

Estimación de parámetros para imágenes digitales, usando clasificadores K-NN y Tesseract



Parameter estimation for digital images, using K-NN and Tesseract classifiers

Lando Stephen Ocaña Pañora¹, Janeth Ileana Arias Guadalupe², Cristian Geovanny Merino Sánchez³ & Víctor Hugo Medina Matute⁴

Recibido: 06-10-2020 / Revisado: 27-10-2020 / Aceptado: 14-11-2020 / Publicado: 05-12-2020

Resumen.

DOI: <https://doi.org/10.33262/concienciadigital.v3i4.1.1476>

Introducción: Para este estudio, se tiene una estimación de parámetros de una comparativa entre dos algoritmos de reconocimiento de caracteres numéricos en imágenes digitales: K-NN y Tesseract, con el fin de determinar el de mayor grado de similitud. **Metodología:** Se utilizó el método inductivo y experimental para adquirir los información y datos como: la precisión, tiempo en el reconocimiento, porcentaje de consumo de memoria RAM y de CPU. Esta investigación es de tipo cuasi experimental por las técnicas escogidas para el reconocimiento de los dígitos aplicadas a imágenes y posteriormente evaluar en K-NN y Tesseract a medidores de energía eléctrica capturados en fotografía para la obtención de una lectura de consumo automática. La investigación es de tipo aplicada ya que se basó en conocimientos existentes de investigaciones previas con dirección al desarrollo tecnológico de mejora de nuevos procesos. También se puede tomar como experimental por la adquisición de datos mediante pruebas de laboratorio donde se puede apreciar elementos importantes y a simple vista una captación de fenómenos del caso. **Conclusión:** Mediante las pruebas para determinar un reconocimiento de caracteres mediante el uso de los algoritmos K-NN y Tesseract se obtuvo como resultados de estimación de precisión del 439.3% en el algoritmo K-NN y

¹ Universidad de las Fuerzas Armadas "ESPE", Latacunga, Ecuador, lsocana@espe.edu.ec, <https://orcid.org/0000-0002-4748-5282>

² Universidad Carlos III de Madrid, Madrid, España, jaariasg@pa.uc3m.es, <https://orcid.org/0000-0001-5540-1395>

³ Escuela Superior Politécnica de Chimborazo, Riobamba, Ecuador, c_merino@epoch.edu.ec, <https://orcid.org/0000-0003-3645-5165>

⁴ Universidad Técnica de Cotopaxi, Latacunga, Ecuador, victor.medina@utc.edu.ec, <https://orcid.org/0000-0002-6149-453X>

un 29.34% con Tesseract utilizando un promedio de tiempo de 1.2 y 0.06 segundos respectivos en cada algoritmo.

Palabras Claves: Algoritmo K-NN, Tesseract, Raspberry Pi, Python, OpenCV.

Summary.

Introduction: For this study, there is an estimation of parameters from a comparison between two algorithms for the recognition of numerical characters in images: K-NN and Tesseract, in order to determine the one with the highest degree of similarity.

Methodology: The inductive and experimental method was used to acquire information and data such as: precision, recognition time, percentage of consumption of RAM and CPU memory. This research is of a quasi-experimental type due to the techniques chosen for the recognition of the digits applied to images and later to evaluate in K-NN and Tesseract electric energy meters captured in photography to obtain an automatic consumption reading. The research is of an applicative type since it was based on existing knowledge from previous research aimed at technological development to improve new processes. It can also be taken as experimental by the acquisition of data through laboratory tests where important elements can be appreciated and a simple view a capture of the phenomena of the case. **Conclusion:** Through tests to determine character recognition using the K-NN and Tesseract algorithms, the precision estimation results of 439.3% were obtained with the K-NN algorithm and 29.34% with Tesseract using a time average of 1.2 and 0.06 seconds in each algorithm.

Keywords: Algorithm K-NN, Tesseract, Raspberry Pi, Python, OpenCV.

Introducción.

Actualmente, se requiere generar aplicaciones enfocadas a adquirir y obtener datos con el fin de realizar diferentes reconocimientos a nivel de caracteres numéricos en imágenes digitales de nuestro entorno. Según López Beltran & Sotter Solano (2001), la visión por computadora o artificial es una de las áreas investigativas con más desarrollo tanto el sector estudiantil como a nivel de empresa, su objetivo fundamental es proveer a los sistemas electrónicos una visión similar a la humana cuyas aplicaciones se enfocan en el reconocimiento facial, el reconocimiento de caracteres, la detección de movimiento, entre otros.

Mediante este tipo de aplicaciones se busca realizar el diseño de algoritmos, se evaluará los algoritmos K-NN y Tesseract aplicados al reconocimiento de una lectura de consumo de energía eléctrica y se medirá su nivel de precisión, tiempo empleado en el reconocimiento, porcentaje de uso de la CPU y de memoria RAM. (Nayak & Ajit Kumar, 2013)

Para la estimación de parámetros para imágenes digitales se utilizará un sistema computacional SBC con el SO Linux, la gestión de imágenes digitales se realizó con OpenCV, el Tesseract se lo utilizo como motor de reconocimiento óptico de caracteres y una tarjeta controladora Raspberry Pi en la cual se programó mediante los lenguajes Python y C++ en los que se implementará cada uno de los algoritmos. (Laganière, 2014)

Según López (2009), el diseño de un algoritmo consiste en “[...] Trazar un plan que lo resuelva por medio de pasos sucesivos y organizados en secuencia lógica.” (p. 21).

Como indican Moeslund & Granum (2001) la visión artificial o visión por computador tiene un amplio campo de aplicación en áreas tan diversas como la medicina, la agricultura, redes eléctricas inteligentes, entre otras; utilizándose en tareas como el reconocimiento de patrones, la detección de movimiento, el reconocimiento de caracteres, etc.

El Reconocimiento de Caracteres, OCR (Optical Character Recognition), es una tecnología que engloba un conjunto de técnicas basadas en estadísticas, en las formas de los caracteres, transformadas y en comparaciones, que, complementándose entre sí, se emplean para distinguir de forma automática entre los diferentes caracteres alfanuméricos existentes en una imagen digital. (Sánchez Fernández & Sadonís Consuegra, 2009)

Para las aplicaciones de reconocimiento de dígitos existen una variedad de algoritmos, el presente trabajo de investigación se centra en el desarrollo de una plataforma basada en un computador de bajo recursos que permita estudiar y evaluar los algoritmos de aprendizaje supervisado: K-NN y ANN, por su simplicidad y grado de precisión en el reconocimiento de caracteres; y contrastarlos frente a Tesseract, una aplicación Open Source de reconocimiento óptico de caracteres, inicialmente desarrollado por HP y que actualmente Google continua con su desarrollo. (Hernández Sampieri, 2010)

Al realizar la segmentación, se tiene una imagen normalizada en la que se encuentra la información susceptible de ser “reconocida”. La información así representada, una matriz bidimensional de valores binarios, niveles de gris o color RGB, no codifica de forma óptima las características más discriminativas del objeto al que representa. La extracción de las características es una de las fases más difíciles en los sistemas de reconocimiento de caracteres, puesto que es muy difícil escoger un conjunto de características óptimo. Para esta fase se utilizan entre otros los siguientes métodos: análisis de componentes principales (PCA), análisis discriminante lineal (LDA), análisis independiente de componentes (ICA) y análisis discriminante no lineal (NDA). (Sánchez Fernández & Sadonís Consuegra, 2009)

Se obtienen las características más importantes de cada uno de los caracteres incluidos en la imagen digital, el siguiente paso es realizar el reconocimiento del mismo en base a un patrón predefinido, básicamente existen dos métodos de reconocimiento: offline y online. Los métodos offline más utilizados son: Clustering, Feature Extraction, Pattern Matching y Redes Neuronales, mientras que el clasificador k-NN y el algoritmo basado en dirección son los métodos online más comunes. (Suruchi & Anjali, 2012)

La visión por computador puede considerarse como un subcampo dentro de la inteligencia artificial y se centra básicamente en intentar expresar el proceso de visión en términos de computación. Algunas de las técnicas más utilizadas en el campo de la visión por computador son el procesamiento de imágenes, los gráficos por computador y el reconocimiento de patrones. (Pajares, 2008)

Es importante destacar que el principal objetivo a alcanzar por todas las investigaciones realizadas en este campo, es intentar conseguir la máxima similitud entre la visión artificial realizada por los computadores y la visión biológica utilizada por los seres vivos.

Desarrollo.

La visión artificial o por computador está asociada a varios conceptos con relación al software, hardware y desarrollos teóricos. La tarea de visión por computador cumple varios pasos fundamentales que se verán a continuación.

La amplia gama de aplicaciones cubiertas por la Visión artificial, se debe fundamentalmente a que esta permite extraer y analizar información espectral, espacial y temporal de los distintos objetos incluidos en imágenes digitales.

La información espectral incluye frecuencia (color) e intensidad (tonos de gris). La Información espacial se refiere a aspectos como forma y posición (una, dos y tres Dimensiones). A su vez, la información temporal comprende aspectos estacionarios (presencia y/o ausencia) y dependientes del tiempo (eventos, movimientos, procesos). (Sobrado, 2003)

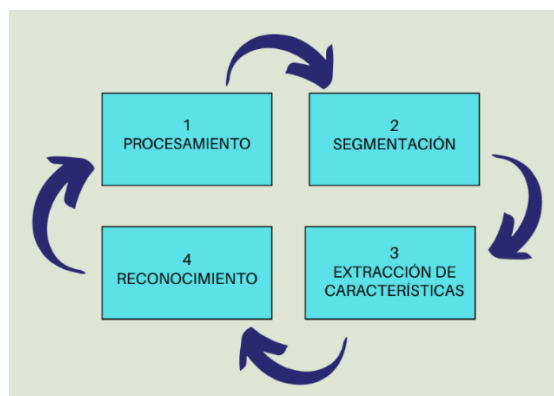
La importancia de todas las investigaciones realizadas en el campo de la visión por computador se hace evidente al observar la alta demanda por parte de la industria y la sociedad de aplicaciones capacitadas para reconocer objetos. Como ejemplo de este tipo de aplicaciones destacaríamos el caso de Fujitsu que ha diseñado robots que tienen la capacidad de reconocer imágenes en tres dimensiones. También es importante destacar el uso de estas tecnologías aplicadas en los nuevos buscadores de internet. Proyectos como Riya ó VIMA entre otros investigan la creación de buscadores de páginas web que además de poder localizar texto como puede hacer Google, permitan buscar contenidos multimedia gracias al uso de algoritmos de reconocimiento de imágenes. (Sanz, 2008)

Como las principales áreas de la visión por computador o visión artificial se tiene:

- **Medición o calibración:** correlación cuantitativa entre datos de diseño, se cumple específicamente con las mediciones preestablecidas.
- **Detección de falla:** análisis cualitativo para detección de artefactos no deseados comúnmente llamados defectos, pueden tener una forma o posición desconocida.
- **Verificación:** chequeo cualitativo de una operación de ensamblaje ha culminado correctamente.
- **Reconocimiento:** se identifica un objeto en base en descriptores asociados con el objeto.

Para llegar a cumplir una tarea de manera satisfactoria, todo el sistema de reconocimiento de caracteres debe realizar cuatro fases.

Figura 1. Proceso para reconocimiento de caracteres.



Fuente: Elaboración propia.

Las técnicas de pre procesamiento pretenden mejorar o realzar las propiedades de las imágenes para facilitar las siguientes operaciones de la Visión Artificial, tales como las etapas de segmentación, extracción de las características y finalmente la interpretación automática de las imágenes. Estas técnicas básicamente son: Realce o aumento del contraste, suavizado o eliminación del ruido presente en la imagen y la detección de bordes. (Dueñas, 2014)

Existen varios algoritmos creados para aplicaciones de reconocimiento ya sea de dígitos, figuras, etc. En esta investigación se tiene como principal objetivo la estimación de parámetros que nos permita evaluar estos algoritmos de aprendizaje supervisado como es el algoritmo K-NN ya que por su simplicidad y alto grado de precisión es aplicado en el reconocimiento de caracteres, también se estudiara el contraste con Tesseract, una aplicación OpenSource de HP y Google.

Existen placas controladoras en el mercado capaces de realizar tareas a nivel de una PC, esta es una tecnología reducida (SBC) como Arduino o Raspberry Pi. En nuestro caso se utiliza un Raspberry Pi, esta placa no se encuentra mucho en el mercado, por esta razón es de gran importancia determinar un algoritmo para reconocimiento de dígitos o caracteres, debe ser óptimo que mantengas un bajo consumo de recursos de la placa controladora Raspberry Pi y un alto grado de precisión.

Metodología.

La presente investigación es del tipo cuasi experimental ya que se escoge la técnica de reconociendo de dígitos más adecuada en imágenes para lo cual se evaluarán: K-NN, Tesseract y Redes Neuronales Artificiales en fotografías tomadas a medidores de energía eléctrica con el objetivo de obtener la lectura de consumo de manera automática, además los datos de las pruebas realizadas son generados por el autor de esta investigación.

Los instrumentos para recopilar los datos de los indicadores son los siguientes:

- **OpenCV:** librería de software libre para aplicaciones de visión artificial, dispone de interfaces de programación para C, C++, Python y JAVA.

(Comunidad OPENCV: <http://www.opencv.org>). La comunidad de OpenCV cuenta actualmente con 47000 usuarios y más de 9 millones de descarga lo que le convierte en la plataforma número uno de investigación y desarrollo en el campo de la visión artificial.

- **Tesseract:** motor libre para aplicaciones de reconocimiento óptico de Caracteres actualmente desarrollado por Google. (<https://github.com/tesseract-ocr/>)
- **Python:** lenguaje de programación orientado a objetos utilizado en una amplia gama de aplicaciones por su facilidad y optimización de código. (Comunidad Python: <https://www.python.org/>)
- **ImageMagick:** conjunto de utilidades de código abierto para mostrar, manipular y convertir imágenes, capaz de reconocer más de 200 formatos. (<http://www.imagemagick.org/script/index.php>)

Reconocimiento de caracteres mediante algoritmo K- NN.

El algoritmo K-NN (K - Nearest Neighbors) o también llamado algoritmo K – Vecinos más cercanos, pertenece a la familia de algoritmos de aprendizaje supervisado, este algoritmo se debe entrenar mediante muestras de dígitos que se necesita reconocer, se toma el valor “K” más cercano entre los datos de entrenamiento para llevar a cabo la tarea, este dato debe estar cercano al valor que se intenta reconocer para esto se toma en cuenta una distancia euclidiana existente entre dos valores.

La verificación del nivel de precisión obtenida del reconocimiento de dígitos aplicados en fotografías o imágenes tiene como población y la muestra tiene varias fotografías tomadas a medidores de energía eléctrica marca HIKING de procedencia china pertenecientes a la Empresa Eléctrica de Riobamba.

Para la obtención de la muestra se tomó 2 fotografías diarias en el horario de la mañana y en la tarde en un lapso de 25 días. En la figura 2 se puede apreciar la muestra de una imagen original de un Monofásico contador electrónico llamado medidor marca HIKING.

Figura 2. Monofásico contador electrónico (medidor)



Fuente: Elaboración propia.

Posteriormente a la selección de herramientas de software y hardware se procede al desarrollo de la aplicación de Visión Artificial la cual nos permitirá verificar el funcionamiento adecuado de ambas técnicas para el reconocimiento de dígitos o caracteres de las imágenes seleccionadas.

Antes de enviar la imagen al algoritmo se debe generar un redimensionamiento de la imagen, el tamaño original de las fotografías tomadas a los medidores HIKING de la Empresa Eléctrica de Riobamba fue en un rango de 15 cm - 30 cm, se apreció que las dimensiones de la imagen variaban de una forma notoria, así que es necesario realizar varias pruebas para determinar las dimensiones optimas, el tamaño adecuado de la imagen precisa para el procedimiento es de 453x487 pixeles Figura 3.

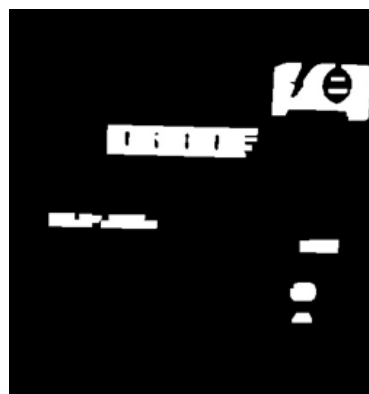
Figura 3. Imagen redimensionada para la detección de caracteres.



Fuente: Elaboración propia.

Se utilizo un método de localización de objetos que se basa en un algoritmo de detección de blobs para realizar la segmentación de la fotografía redimensionada de un medidor de energía eléctrica, Se delimitaron los objetos que se encontraban dentro de la imagen, estos límites son en forma rectangular en los cuales pueden contener áreas de interés. Así se delimito la zona de muestra donde se encuentran los dígitos del consumo de energía eléctrica de la fotografía. Este preprocesamiento utilizo las funciones de erosión de OpenCV y dilatación. Se puede apreciar en la Figura 4 el resultado de este proceso.

Figura 4. Detección de Blobs



Fuente: Elaboración propia.

En esta etapa se recorta el blob en secciones rectangulares seleccionando nuestras zonas de interés las cuales son las de los dígitos del consumo eléctrico del medidor HIKING, se tiene una posición de inicio de 144 x 170 pixeles y una posición final de 315 x 201 pixeles. En la Figura 5 se muestra el sector del blob de muestreo, se utiliza una función llamada getRectSubPix

Figura 5. Lectura obtenida del blob



Fuente: Elaboración propia.

Se puede mencionar los problemas que se presentaron al analizar las imágenes obtenidas mediante dispositivos como cámaras fotográficas con un alto nivel de compresión, de cuantización y con un sensor con alta sensibilidad del dispositivo de captura. Este problema se puede resolver con el uso de varias técnicas de fotografía, por ejemplo, el suavizado de una imagen mediante filtros o AI, en nuestro caso se utilizó un filtro llamado Filtro Gaussiano. En la figura 6 se puede apreciar la lectura del blob aplicado el Filtro Gaussiano.

Figura 6. Post procesamiento mediante filtro Gaussiano



Fuente: Elaboración propia.

Como un proceso previo a la utilización de las funciones de reconocimiento de dígitos o caracteres por medio de patrones se tiene la binarización inversa, esta técnica consiste en una detección de puntos de color blanco ubicados sobre un fondo negro. En la Figura 6 se puede apreciar que los dígitos son blancos en un fondo negro, así que es necesario realizar la técnica de binarización inversa sobre nuestra imagen de muestra seleccionada y así obtener su complemento. En la Figura 7 se puede apreciar el resultado de aplicar la binarización inversa.

Figura 7. Proceso de binarización inversa aplicada a la imagen referencial



Fuente: Elaboración propia.

Resultados.

En la estimación de parámetros para imágenes digitales, usando clasificadores K-NN y Tesseract se establecieron dos bancos de pruebas las cuales contaron con 30 muestras de fotografías tomadas a los medidores de energía eléctrica de marca HIKING

pertenecientes a la Empresa Eléctrica de Riobamba para determinar la técnica que sea más adecuada para nuestra investigación, estos bancos de pruebas se describen a continuación:

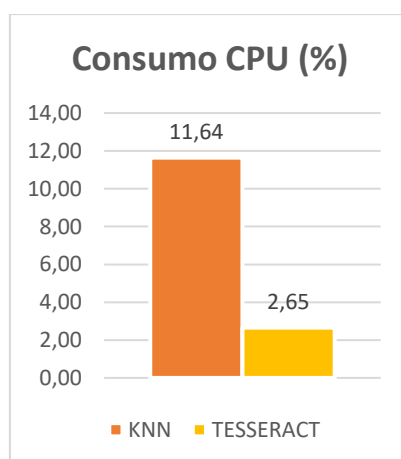
- Banco de prueba 1: Reconocimiento de caracteres con algoritmo K-NN.
- Banco de prueba 2: Reconocimiento de caracteres con Tesseract.

Los parámetros indicadores para la estimación son los siguientes:

- **Consumo de CPU:** El consumo del CPU se representa con el porcentaje de uso de CPU de la placa Raspberry Pi (SBC) utilizado para el reconocimiento de cada una de las técnicas K-NN y Tesseract.
- **Consumo de memoria RAM:** El consumo de memoria RAM se representa con el porcentaje en MB de uso de la memoria RAM integrada en la placa Raspberry Pi (SBC) utilizado para el reconocimiento de cada una de las técnicas K-NN y Tesseract.
- **Porcentaje de precisión:** Se toma en cuenta una relación entre el número de dígitos reconocidos por ambas técnicas K-NN y Tesseract sobre el número total de dígitos en la imagen de muestra.
- **Tiempo de procesamiento:** Es el tiempo en unidad de segundos que se demora cada una de las técnicas K-NN y Tesseract para culminar con el proceso de reconocimiento de dígitos o caracteres de la imagen de muestra.

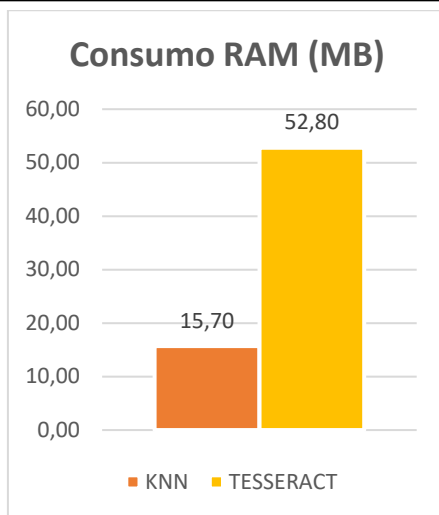
A continuación, se muestran las figuras 8, 9,10 y 11 en las cuales se puede apreciar de manera gráfica los resultados obtenidos de ambos métodos K-NN y Tesseract de las pruebas realizadas.

Figura 8. Consumo de CPU por algoritmo.



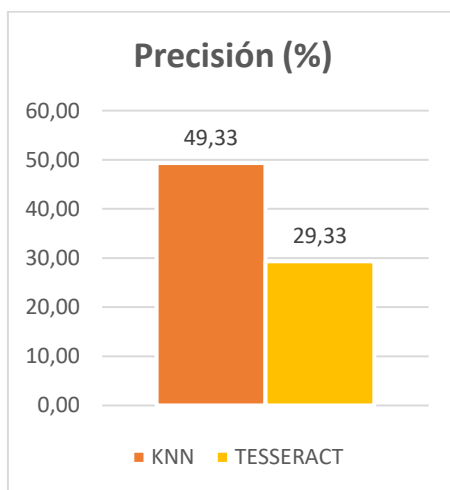
Fuente: Elaboración propia.

Figura 9. Consumo de RAM por algoritmo.



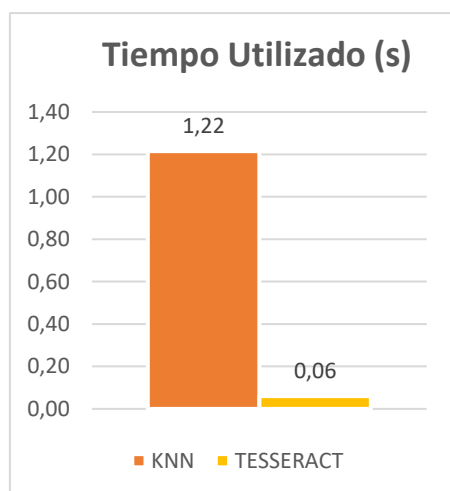
Fuente: Elaboración propia.

Figura 8. Precisión del uso de K-NN.



Fuente: Elaboración propia.

Figura 9. Tiempo de cada algoritmo.



Fuente: Elaboración propia.

Conclusiones.

- De las pruebas realizadas con los clasificadores K-NN y Tesseract, para medir el mayor grado de precisión de reconocimiento de caracteres numéricos, se utilizó una muestra de 30 fotografías de medidores de energía eléctrica, dando como resultado que el clasificador con mayor grado de precisión es K-NN con 49.34%, utilizando un tiempo promedio de 1.23s y para el clasificador Tesseract de 29.24%, utilizando un tiempo promedio de 0.061s, datos que se obtuvieron aplicando el test estadístico de Kruskal-Wallis.
- A través de la aplicación del mismo test estadístico se determinó que: K-NN presenta un consumo promedio de 15.7MB y 11.64% de consumo de CPU; para el clasificador Tesseract 52.8MB de memoria RAM y 2.65% de CPU, que son los parámetros relacionados con el consumo de recursos utilizados.

Recomendaciones:

- Tomar en cuenta realizar las capturas de imágenes en horas de la mañana y tarde donde no esté presente alguna condición climatológica que pueda afectar la precisión de la lectura de datos o el reconocimiento posterior de la zona de interés.
- Para una mejor precisión en este reconocimiento de la zona de interés como dígitos se debe incrementar la muestra de dígitos para el entrenamiento de K-NN ampliando las fuentes tipográficas variadas.
- Profundizar el tema de estudio sobre la técnica K-NN & SVM con enfoque los diferentes campos enfocados a la industria en aplicaciones de visión artificial y evaluar técnicas de reconocimiento usando patrones tales como SVM, adicionalmente con un enfoque en Deep Learning y un aprendizaje no supervisado.

Referencias Bibliográficas.

- Hernández Sampieri, R. (2010). Metodología de la Investigación. México: McGRAW-HILL.
- Laganière, R. (2014). OpenCV Computer Vision Application. Packt Publishing.
- López Beltrán, R., & Sotter Solano, E. (2001). Aplicación del sistema Robot Visión PRO para operaciones automáticas de control de calidad. Revista Científica Ingeniería y Desarrollo (9).
- López García, J. C. (2009). Educación Básica Algoritmos y programación guía para docentes. Eduteka.
- Moeslund, T., & Granum, E. (2001). A Survey of Computer Vision-Based Human Motion Capture. Computer Vision and Image Understanding.
- Nayak, M., & Ajit Kumar, N. (2013). Odia Characters Recognition by Training Tesseract OCR. International Journal of Computer Applications.

- Pajares, G. (2001). VISION POR COMPUTADOR: IMAGENES DIGITALES Y APLICACIONES. México: RA-MA.
- Sánchez Fernández, C., & Sardonís Consuegra, V. (2014). Reconocimiento Óptico de Caracteres (OCR). Universidad Carlos III.
- Sanz, J. (2008). Reconocimiento de objetos por descriptores de forma. Universidad de Barcelona.
- Sobrado, A. (2003). Sistema de Visión Artificial para el reconocimiento y manipulación de objetos utilizando un brazo Robot. Pontificia Universidad Católica del Perú.
- Suruchi, D., & Anjali, C. (2012). Survey of Methods for Character Recognition. International Journal of Engineering and Innovative Technology (IJEIT), 1.

PARA CITAR EL ARTÍCULO INDEXADO.

Ocaña Pañora, L. S., Arias Guadalupe, J. I., Merino Sánchez, C. G., & Medina Matute, V. H. (2020). Estimación de parámetros para imágenes digitales, usando clasificadores K-NN y Tesseract. *ConcienciaDigital*, 3(4.1), 103-115.

<https://doi.org/10.33262/concienciadigital.v3i4.1.1476>



El artículo que se publica es de exclusiva responsabilidad de los autores y no necesariamente reflejan el pensamiento de la **Revista Conciencia Digital**.

El artículo queda en propiedad de la revista y, por tanto, su publicación parcial y/o total en otro medio tiene que ser autorizado por el director de la **Revista Conciencia Digital**.

