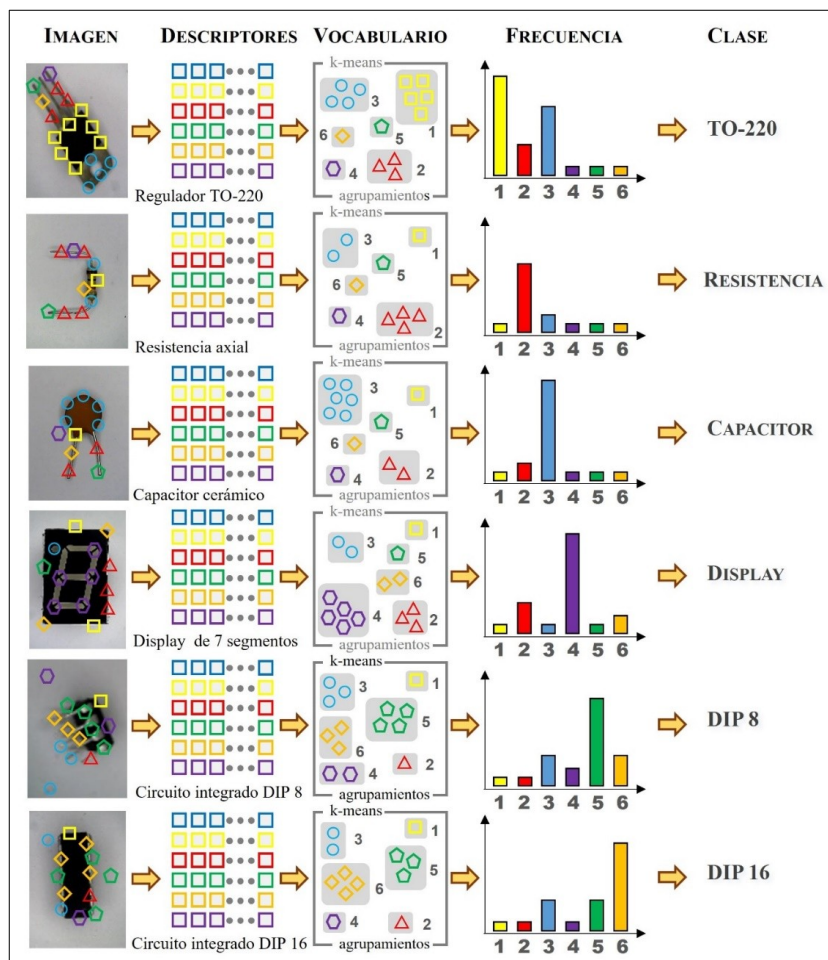


Figura 1. Proceso para el reconocimiento de componentes electrónicos etiquetados como TO-220, Resistencia, Capacitor, Display, DIP 8 y DIP 16, mediante el método BoVW. Fuente propia.

3. TRABAJOS RELACIONADOS

En esta sección se presenta una revisión en la literatura científica de trabajos relacionados sobre el uso de algoritmos utilizados para la identificación y clasificación de componentes electrónicos. En la mayoría de los casos, se utilizó la métrica *accuracy* para comparar el rendimiento de los modelos entrenados de

aprendizaje automático.

Para un sistema de visión por computadora, en [2] reportan la creación de un *dataset* con imágenes estáticas de 960x720 píxeles tomadas desde una cámara conectada a MATLAB, con el objetivo de clasificar componentes electrónicos reciclados de tarjetas reguladoras de voltaje. Las partes definidas como reutilizables fueron

capacitores, reguladores de tensión con empaquetado TO-220 y potenciómetros. En los resultados sobre el rendimiento de tres clasificadores diferentes, con SNN se logró un *accuracy* de 93.5%, con SVM del 94.9% y con CNN del 98.4%.

Huang *et al.*, usaron el algoritmo YOLO-v3, reportando altas velocidades de detección de 0.0794s y una métrica *accuracy* de 95.21% para identificar cuatro tipos de componentes electrónicos que tiene una forma similar, bajo las etiquetas de capacitores de 220 μ F, 47 μ F, 22 μ F e inductor [3].

Xu *et al.*, señalan que aplicando métodos tradicionales en la industria como SVM lograron clasificar el tipo de componente a partir de la forma de su contorno, reportan una métrica de *accuracy* de 87% para identificar resistencias, 82.6% para capacitores y 100% para inductores. Con algoritmos de aprendizaje profundo basados en CNN lograron predicciones de verdaderos positivos cercanos al 100%, pero con tiempos de procesamiento largos de entre 2.57ms y 4.85ms [4]. Usando un *dataset* de 2,666 imágenes de componentes electrónicos, distribuidos como 25% en capacitores, 45% en resistencias y 30% en diodos, Atik [5] entrenó un método de aprendizaje profundo basado en CNN, y logró un *accuracy* de 98.99%. Sin embargo, entrenar modelos con un *dataset* desbalanceado crea un sesgo en el aprendizaje que afecta a las clases minoritarias.

Chigateri *et al.*, desarrollaron una interfaz gráfica de usuario programada en MATLAB para identificar capacitores, diodos, resistencias y transistores a través de tres métodos. En el primero, identificaron el tipo de componente por la cantidad de terminales. En el segundo, identificaron el tipo de componente por la redondez de la forma del objeto, calculando la relación entre el área de la superficie y el área de un círculo cuyo diámetro es igual al diámetro

máximo inscrito en el objeto. En el tercero, calcularon la correlación de la imagen de muestra contra una imagen de referencia. En este trabajo los investigadores no utilizaron algún método moderno basado en el aprendizaje automático [8].

Simulando un entorno industrial con cámaras CCD y aplicando el algoritmo YOLO-v4 para la detección de componentes electrónicos que se desplazaban sobre una banda transportadora, en [9] calcularon un *accuracy* entre 93.94% y 98.6%. Las clases definidas por los investigadores fueron: resistencias, capacitores de mica, capacitores electrolíticos, capacitores de tantalio, conectores de pin, conectores plug, conectores USB, micro-motores, fotorresistencias, circuitos integrados, baterías, micro-interruptores, LED (*Light Emitting Diode*), resonadores de cuarzo, display de 7 segmentos, inductores, fusibles, tríodos en empaquetado TO-92 y transistores en empaquetado T0-5.

Kaya *et al.*, establecieron tres clases de entrenamiento entre capacitores, diodos y resistores [10]; obtuvieron con SVM un *accuracy* del 95.24%, con Naives Bayes del 67.01% y con Random Forest del 90.14%. En el mismo año 2022, Chand *et al.*, tomaron imágenes estáticas usando una cámara web para clasificar mediante un manipulador robótico capacitores, potenciómetros y circuitos reguladores de voltaje sobre un área de trabajo horizontal; lograron un *accuracy* del 85.6% con SNN, del 95.2% con SVM y del 98.1% con CNN [11].

Varna *et al.*, desarrollaron un sistema de bajo costo para contar cuatro tipos de componentes electrónicos de montaje superficial que se desplazaban a través de una banda transportadora [12]. Implementaron un algoritmo de aprendizaje profundo CNN en una microcomputadora Nvidia Jetson Nano. Usando un *dataset* de imágenes de capacitores, resistencias, diodos y transistores, calcularon un *accuracy* de 97.3% para la detección de capacitores, 98.6% para

resistencias, 97.2% para transistores y 88.7% para diodos.

Los métodos de aprendizaje profundo como CNN alcanzan resultados aceptables de rendimiento, pero su entrenamiento exige un alto costo computacional así como grandes cantidades de datos etiquetados, del orden de los miles de datos; por lo que demandan una inversión de tiempo y trabajo humano muy grande para completar esta tarea. Por otro lado, en la mayoría de los trabajos relacionados al reconocimiento de componentes electrónicos, se establecen clases específicas para ciertos componentes de interés para el proyecto. Para aplicaciones muy particulares, se necesita que el conjunto de datos de entrenamiento incluya las imágenes de los componentes que se desean identificar, sin embargo es probable que no estén disponibles para su uso libre, porque no se consideraron o no forman parte del *dataset*. En consecuencia, es recomendable que el usuario de este tipo de proyectos construya su propio *dataset*, según sus necesidades y los tipos de objetos que desea identificar, así como los recursos tecnológicos que tenga a su disposición [14]. En este orden de ideas, el método de aprendizaje supervisado BoVW se presenta como una alternativa viable para aplicaciones de bajo costo, con recursos limitados y para velocidades de predicción moderadas, que no requieren muchos datos en la etapa de entrenamiento, pero que logran una buena confiabilidad en la etapa de predicción.

4. METODOLOGÍA

En esta sección se describen los materiales utilizados en el sistema de adquisición de imágenes para entrenamiento del modelo BoVW implementado en MATLAB. También se describen las técnicas aplicadas para la reducción de descriptores no informativos que afectan el rendimiento del método BoVW.

La metodología propuesta en la investigación se muestra en la Figura 2. Para facilitar la adquisición y etiquetado de imágenes, el proyecto se inició con la programación en MATLAB de una interfaz gráfica de usuario GUI que consta de dos ventanas. En la primera ventana el usuario captura las imágenes del *dataset* vinculando cada imagen a una etiqueta para aprendizaje supervisado, al finalizar la captura de todos los datos, se entrena el modelo BoVW. Para comparar la eficiencia del modelo BoVW respecto al número de imágenes del *dataset*, se realizaron cuatro ensayos con 324, 378, 432 y 486 imágenes distribuidas en seis clases. En el cuarto ensayo se alcanzó la mejor métrica promedio de accuracy. El componente electrónico se colocó en todas las posibles combinaciones de orientación, posición y ubicación dentro del plano focal de la cámara, evitando usar imágenes similares en la etapa de entrenamiento para la misma clase.

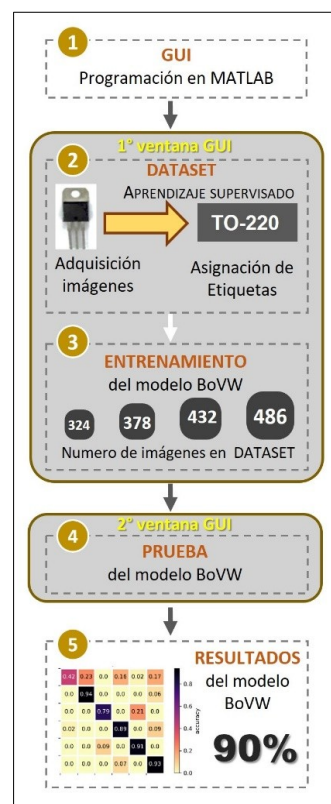


Figura 2. Metodología propuesta en el proyecto de investigación conformada por 1) la programación de la GUI, 2) adquisición de las imágenes de entrenamiento, 3) entrenamiento, 4) prueba y 5) presentación de resultados. Fuente propia.

En la segunda ventana, se prueba el modelo BoVW entrenado con imágenes distintas a las utilizadas en la etapa de entrenamiento. Cabe mencionar que tanto las imágenes usadas en la etapa de entrenamiento como de prueba o predicción fueron adquiridas directamente de fotogramas de una señal de video en tiempo real. Finalmente, los resultados de los ensayos se graficaron en matrices de confusión normalizadas para la métrica *accuracy*. Del análisis de los resultados, se observó que para cinco de las seis clases de componentes electrónicos seleccionados para reconocimiento mediante el modelo BoVW, se logra una exactitud en las predicciones cercana o superior al 90% pero con una menor cantidad de datos y costo computacional en comparación a otros algoritmos de aprendizaje automático de mayor complejidad.

En la Figura 3 se muestra el esquema del sistema de adquisición de imágenes usado en el presente trabajo. Se utilizó una cámara COM-121 de marca Steren con una resolución de 1280x720 pixeles y velocidad de 30 fotogramas por segundo, las características de esta cámara de bajo costo son genéricas pero requiere la instalación de una librería en MATLAB. Se ajustó la distancia focal de la cámara a 100 mm de la posición del componente electrónico. La cámara se conectó mediante un cable USB hacia una computadora de marca HP con procesador AMD-A9-9410, con 12 GB de RAM y unidad de procesamiento gráfico RADEON R5. Para la iluminación del componente electrónico se instaló enfrente a la cámara, una lámpara de anillo de luz LED de 60 mm de diámetro con una alimentación de 12 V de corriente directa y 7 W de potencia, la temperatura de color de la lámpara es de 6500 K con emisión de luz blanca de tipo fría. Se cuidó que la posición del anillo de luz no obstruyera el ángulo de visión de la lente de la cámara.

La presencia de ruido en las imágenes de entrenamiento es un problema que puede

afectar significativamente el rendimiento del método BoVW para hacer predicciones correctas, debido a la generación de descriptores no informativos. En el presente trabajo, este problema se disminuyó mediante la lámpara de iluminación de anillo de luz LED y la aplicación de un filtro de detección de bordes en las imágenes. También se usó un material liso en el fondo con poca reflexión de luz, para reducir la detección de descriptores no informativos asociados a la rugosidad de la superficie. Además se evitaron objetos extraños al componente electrónico en la composición de la imagen, como partículas de polvo y partículas contaminantes. Las imágenes a color se transformaron en imágenes en escala de grises mediante la instrucción `rgb2gray` de MATLAB, después se aplicó un filtro de detección de bordes mediante la instrucción `edge` para detectar los contornos en los puntos en los que el gradiente de la imagen alcanza su valor máximo, usando la aproximación de Sobel [13].

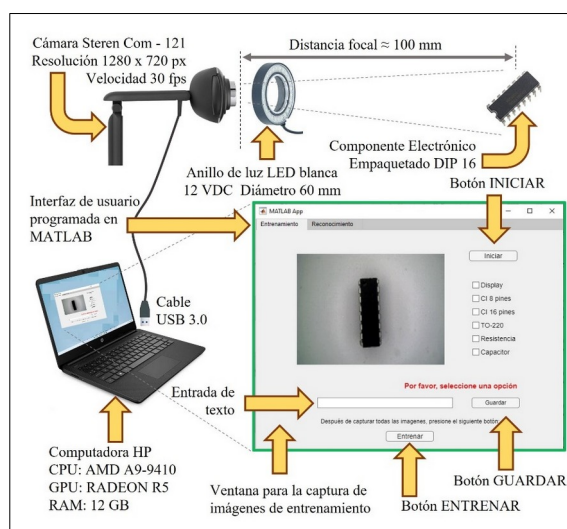


Figura 3. Esquema del sistema de adquisición de imágenes, donde se muestra la primera ventana de la GUI, para la captura de imágenes antes de la etapa de entrenamiento del modelo BoVW. Fuente propia.

Durante la construcción del sistema de adquisición de imágenes de la Figura 3, se experimentó con diferentes técnicas de iluminación para reducir la formación de sombras. En el inciso a) de la Figura 4 se

utilizó una iluminación lateral y se observa la formación de una sombra en la imagen, pero después al aplicar el filtro de detección de bordes, en el inciso b) se observa un incremento significativo de puntos irrelevantes. Por el contrario, en el inciso c) se utilizó una iluminación de anillo de luz LED, y después de aplicar el filtro de detección de bordes, se observa en la imagen del inciso d) únicamente puntos relevantes para el contorno del encapsulado, reduciendo notablemente el ruido del fondo de la imagen, que se forma alrededor del objeto. El procesamiento digital de imágenes mediante el uso de filtros, máscaras, ajuste del brillo, y conversión del formato de color a binario, es común en los sistemas de visión artificial, para eliminar elementos no deseados que puedan afectar el algoritmo [15], [16].

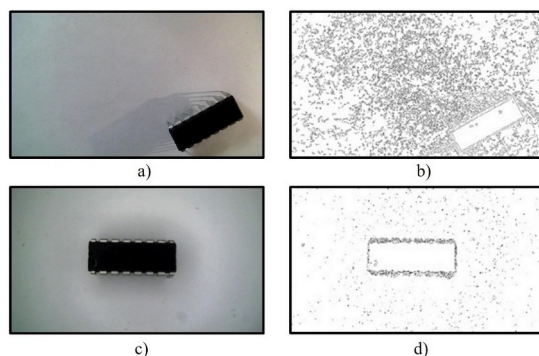


Figura 4. Imagen del componente DIP 16. a) Con iluminación lateral, b) ruido de fondo alrededor del objeto, c) con iluminación por anillo de luz LED y d) con reducción en el ruido en la imagen. Fuente propia.

Para facilitar y agilizar el reconocimiento de componentes electrónicos basado en el método BoVW, se programó en MATLAB una interfaz gráfica de usuario que consta de dos ventanas. En la primera ventana, el usuario realiza la captura de las imágenes de los componentes de acuerdo al siguiente procedimiento. Primero se oprime el botón INICIAR para activar la transmisión de video de la cámara hacia la interfaz gráfica. Una vez seleccionada la composición de la imagen deseada, respecto a la orientación, posición e iluminación del componente electrónico se captura un fotograma de la

señal de video en el recuadro de la ventana al oprimir el botón GUARDAR, previamente el usuario asocia una etiqueta, seleccionando una de las seis clases enlistadas como Display, CI 8 pines o DIP 8, CI 16 pines o DIP 16, TO-220, Resistencia y Capacitor. A continuación el usuario debe colocar el cursor en el cuadro de entrada de texto y escribir el nombre del archivo con el que se guardará la imagen en formato JPG y finalmente oprimir el botón GUARDAR. Este procedimiento se repite cambiando la posición y orientación del componente electrónico en el plano focal de la cámara, hasta completar la cantidad de imágenes deseadas para cada tipo de componente. Para borrar algún archivo guardado erróneamente, el usuario debe ir a la carpeta de almacenamiento, seleccionar el archivo y eliminarlo manualmente. Al finalizar este procedimiento, el usuario termina de construir un *dataset* balanceado.

Al oprimir el botón ENTRENAR se ejecuta la instrucción:

```
mBoVW = trainImageCategoryClassifier(DATASET,bag)
```

en MATLAB. La variable de entrada DATASET señala la ruta de la carpeta en donde se almacenaron los archivos de las imágenes etiquetadas para entrenamiento bajo aprendizaje supervisado. La entrada bag se asocia a la ejecución de la instrucción bagOfFeatures que tiene la función de extraer los descriptores de la imagen. El modelo entrenado se almacena en mBoVW.

En la segunda ventana de la interfaz gráfica que se muestra en la Figura 5, el usuario realiza el reconocimiento del componente electrónico usando el modelo previamente entrenado y guardado en mBoVW al final del procedimiento de la primera ventana. En otras palabras, en esta etapa el usuario pone a prueba el modelo BoVW con imágenes inéditas que no se usaron en la etapa de entrenamiento. El procedimiento de la segunda ventana o de reconocimiento, se describe a continuación. En el primer paso, el usuario oprime el botón INICIAR para activar la transmisión de la señal de video en tiempo

real, de la cámara hacia la interfaz gráfica de la computadora. En el segundo paso, se almacena en la memoria el fotograma que se muestra en la interfaz, en el instante en que se oprime el botón CAPTURAR y RECONOCER, como la variable `img`; después se ejecuta la instrucción:

```
[label, score] = predict(mBoVW,img)
```

donde la variable `label` es la predicción para la clase con la pérdida binaria promedio más baja del clasificador y la variable `score` es el valor de la pérdida binaria promedio más pequeña por clase que se muestra en un vector de `N` elementos, donde `N` es el número de clases. De acuerdo al valor de la variable `label`, se despliega en la interfaz gráfica el resultado como un enunciado de texto de color verde que señala, que la imagen de prueba corresponde a un Encapsulado tipo "clase". Por ejemplo, en el inciso a) de la figura 5, la imagen de prueba es un regulador con empaquetado TO-220 y el sistema hace una predicción correcta al señalar que la imagen corresponde a un encapsulado tipo TO-220.

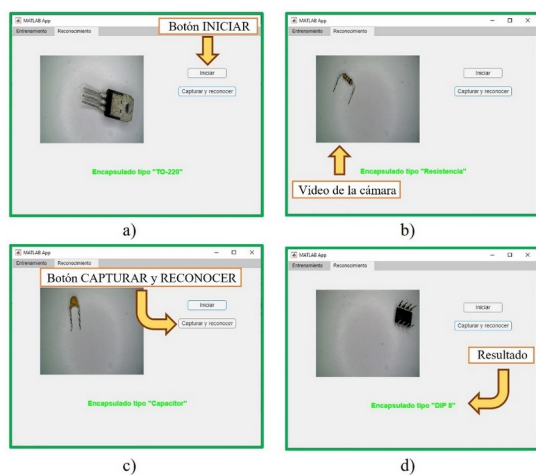


Figura 5. Segunda ventana de reconocimiento para a) predicción del componente de clase TO-220, b) predicción del componente de clase resistencia, c) predicción del componente de clase capacitor y d) predicción del componente de clase DIP 8. Fuente propia.

El sistema siempre vinculará la imagen a uno de los seis tipos de componentes

electrónicos establecidos como clases en la primera ventana. Cabe señalar, que no se incluyó una clase como no definida para un objeto distinto a las seis clases. Es decir, en caso de presentarse una falla en la predicción, la imagen de prueba real puede ser una resistencia y el sistema indicar que es un encapsulado tipo DIP 8, reconociendo erróneamente la imagen como un circuito integrado de 8 pines. Se decidió programar de esta forma la manera de presentar los resultados de predicción, para facilitar la identificación de los aciertos y errores durante la etapa de prueba. El tiempo que le toma al sistema hacer una predicción es menor a 0.5 s.

En la Figura 6 se muestra un ejemplo de las imágenes usadas en la etapa de entrenamiento del modelo BoVW, etiquetadas en seis clases distintas: DIP 16, DIP 8, Capacitor, Display, Resistencia y TO-200. Se construyó un *dataset* balanceado de 540 imágenes en total, con 90 imágenes de cada clase. El componente electrónico se colocó en todas las posibles combinaciones de orientación, posición y ubicación dentro del plano focal de la cámara, evitando usar imágenes similares y donde los componentes se mostraban parcialmente en el plano focal. En los fotogramas de la Figura 6 se observa la reflexión de la luz de la lámpara de anillo pero con pocas zonas sombreadas. No obstante, la aplicación de filtros de detección de bordes antes de la etapa de entrenamiento, contribuye a reducir el ruido y la detección de puntos descriptivos no informativos que pueden afectar la exactitud del método BoVW.

5. RESULTADOS

En la Figura 7 se observa una mejor métrica de *accuracy* para los datos de prueba, cuando se incrementó la cantidad de imágenes en la etapa de entrenamiento. La métrica promedio de *accuracy* se calculó como la suma de los resultados normalizados de la diagonal de

predicciones verdaderas positivas y negativas de la matriz de confusión, entre la cantidad de clases.

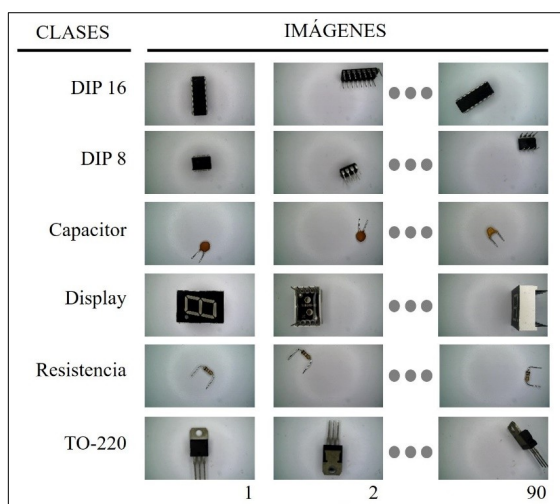


Figura 6. En el dataset de la etapa de entrenamiento del modelo BoVW, se usaron hasta 90 imágenes distintas de cada clase, cambiando la orientación, la posición y el lugar del componente electrónico en el plano focal de la cámara. Fuente propia.

Los resultados de los ensayos, muestran que el tiempo de entrenamiento para 324 imágenes fue de 3 horas y 12 minutos; calculando una métrica de *accuracy* promedio de 73.5% para 216 imágenes en la etapa de validación, ver inciso a). El tiempo de entrenamiento para 378 imágenes fue de 4 horas y 18 minutos; calculando una métrica de *accuracy* promedio de 79.5% para 162 imágenes en la etapa de validación, ver inciso b). El tiempo de entrenamiento para 432 imágenes fue de 5 horas y 48 minutos; calculando una métrica de *accuracy* promedio de 81.3% para 108 imágenes en la etapa de validación, ver inciso c). Finalmente, el tiempo de entrenamiento para 486 imágenes fue de 7 horas y 3 minutos; calculando una métrica de *accuracy* promedio de 88.3% para 54 imágenes en la etapa de validación, ver inciso d).

Con el 90% de las imágenes del dataset utilizadas en la etapa de entrenamiento, en la matriz de confusión del inciso d) de la Figura 7, se observa que los resultados de validación muestran que el 97% de las

imágenes de componentes con empaquetado DIP 8 se identificaron correctamente, así como el 94% de los componentes con empaquetado TO-220, mientras que el 93% y 92% de las imágenes de resistencias y display se identificaron sin error, respectivamente. El 89% de las imágenes de capacitores se identificaron correctamente.

Para la mayoría de los casos es una ventaja que el modelo BoVW sea capaz un reconocer un objeto independientemente de su tamaño en la imagen que se está analizando. Sin embargo, para el caso de circuitos integrados con empaquetado tipo DIP, esto representa una desventaja debido a su similitud, como muestran los resultados, cuando el 65% de los componentes con empaquetado DIP 16 se identificaron en la clase correcta. En otras palabras, el 35% de los componentes DIP 16 fueron predichos erróneamente como DIP 8, es decir, el modelo BoVW, no logró obtener una mayor cantidad de descriptores de la imagen, a partir de la mayor cantidad de terminales que distinguen al componente DIP 16 del DIP 8.

6. DISCUSIÓN

Los resultados experimentales, muestran que los sistemas de visión por computadora basados en aprendizaje automático [3], [4], [9], [12] son más eficientes para el reconocimiento de objetos que los métodos basados en la medición de parámetros geométricos y de correlación [8]. Además, en la industria se requiere que la detección de los componentes electrónicos se realice rápidamente cuando se desplazan en bandas transportadoras [3]. Sin embargo, su eficiencia disminuye para velocidades muy altas y no es igual para todos los componentes [12]. Por otro lado, trabajar con muchas clases requiere tiempos de entrenamiento muy largos y grandes cantidades de datos [9]. En consecuencia, se propone limitar el entrenamiento para detectar únicamente partes que tengan un valor económico y sean reutilizables [2].

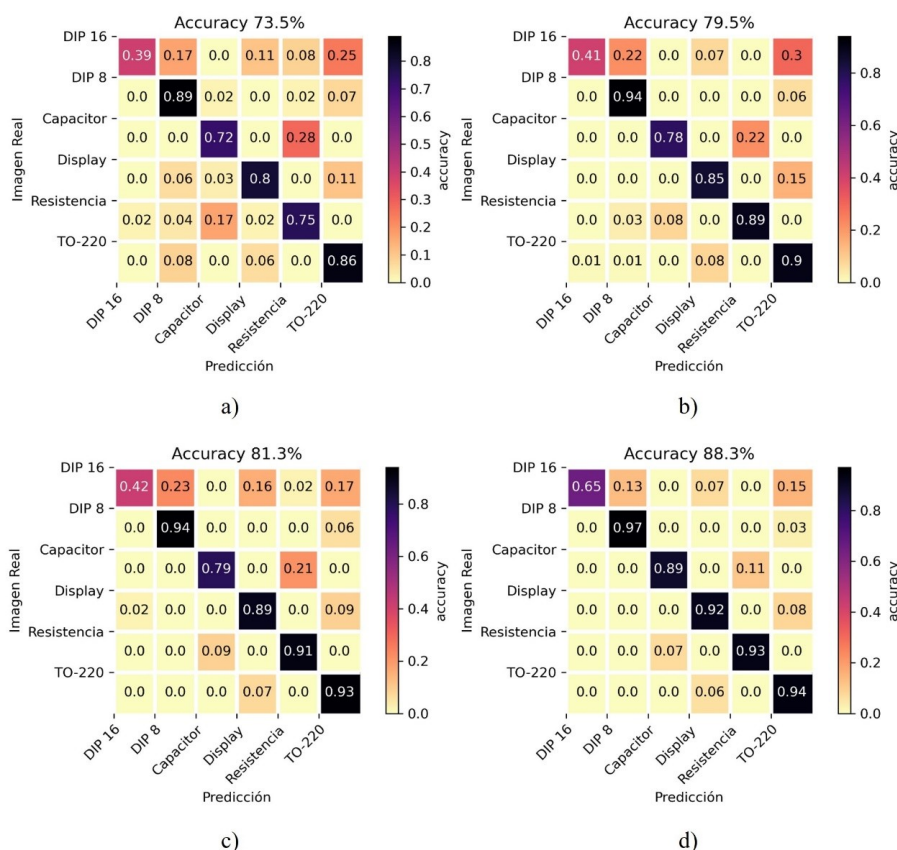


Figura 7. Matriz de confusión normalizada para datos de validación usando: a) 324 imágenes de entrenamiento, b) 378 imágenes de entrenamiento, c) 432 imágenes para entrenamiento y d) 486 imágenes de entrenamiento. Fuente propia.

En los resultados de la Figura 7, se observó que incrementar el número de imágenes para entrenamiento, mejora el rendimiento del modelo de aprendizaje, pero cantidades desiguales en los datos de cada clase, es la causa de problemas, como sesgos en el proceso de aprendizaje [5]. Por tal razón, se recomienda usar conjuntos de datos balanceados con la misma cantidad de datos por clase.

La aplicación de modelos basados en redes neuronales convolucionales CNN logran hacer predicciones con una mayor exactitud que BoVW, pero requieren equipos de cómputo de gran rendimiento y de costo elevado [2], [9], [11], [12]. Sin embargo, no todos están orientados al reciclaje de componentes electrónicos [2], [9], y también

pueden incurrir en errores de clasificación cuando la imagen de algún componente es muy diferente de las imágenes usadas en el entrenamiento.

En este proyecto se demostró que el método BoVW se puede implementar en equipos de cómputo modestos y alcanzar predicciones con un rendimiento aceptable con una métrica de *accuracy* promedio del 88.3% para imágenes de componentes electrónicos con distinta orientación, posición y ubicación en el plano focal de la cámara. No obstante, para lograr reducir la generación de desechos electrónicos, mediante procesos más eficientes de reconocimiento y clasificación, consideramos que es necesario rediseñar los productos para que sean fáciles de desensamblar.

7. CONCLUSIONES

Se programó una interfaz gráfica que consta de dos ventanas. En la primera ventana, el usuario captura las imágenes y asigna la etiqueta de la clase. El modelo BoVW se entrenó bajo aprendizaje supervisado. Las clases de componentes electrónicos definidos en este trabajo son diferentes a lo reportado en investigaciones anteriores. Cabe mencionar, que se escogieron seis clases para agilizar el proceso de construcción del *dataset*. No obstante, se pueden agregar más opciones al código fuente de la interfaz gráfica para incluir más etiquetas, según las necesidades del usuario para identificar más clases de componentes o enfocar el trabajo de clasificación únicamente para determinados componentes electrónicos que conservan algún valor económico, en procesos de reciclaje.

Después de finalizar la construcción del *dataset*, se entrena el modelo BoVW y se activa la segunda ventana de la interfaz gráfica donde se pueden hacer predicciones para imágenes inéditas, no utilizadas en la etapa de entrenamiento, que se adquieren directamente de fotogramas de la señal de video que la cámara envía en tiempo real a la interfaz. Se observó que una iluminación adecuada, disminuye las sombras en la composición de las imágenes, que afectan el rendimiento del modelo BoVW. Además el procesamiento digital de las imágenes mediante filtros también contribuye a reducir la detección de descriptores no informativos. En conjunto, la iluminación por medio de una lámpara de anillo de luz LED, la aplicación de un filtro de detección de bordes y la selección del usuario para evitar imágenes con componentes parciales en el plano focal de la cámara contribuyó a lograr altos valores en la métrica *accuracy*, cercanos o superiores al 90%, en la predicción de cinco de los seis componentes electrónicos seleccionados.

Los resultados del método BoVW muestran un rendimiento comparable a la exactitud de otros métodos más sofisticados de aprendizaje profundo para reconocer componentes eléctricos pero con una menor cantidad de datos de entrenamiento. Usando un *dataset* de 486 imágenes en la etapa de entrenamiento y con una distribución balanceada en 81 imágenes por clase, se reporta una exactitud del 97% para reconocer sin error imágenes de circuitos integrados con empaquetado DIP 8, del 94 % para componentes con empaquetado TO-220, del 93% para resistencias, del 92% para display de 7 segmentos y del 89% para capacitores. Sin embargo, se reporta una métrica *accuracy* del 65% para reconocer circuitos integrados con empaquetado DIP 16. En la etapa de validación, el método BoVW fracasó en reconocer correctamente tres imágenes de nueve etiquetadas como DIP 16, confundiendo las imágenes verdaderas con un circuito integrado DIP 8, un display y un regulador TO-220. Para algunas aplicaciones como el reciclaje de componentes, se considera que un porcentaje de efectividad inferior al 90% en el reconocimiento de objetos, no es aceptable porque implica un trabajo manual en la separación de los componentes clasificados erróneamente.

Los resultados muestran que incrementar el número de imágenes usadas en la etapa de entrenamiento mejora la exactitud en la predicción del tipo de componente electrónico pero también aumenta el tiempo de entrenamiento. Sin embargo, el tiempo que toma entrenar el modelo BoVW se puede reducir utilizando un equipo de cómputo con mejores características técnicas que el equipo utilizado en el presente trabajo. Por otro lado, los resultados muestran que el modelo BoVW logra hacer predicciones exactas con menos imágenes de entrenamiento que otros métodos más sofisticados de aprendizaje, lo cual es una ventaja que impacta en la reducción de costos del equipo de cómputo y en el esfuerzo que se debe destinar para etiquetar grandes cantidades de imágenes en métodos de aprendizaje supervisado.

8. TRABAJOS FUTUROS

Como trabajo para futuras investigaciones, se propone incrementar la exactitud del sistema de reconocimiento de componentes con empaquetados similares en su forma, como los encapsulados tipos DIP 4, DIP 6, DIP 8, DIP 10 y hasta DIP 48, colocando una marca distintiva mediante una máquina de grabado laser, antes de desensamblar los componentes de la tarjeta electrónica. Además, se pueden limitar las posiciones de los componentes tipo DIP, solo a dos vistas, desde arriba y desde abajo, evitando que el componente electrónico se coloque de costado, debido a que las terminales generan sombras y en consecuencia descriptores no informativos que afectan la exactitud del modelo BoVW. También se puede incluir un segundo criterio de clasificación, calculando el área superficial que abarca un componente DIP y comparándolo contra el área del fondo de la escena.

AGRADECIMIENTOS

Agradecemos al Tecnológico Nacional de México y a la unidad académica Zapopan del Instituto Tecnológico José Mario Molina Pasquel y Henríquez por el apoyo recibido para la divulgación de este trabajo, así como al estudiante de ingeniería electrónica Fernando Alcantara Espinoza por su valiosa contribución en la etapa de adquisición de datos del presente proyecto.

REFERENCIAS

- [1] Islam MT, Huda N, Baumber AR, Shumon R, Zaman A, Ali F, Hossain R, Sahajwalla B. A global review of consumer behavior towards e-waste and implications for the circular economy. *Journal of Cleaner Production*. 2021;316:128297. doi: [10.1016/j.jclepro.2021.128297](https://doi.org/10.1016/j.jclepro.2021.128297)
- [2] Chand P. A Low-Resolution Used Electronic Parts Image Dataset for Sorting Application. *Data*. 2023;8(1):20. doi: [10.3390/data8010020](https://doi.org/10.3390/data8010020)
- [3] Huang R, Gu J, Sun X, Hou Y, Uddin S. A Rapid Recognition Method for Electronic Components Based on the Improved YOLO-V3 Network. *Electronics*. 2019;8(8):825. doi: [10.3390/electronics8080825](https://doi.org/10.3390/electronics8080825)
- [4] Xu Y, Yang G, Luo J, He J. An Electronic Component Recognition Algorithm Based on Deep Learning with a Faster SqueezeNet. *Mathematical Problems in Engineering*. 2020;2020:2940286. doi: [10.1155/2020/2940286](https://doi.org/10.1155/2020/2940286)
- [5] Atik I. Classification of Electronic Components Based on Convolutional Neural Network Architecture. *Energies*. 2022;15(7):2347. doi: [10.3390/en15072347](https://doi.org/10.3390/en15072347)
- [6] Karim AAA, Sameer RA. Image Classification Using Bag of Visual Words (BoVW). *Al-Nahrain Journal of Science*. 2018;21(4):76-82. doi: [10.22401/ANJS.21.4.11](https://doi.org/10.22401/ANJS.21.4.11)
- [7] Cortés X, Conte D, Cardot H. A new bag of visual words encoding method for human action recognition. *Proc. of the 24th International Conference on Pattern Recognition (ICPR)*, Beijing, China. 2018;2480-2485. doi: [10.1109/ICPR.2018.8545886](https://doi.org/10.1109/ICPR.2018.8545886)
- [8] Chigateri MK, Moinuddin K. The Detection of Electrical and Electronics Components using K Nearest Neighbour (KNN) Classification Algorithm. *International Research Journal of Engineering and Technology (IRJET)*. 2016;3(5):169-175.
- [9] Guo C, Lv XI, Zhang Y, Zhang MI. Improved YOLOv4-tiny network for real-time electronic component detection. *Scientific Report*. 2021;11:22744. doi: [10.1038/s41598-021-02225-y](https://doi.org/10.1038/s41598-021-02225-y)
- [10] Kaya V, Akgül I. Classification of Electronic Circuit Elements by Machine Learning Based Methods. *Proc. of the 6th International Conference on Advances in Natural and Applied Sciences*, Ağrı, Türkiye. 2022;136-140.
- [11] Chand P, Lal S. Vision-Based Detection and Classification of Used Electronic Parts. *Sensors*. 2022;22(23):9079. doi: [10.3390/s22239079](https://doi.org/10.3390/s22239079)
- [12] Varna D, Abromavičius V. A System for a Real-Time Electronic Component Detection and Classification on a Conveyor Belt. *Applied Sciences*. 2022;12:5608. doi: [10.3390/app12115608](https://doi.org/10.3390/app12115608)
- [13] MathWorks. Encontrar los bordes de una imagen 2D en escala de grises [Consultado el 4 de julio de 2024]. Disponible en: <https://la.mathworks.com/help/images/ref/edge.html>
- [14] Rojas-Cuevas ID, Méndez-López LG, Mendoza-Vázquez JR, Ramírez-Palacios V, Torres Méndez SJ. Sistema multiplataforma de gestión de datos para visión por computadora. *Programación Matemática y Software*. 2022;14(2):60-68. doi: [10.30973/progmat/2022.14.2/7](https://doi.org/10.30973/progmat/2022.14.2/7)
- [15] Organista VD, Montero JA, Martínez MC. Extracción y selección de características en imágenes digitales mostrando lesiones de piel. *Programación Matemática y Software*. 2021;13(3):91-104. doi: [10.30973/progmat/2021.13.3/7](https://doi.org/10.30973/progmat/2021.13.3/7)
- [16] Hernández Rabadán DL, Sánchez Hernández JP. Redes neuronales convolucionales para el reconocimiento de imágenes con presencia de cenicilla polvorienta en cultivos de tomate. *Programación Matemática y Software*. 2022;14(3):21-28. doi: [10.30973/progmat/2022.14.3/3](https://doi.org/10.30973/progmat/2022.14.3/3)

ACERCA DE LOS AUTORES



Teth Azrael Cortes-Aguilar (Dr.) Recibió el grado de ingeniero en comunicaciones y electrónica por la Universidad de Guadalajara en 2003. Obtuvo el grado de maestro en ciencias en óptica por el centro de investigación CICESE, Ensenada, México en 2005. En 2022 recibió el grado de doctor en manufactura avanzada por el centro de tecnología avanzada CIATEQ. Actualmente se desempeña como profesor investigador de tiempo completo en el departamento de ingeniería electrónica, con perfil PRODEP del Tecnológico Nacional de México en la unidad académica Zapopan del Instituto Tecnológico José Mario Molina Pasquel y Henríquez, Jalisco, México. En 2024 se le otorgó el nivel de candidato en el SNII por el Consejo Nacional de Humanidades, Ciencias y Tecnologías CONAHCYT del gobierno del México.



Adriana Tovar-Arriaga (MSC.) Recibió en grado de ingeniera en sistemas computacionales por el Instituto Tecnológico de Querétaro, México, en 1999, y el grado de maestría en sistemas computacionales por el Instituto Tecnológico de Zapopan en 2016. Desde 2007 es profesora de la carrera de ingeniería en sistemas computacionales con perfil PRODEP del Tecnológico Nacional de México en la unidad académica Zapopan del Instituto Tecnológico José Mario Molina Pasquel y Henríquez, Jalisco, México. También es profesora en el departamento de sistemas computaciones en la universidad ITESO, Tlaquepaque, Jalisco, México.