

Universidad de Lima
Facultad de Ingeniería y Arquitectura
Carrera de Ingeniería de Sistemas



RECUPERACIÓN DE INFORMACIÓN DE PLACAS VEHICULARES EN BASE AL MÉTODO OPTICAL CHARACTER RECOGNITION (OCR)

Trabajo de investigación para optar el grado académico de bachiller en Ingeniería de Sistemas

Carlos Enrique Castañeda Haro

Código 20161905

Asesor

Víctor Hugo Ayma Quirita

Lima – Perú

Setiembre de 2020

Recuperación de información de Placas Vehiculares en base al método de Optical Character Recognition (OCR)

Castañeda Haro, Carlos Enrique

20161905@aloe.ulima.edu.pe

Universidad de Lima

Resumen: La suplantación vehicular es un problema basado en la adulteración de las matrículas vehiculares, las cuales han sido campo de estudio de la visión computacional, debido al alto impacto que supone el reconocimiento de placas vehiculares en temas de transporte inteligente, seguridad vial, etc. En esta investigación se presenta la implementación de un proceso de OCR para el reconocimiento de las matrículas vehiculares. Para esto, se ha establecido una base de datos con imágenes de placas vehiculares rectificadas para la realización de pruebas. El proceso consiste en un filtrado de ruido inicial a las imágenes, para posteriormente segmentar la región de la placa que contiene los caracteres, realizar un preprocesamiento a dicha imagen y aislar o segmentar a cada uno de los caracteres para aplicar el OCR de manera individual a cada uno de ellos. Los resultados nos muestran una precisión del 89.33% en el reconocimiento de matrículas vehiculares lo cual se encuentra acorde a investigaciones similares y una precisión de 96.11% en la identificación de cada carácter individualmente.

Palabras Clave: Identificación, Detección, Visión Computacional, Reconocimiento de placas vehiculares

Abstract: Vehicle impersonation is one of the main problems based on license plates adulteration. Furthermore, license plate recognition has been field of study of computer vision, due to the high impact it has on intelligent transport, road safety, etc. This work presents an OCR process based on license plate recognition. To achieve this, we have established a rectified license plate database to experiment. The process consists of an initial noise filtering to the images, to later segment the region of the plate that contains the characters, then, perform a pre-processing on the images and, finally, isolate each of the characters to apply the OCR individually to each one of them. The results show an 89.33% accuracy in the license plate recognition in accordance with similar researches, and a 96.11% accuracy in the recognition of each character individually.

Keywords: Identification, Detection, Computer Vision, License Plate Recognition

1. INTRODUCCIÓN

De manera general el problema de la inseguridad ciudadana es uno de los principales problemas del mundo; y en las últimas décadas, este fenómeno ha mostrado un incremento, generalización y diversificación que conlleva a altos impactos sociales debido a causas como la pobreza, el desempleo y la injusticia (Patriau, 2016). En este sentido, la seguridad ciudadana es una acción desarrollada por el Estado, la ciudadanía y otras entidades públicas para asegurar la convivencia y desarrollo pacífico, la erradicación de la violencia y, en general, evitar la comisión de delitos y faltas contra las personas y sus bienes (Páez et. Al., 2018).

Podemos indicar que dentro de los problemas que afectan a la seguridad ciudadana, uno de los que más resaltan es el relacionado con los delitos de suplantación vehicular. La suplantación vehicular no solo es un problema por el hecho de la adulteración de la placa del vehículo por sí misma, sino que normalmente inicia con el hurto de estas unidades, y llega hasta la falsificación de sus placas, reventa de sus autopartes, incluso al uso de estos vehículos para la realización de otros delitos (Zapata, 2017).

Según Bu, Lan y Xu (2017), en el problema de la suplantación vehicular existen dos campos claves a analizar, para corroborar si en efecto un vehículo es suplantado o no; estos campos se refieren al reconocimiento y análisis de la placa vehicular, y a la detección y clasificación del vehículo. En relación a las dos aproximaciones mencionadas para un correcto reconocimiento del vehículo, el presente artículo pretende centrarse solo en la identificación y análisis de la placa vehicular, ya que el tratamiento de placas vehiculares supone un reto mayor a la clasificación de vehículo en base a características físicas. Para ello, la visión computacional contiene los tópicos adecuados para desarrollar una herramienta que permita encontrar cierta correspondencia entre las imágenes de vehículos hacia su registro o matrícula vehicular (Bu et al., 2017).

La estructura del siguiente artículo está organizada de la siguiente manera. La sección 2 introduce el estado del arte de investigaciones relacionadas a algoritmos utilizados para el reconocimiento de objetos y caracteres, así como casos específicos relacionados a placas vehiculares. La sección 3 muestra antecedentes relacionados a técnicas utilizadas en el campo de visión computacional que son importantes para el tratamiento de imágenes y reconocimiento

de objetos y caracteres. La sección 4 indica la metodología desarrollada para realizar esta investigación, así como el detalle de la construcción de la base de datos con la que se experimentó. La sección 5 presenta los principales resultados obtenidos durante la experimentación. La sección 6 es una discusión de los principales resultados en comparación a estudios similares. La sección 7 presenta las principales conclusiones desarrolladas en base a los resultados obtenidos, así como en base al proceso de experimentación realizado. Finalmente, la sección 8 detallan algunos posibles trabajos futuros para ampliar el campo de investigación desarrollado.

2. ESTADO DEL ARTE

En esta sección se pretende hacer una revisión de las principales técnicas que se puedan utilizar en búsqueda de diferentes soluciones para el análisis e identificación de placas vehiculares. Así, en lo que respecta al reconocimiento e identificación de placas vehiculares, con el desarrollo de la visión computacional, múltiples algoritmos de reconocimiento de imágenes han aparecido (Bu et al., 2017), como el algoritmo de reconocimiento de placas basado en el descriptor SIFT (Scale Invariant Feature Transform), o el algoritmo SURF (Speed Up Robust Features) basado como mejora del algoritmo SIFT (Bay, Tuytelaars, & Van Gool, 2008), entre muchos otros algoritmos y desarrollo de tecnologías (Bu et al., 2017). También, es necesario señalar que el algoritmo SIFT (Scale Invariant Feature Transform) es considerado el más importante detector y descriptor de características y que ha sido, por ello, utilizado en múltiples aplicaciones. (Khalifa, Semary, El-Sayed, & Hadhoud, 2015)

Cuando se hace referencia al reconocimiento de imágenes se puede distinguir entre los métodos tradicionales y los basados en machine learning. Los métodos tradicionales se basan en la extracción de características específicas de la región de interés (ROI, de sus siglas en inglés) de los vehículos, de aquellas que justamente los diferencien independientemente de variantes que puedan afectar a la imagen como la iluminación, rotación, escala, etc. Tras esto, se busca emparejar el modelo de los vehículos en base a estas características. (Bu, et al., 2017). A continuación, se mencionan los principales algoritmos utilizados para el reconocimiento de objetos, los cuales están orientados en mayor medida al reconocimiento de la placa vehicular.

En primer lugar, SIFT es un algoritmo de reconocimiento de imágenes, el cual está basado en la detección de puntos clave que ofrecen características independientes de la iluminación, escala, rotación, entre otros factores que puedan alterar la imagen, tal como se muestra en diversas investigaciones (Cao, Rui y Zhang, 2012). La manera en cómo opera el algoritmo se describe según lo expuesto por Cao et al. (2012). Primero se establece, el espacio de la imagen de escala múltiple, así como la pirámide Gaussiana, se obtiene luego la Diferencia de las Gaussianas (DoG, de sus siglas en inglés) del espacio de escala múltiple, al restar las dos imágenes Gaussianas de escala adyacente, se detecta el umbral en la pirámide DoG para ubicar los puntos de características iniciales. Luego se realiza un ajuste a la superficie para obtener la localización y escala de los puntos de características precisos. Finalmente, se toma el gradiente de píxeles del conglomerado de puntos de características para que el operador SIFT tenga invarianza de rotación (Cao et al., 2012). A partir de este proceso se puede identificar y diferenciar cuatro etapas dentro de la operación del algoritmo, i) la detección del espacio escala extremo, ii) la localización de los puntos clave, iii) la orientación de la asignación, y iv) la descripción de puntos clave.

Asimismo, en razón a la identificación de la placa vehicular Zahedi & Mahdi (2010) proponen utilizar el descriptor SIFT (Scale Invariant Feature Transform) para el reconocimiento de la matrícula del vehículo. Para esto se contó con un dataset de 100 imágenes para 20 tipos de vehículos diferentes (2000 imágenes en total). Las pruebas mostraron que su método obtuvo un 84% de éxito en casos en los que las imágenes muestran aspectos pobres en iluminación, renderizado, líneas verticales en ellas, etc. Pruebas en las mismas imágenes de los vehículos, tras haber sido preprocesadas, muestran un 89% de éxito.

Para este caso, se compararon un set de características extraídas de las imágenes de pruebas con los puntos clave indicados en el set de entrenamiento de la base de datos. Se pudo observar un intento satisfactorio en identificar los puntos clave. Por otro lado, se les aplica un preprocesamiento a las imágenes ya que, al haber sido tomadas en un ambiente abierto, estas poseen pobres condiciones de iluminación, contraste, etc. A su vez, antes de realizar la extracción por medio de SIFT, se debe realizar una detección al área de la placa para evitar que se realice una relación incorrecta con otras partes del carro.

Como otra alternativa, tenemos al algoritmo Speed Up Robust Features (SURF). Se debe señalar que el descriptor SURF es un algoritmo de reconocimiento de imágenes basado en propiedades similares a las del algoritmo SIFT, con la diferencia que este es más veloz que SIFT (Bay et al., 2008).

En un esquema similar, Chaturvedi (2013) utiliza el descriptor SURF para la detección y extracción de características de la placa, en conjunto con una modificación de la red RBFN (Radial Basis Function Network) para la identificación de caracteres como alternativa al reconocimiento de matrículas vehiculares. Para esto, se desarrolló una base de datos en base a múltiples imágenes de las 26 de letras del alfabeto para la comparación de caracteres,

así como una base de datos de múltiples imágenes de vehículos y sus matrículas encontrados en Google y otras bases de datos abiertas al público para la experimentación como Medialab. En las pruebas, se logró superar el 95% en la detección de la zona de interés a través del uso de la técnica SURF y, se logró determinar el carácter con un 90% de éxito utilizando RBFN. Para ello, todos los experimentos se simularon en la plataforma de Matlab 2011.

Otro método de procesamiento convencional es el algoritmo Oriented FAST and Rotated BRIEF (ORB), que es un algoritmo de reconocimiento de imágenes que, en cierta medida, funciona de mejor manera que otros algoritmos que se encuentran en dicho campo de estudio, tales como SIFT o SURF, puesto que muestra que es afectado en menor medida por el “ruido” que pueda poseer la imagen, así como por mostrar un mejor desempeño en la identificación en tiempo real (Rublee, Rabaud, Konolige, & Bradski, 2011). Asimismo, la principal motivación para el desarrollo de ORB es mejorar las aplicaciones de coincidencia de imágenes para permitir que los dispositivos de bajo consumo de energía realicen un trabajo similar al de aquellos con alto consumo, así como también para reducir el tiempo de detección de objetos basados en características en equipos estándar (Rublee et al., 2011).

Según Dwarakanath et al. (2012), el algoritmo ORB modifica el detector FAST, que es usado para identificar los puntos clave en la imagen, añadiendo también un componente de orientación rápido y preciso, adicionalmente utiliza también al descriptor BRIEF rotado para lograr una mejor aproximación en la orientación de la imagen. En el algoritmo ORB, se lleva a cabo la detección de esquinas usando FAST y eso resulta en N puntos que se ordenan en base a la medida de Harris. Una pirámide de la imagen es construida, y los puntos clave son detectados en cada nivel de la pirámide. Se asume la intensidad de la esquina detectada para tener un desplazamiento desde su centro. Esta representación de desplazamiento, como un vector, es usada para calcular también la orientación. Las imágenes se suavizan con el parche de 31 x 31 píxeles y la orientación de cada parche de pixel se utiliza para dirigir el descriptor BRIEF para obtener la invarianza rotacional. De lo descrito previamente, se puede observar que ORB se presenta como una mejora de los algoritmos anteriormente presentados, y que a su vez permite operar con la misma calidad en cuanto al reconocimiento de objetos en equipos de bajo consumo de energía.

Finalmente, el algoritmo de Optical Character Recognition (OCR), que es utilizado para el reconocimiento de caracteres de imágenes digitales tal como señalan Lee et al. (2018). Ellos establecen el uso de OCR para determinar el contenido de información personal de manera eficiente y rápida en documentos almacenados como archivos de imágenes en computadoras personales o empresariales para prevenir el robo de estos. Realizaron pruebas con 50 archivos de imágenes y se comparó el método actual usado contra el propuesto por OCR. En estas pruebas, se pudo observar que el modelo tradicional brindaba un éxito de 90.9% en el reconocimiento, mientras que el OCR obtuvo una tasa de éxito de 94% en su mejor prueba. Asimismo, se observó que el método actual poseía un tiempo de 158.7 segundos mientras que el basado en OCR, logró esta tarea en 103.5 segundos.

También, Kakani, Gandhi y Jani (2017) utilizan la metodología que plantea el algoritmo de OCR en una red neuronal artificial de tipo feed-forward para el reconocimiento de los caracteres en una placa vehicular. Para ello, primero se localiza la región de la placa y se aplican técnicas de preprocesamiento para mejorar la imagen de la matrícula. Segundo, se procede a segmentar los caracteres de la placa vehicular de modo que solo se mantenga la información que sea importante para el reconocimiento. Finalmente, se aplica el OCR para transformar los caracteres de la imagen en información. Las pruebas se realizaron en Matlab en 300 imágenes de diferentes vehículos con diferentes ángulos de rotación, vistas, así como condiciones ambientales. Estas pruebas mostraron una precisión del 92.2% en el reconocimiento de los caracteres de la placa vehicular.

Otra investigación relacionada al uso del OCR para el reconocimiento de placas vehiculares se muestra por Jing et al. (2017). Se establece que para el reconocimiento de una placa vehicular es necesario la detección de la placa, la segmentación y extracción de características de esta y el reconocimiento de los caracteres de la placa vehicular. Este estudio se centra en el proceso del reconocimiento de caracteres, para el cual utilizaron una red neuronal artificial de tipo feed forward para la aplicación del OCR. Se realizaron experimentos en Matlab en alrededor de 665 imágenes imperfectas (imágenes borrosas o de visibilidad difícil) de caracteres en la placa de un vehículo de manera individual. De estos 665 caracteres se logró reconocer de manera acertada 653 caracteres lo cual representó un 98.2% de precisión en dicho reconocimiento.

En conclusión, debido a los altos índices de éxitos al trabajar con caracteres en formatos digitales, consideramos que sería interesante replicar el uso del algoritmo de OCR para la identificación de la placa vehicular, puesto que la literatura revisada nos sugiere eso. Para ello planteamos desarrollar experimentos en la plataforma de Matlab utilizando una base de datos de 150 imágenes de placas vehiculares rectificadas, a las que se aplicarán técnicas de preprocesamiento para mejorar la imagen, segmentación a los caracteres de la placa en sí y finalmente la aplicación de un OCR implementado en Matlab para el reconocimiento de los caracteres del vehículo.

3. ANTECEDENTES

Para entender la manera en cómo se ha trabajado respecto a la identificación de placas y vehículos, primero es necesario diferenciar los campos claves para la realización de estas tareas. En relación a la identificación de placas, debemos considerar la segmentación de la zona donde se ubica (Región de Interés) y la identificación del texto que contiene (Reconocimiento de caracteres). Por otro lado, para la identificación del modelo vehicular, será necesario considerar la segmentación del vehículo, la extracción de sus atributos y lograr la clasificación del vehículo según marca-modelo para posteriormente realizar el emparejamiento y comparación entre la placa y el modelo vehicular.

Así, a continuación, revisaremos los fundamentos teóricos que nos permitan definir de manera adecuada a los principales conceptos que serán utilizados en la presente investigación.

3.1 Pre procesamiento de la imagen

Se entiende por pre procesamiento a una mejora a la imagen, o mejor dicho como una meta de las diversas técnicas para lograr que esta se visualice mejor en algunos sentidos o características específicas de esta (Gonzales y Woods, 1996). En este sentido, debemos entender de que el pre procesamiento se muestra como una mejora netamente subjetiva de la persona que la lleva a cabo, pues se buscan resaltar con este método las características que sean consideradas como más relevantes de la imagen según se requiera (Gonzales y Woods, 1996). Entre algunos ejemplos de pre procesado de la imagen, al menos de una forma general y básica, encontramos lo que sería la ampliación o reducción de la escala, el manejo de la escala de grises de esta, la modificación del contraste, entre otros. A continuación, se explicará aquellos métodos utilizados para el preprocesamiento de la imagen que serían relevantes para el desarrollo de este trabajo.

a. Filtros de ruido en imágenes (Suavización)

Los filtros de suavizado son usados para difuminar la imagen y para la reducción del “ruido” de la imagen. Esta difuminación, es usada en los pasos de preprocesamiento como el remover pequeños detalles de la imagen para priorizar la extracción del objeto y cerrar pequeñas brechas en las líneas o curvas de la imagen. La reducción del ruido puede ser lograda a través de la difuminación a través de filtros lineales (Gonzales y Woods, 1996).

En este sentido, según Gonzales y Woods (1996), la respuesta de un filtro lineal espacial de suavizado es simplemente el promedio de los pixeles contenidos en el alrededor de la máscara filtro. La idea detrás de los filtros de suavizado es directa, se busca reemplazar el valor de cada pixel en una imagen por el promedio de los niveles grises en los alrededores definidos por el filtro de la máscara, este proceso resulta en una imagen con transiciones de “forma” reducida en los niveles grises. Asimismo, los límites son caracterizados por las transiciones de forma en los niveles grises, así que promediar los filtros tiene un efecto no deseable que difumina los límites. Otra aplicación de este proceso incluye el suavizado de los falsos contornos que resulta de usar un número insuficiente de niveles grises. Un uso mayor del promediado de filtros se encuentra en la reducción de detalles irrelevantes en una imagen. Por irrelevante, hacemos referencia a las regiones de pixeles que son pequeñas en comparación al tamaño del filtro de la máscara.

$$\frac{1}{9} \times \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array} \quad \frac{1}{16} \times \begin{array}{|c|c|c|} \hline 1 & 2 & 1 \\ \hline 2 & 4 & 2 \\ \hline 1 & 2 & 1 \\ \hline \end{array}$$

Figura 1. Ejemplos de Filtros de máscara (Gonzales y Woods, 1996).

La figura 1 muestra dos filtros de suavizado de tipo 3 x 3. El uso del primer filtro cede el promedio estándar de los pixeles debajo de la máscara. Este puede ser visto al sustituir los coeficientes de la máscara en la ecuación 1.

$$R = \frac{1}{9} \sum_{i=1}^9 Z_i \quad (1)$$

Este es el promedio de los niveles grises de los pixeles en los alrededores de la imagen 3 x 3 definida por la máscara. Asimismo, se observa que en la matriz con el coeficiente 1/9, todos los coeficientes del filtro son 1. La idea es que computacionalmente es más eficiente tener coeficiente valuados en 1. Al final del proceso de filtrado, la imagen entera es dividida por 9, Una máscara de tipo m x n hubiera normalizado esa constante a un valor de 1/(m * n). La segunda máscara mostrada en la figura 1 nos muestra cómo se cede cuando los valores en esta máscara presentan diferentes pesos en sus coeficientes, lo que significa que hay pixeles que son más “importantes” que otros en la imagen. En esta máscara. Los pixeles al centro de la máscara son multiplicados por un valor más alto que el resto, los

otros píxeles por otra parte son calculados de manera inversa como una función de distancia de ellos al centro de la máscara. La estrategia principal detrás de esta valuación es lograr un intento para reducir la difuminación en el proceso del suavizado. Asimismo, es necesario señalar que el coeficiente de $1/16$ por el que es multiplicado la segunda máscara es resultado de la suma de todos los pesos de los píxeles de la máscara. En general, es difícil ver diferencias entre las imágenes suavizada por ambos métodos de máscara o arreglos similares porque el área que estas máscaras ocupan en cualquier lugar de la imagen es muy pequeña. En referencia a esta segunda máscara, la ecuación 2 nos muestra la manera en cómo se opera cuando se posee una implementación de este tipo de filtro cuando los promedios pesados son diferentes en una imagen máscara tipo $m \times n$.

$$g(x, y) = \frac{\sum_{s=-a}^a \sum_{t=-b}^b w(s, t) f(x + s, y + t)}{\sum_{s=-a}^a \sum_{t=-b}^b w(s, t)} \quad (2)$$

El denominador en la ecuación 2 es simplemente la suma de los coeficientes de las máscaras y es una constante que necesita ser computada solo una vez. Por lo general, este factor de escala es aplicado a todos los píxeles de la imagen de salida después de que el proceso está completado.

Como se mencionó al inicio, una importante aplicación de este promediado espacial es difuminar la imagen para obtener una mejor representación de los objetos de interés, como la intensidad de objetos pequeños con el fondo u objetos grandes que se vuelven fáciles de detectar. El tamaño de la máscara establece el tamaño relativo de los objetos que serán mezclados con el fondo. Como ejemplo de cómo funciona este tipo de pre-procesado mostramos la siguiente figura 2 para que se pueda observar de manera gráfica como queda la imagen tras aplicar este suavizado y posteriormente un umbralizado.



Figura 2 Imagen en formato RGB de una placa, Imagen con suavizado e Imagen umbralizada.

b. Sharpening

Tal como menciona Gonzales y Woods (1996), el principal objetivo del sharpening es resaltar los detalles más finos en una imagen o mejorar aquellos detalles que han sido difuminados, ya sea por un error o por un efecto natural al momento de adquirir la imagen. El uso del sharpening o afinamiento de la imagen varía e incluye múltiples aplicaciones. Asimismo, es necesario señalar que este afinamiento de la imagen será logrado por medio de la diferenciación espacial de la imagen. Fundamentalmente, la fuerza de la respuesta de una derivada del operado es proporcional al grado de discontinuidad de la imagen en el punto en el que el operado es aplicado. A partir de ello, la diferenciación de la imagen mejora en límites y otras discontinuidades como el ruido.

Para este afinamiento, se considera lo que sería el comportamiento de las derivadas en áreas de constantes niveles grises y en aquellas que se encuentran al final de las discontinuidades. Este tipo de discontinuidades pueden ser usadas para modelar puntos de ruido, línea y límites en una imagen. De esta forma, las derivadas de una función están definidas en términos de diferencias. Hay varias maneras para definir estas diferenciales. Sin embargo, se necesita que cualquier definición a usar para la primera derivada debe ser cero en las áreas grises y no debe ser cero en aquella que no son áreas grises. A continuación, en la ecuación 3 se muestra la manera en cómo se definiría una derivada de primer orden en una función $f(x)$ de una dimensión.

$$\frac{\partial f}{\partial x} = f(x + 1) - f(x) \quad (3)$$

Se usa una derivada parcial aquí para mantener la notación de la misma forma como si se trabajara en una función de dos variables de tipo $f(x,y)$, la cual se usara cuando se trabaja en un espacio de dos dimensiones.

Una vez se posee este concepto, se desarrolla el concepto del operador Laplaciano, el cual es definido para una función (imagen) $f(x,y)$ de dos variables de la forma mostrada en la ecuación 4.

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \quad (4)$$

Para que esta ecuación se útil para el procesamiento de la imagen, esta ecuación necesita ser expresada en una forma discreta. De esta forma utilizamos las siguientes definiciones para para esta forma en x e y según lo mostrado en las ecuaciones 5 y 6.

$$\frac{\partial^2 f}{\partial^2 x^2} = f(x + 1, y) + f(x - 1, y) - 2f(x, y) \quad (5)$$

$$\frac{\partial^2 f}{\partial^2 y^2} = f(x, y + 1) + f(x, y - 1) - 2f(x, y) \quad (6)$$

De esta forma la ecuación laplaciana en dos dimensiones que apoyará este procesamiento está mostrado como la suma de las ecuaciones 5 y 6.

Esta ecuación puede ser implementada para el uso en la máscara de la imagen de manera que brinda un resulta isotrópico a la rotación, incrementándola en 90 grados.

Debido a que la ecuación laplaciana es un operado de derivación, se usó para resaltar las discontinuidades de los niveles grises en una imagen y enfatizar las regiones con una variación moderada en estos niveles grises. Esto tenderá a producir imágenes que tienen líneas de borde grisáceas y otras discontinuidades, todas superpuestas sobre un fondo oscuro y sin rasgos distintivos. Esto podemos representarlo en lo mostrado en la figura 3, donde observamos primero una imagen del polo norte de la luna, luego dicha imagen con un filtro laplaciano, después dicha imagen con el filtro laplaciano de imagen escalado y finalmente la imagen mejorada.

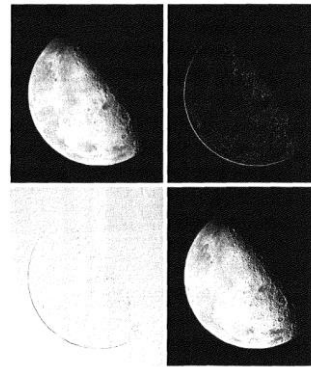


Figura 3 Imagen del Polo Norte de la luna, Imagen con filtro laplaciano, imagen escalada con filtro laplaciano e imagen mejorada usando la ecuación 5 y 6 (Gonzales y Woods, 1996).

3.2 Métodos de segmentación

La segmentación consiste en subdividir una imagen en regiones u objetos que la constituyen. El nivel al que se desea llegar a subdividir la imagen depende del problema que se vaya a resolver. Por lo general, esta subdivisión debe terminar cuando el objeto de interés en la aplicación ha llegado a ser separado del resto. En este sentido, la segmentación de imágenes no triviales es una de las tareas más difíciles en el marco del procesamiento de imágenes, ya que la precisión de la segmentación determina el fallo o logro adecuado del análisis de los procedimientos computarizados. Por esta razón, se debe tener un cuidado importante durante todo el proceso de segmentación de una imagen. A continuación, se mostrarán algunos tipos de segmentación a considerar.

a. Binarización por umbralización

Para el caso de estudio según lo expuesto por Mundaca et Al. (2016), luego que se tiene la imagen corregida, se binariza la imagen, este proceso consiste en hacer 1 ó 0 los pixeles que están en orden matricial de cierto rango en la escala de grises 0-255, las funciones que se usan para esto se basa en una estadística avanzada donde mide el gradiente de la figura en escala de grises encontrando así un valor X, la función que binariza convierte a en 1 a todos los pixeles mayores que el número X y convierte en 0 a todos los pixeles de la imagen que sea menores a ese valor X en la escala de grises, obteniendo así una imagen binarizada, es decir un matriz de 0 y 1 que facilita su posterior análisis. Luego se aplica el filtro de suavizado con la finalidad de eliminar cierto ruido que puede presentar aun dicha placa como por ejemplo pernos letras, figuras pegadas en la placa o las letras donde señala el país de procedencia de la placa, que no son necesarios en el análisis de la placa. El resultado final lo podemos observar en la figura 4, la cual nos muestra el proceso por el que pasa la placa desde su ingreso hasta su forma binarizada.



Figura 4 Imagen de la placa de automóvil binarizada (Mundaca et Al, 2016)

b. Umbralización OTSU

Según lo señalado por August y Woods (1996), se supone que el histograma de niveles grises mostrado en la figura 5, corresponde a un imagen $f(x,y)$ compuesta de objetos luminosos en un fondo oscuro, de una forma en que los pixeles del objeto y del fondo tienen niveles grises agrupados en dos modos dominantes. Una forma para extraer los objetos del fondo es seleccionar el umbral T que separa estos modos. Luego, cualquier punto (x,y) para el que $f(x,y) > T$ es llamado un punto de objeto. Caso contrario, el punto es denominado como punto de fondo. El segundo histograma mostrado en la figura 5, muestra una aproximación más general al caso, donde los tres modos dominantes caracterizan el histograma de la imagen. En esta, el multinivel del umbral clasifica un punto (x,y) como perteneciente a una clase objeto si $T_1 < f(x,y) < T_2$, o a otra clase de objeto si $f(x,y) > T_2$, y el fondo si $f(x,y) < T_1$. En general, los problemas de segmentación que requieren múltiples umbrales son los mejor resueltos usando métodos de crecimiento de regiones.

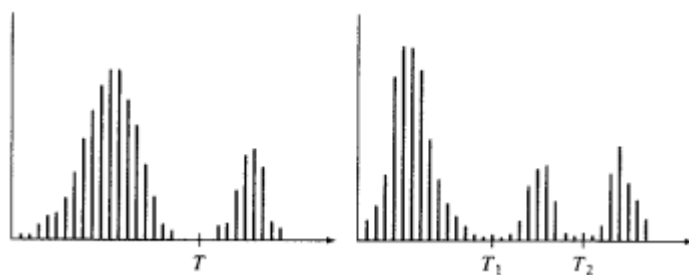


Figura 5. Histograma de escala de grises (Gonzales y Woods, 1996).

Asimismo, la umbralización debe ser vista como una operación que envuelve pruebas con una función T de la forma de la ecuación 7.

$$T = T[x, y, p(x, y), f(x, y)] \quad (7)$$

En la ecuación 7, $f(x,y)$ representa el nivel de grises del punto (x,y) y $p(x,y)$ denota algunas propiedades locales de este punto. Por ejemplo, el promedio de los niveles grises de los alrededores centrados en (x,y) . Tras esto, una imagen umbralizada de tipo $g(x,y)$ es definida según la forma de la ecuación 8.

$$g(x, y) = \begin{cases} 1 & \text{if } f(x, y) > T \\ 0 & \text{if } f(x, y) \leq T. \end{cases} \quad (8)$$

En esta ecuación 8, los pixeles etiquetados como 1 corresponden a los objetos. Por otro lado, los pixeles etiquetados con el valor de 0 corresponden al fondo. Cuando T depende solo de $f(x,y)$ y $p(x,y)$, el umbral es llamado local. Si en adición T depende de las coordenadas espaciales x y y , el umbral es llamado como dinámico o adaptativo.

3.3 Reconocimiento de objetos

Cuando se hace referencia al reconocimiento de objetos, hacemos referencia a la forma y características físicas que nos puede ofrecer el vehículo. Según lo señalado por August y Woods (1996), existen dos tipos de reconocimiento de patrones, los de decisión-teóricos y los estructurales. La primera categoría trata con patrones descritos usando descriptores cuantitativos, como la longitud, el área y la textura. La segunda categoría trata con patrones que son descritos de mejor manera por descriptores cualitativos. El valor central en este tema de reconocimiento de objetos es aprender de patrones ejemplo que se posean.

Las categorías basadas en decisiones teóricas están basadas en funciones o discriminantes de decisión. Un vector x ($x_1, x_2, x_3, \dots, x_n$) representa un patrón de dimensión n . Asimismo, para el patrón de clases w (w_1, w_2, \dots, w_n) el principal problema radica en encontrar las funciones de decisión de w , $d_1(x), d_2(x), \dots, d_w(x)$ con la propiedad de que si un patrón x pertenece a la clase w , entonces se describe la ecuación 9.

$$d_i(x) > d_j(x) \quad j = 1, 2, \dots, W; j \neq i. \quad (9)$$

En otras palabras, un patrón desconocido x está llamado a pertenecer al patrón de clase i , si el sustituto de x en las funciones de decisión $d_i(x)$ cubre el valor numérico más grande. En caso de empate estos son resueltos de manera arbitraria.

Los límites de decisión al separar las clases w_i de la w_j están representados por los valores de x para los que $d_i(x) = d_j(x)$ o equivalentemente por valores para los que x cumple la ecuación 10.

$$d_i(x) - d_j(x) = 0. \quad (10)$$

La práctica común es identificar los límites de decisión entre dos clases por la función simple representada

en la ecuación 11. En la que observamos que $d_{ij}(x) > 0$ para los patrones de la clase w_i y $d_{ij}(x) < 0$ para los patrones de la clase w_j .

$$d_{ij}(x) = d_i(x) - d_j(x) = 0 \quad (11)$$

Ahora, es necesario enfatizar en una de las técnicas de este tipo de reconocimiento de objetos como lo es la técnica de emparejamiento. Según lo señalado por August y Woods (1996), las técnicas de reconocimiento basadas en emparejamiento representan a cada clase con un prototipo de vector patrón. Un patrón desconocido es asignado a la clase más cercana en términos de métrica predefinida. El aproximamiento más simple es el clasificador de distancia mínima el cual a su vez como señala su nombre, computa la distancia euclidiana entre el patrón desconocido y el prototipo de vectores. Tras esto, elige la distancia más pequeña para la toma de la decisión.

a. Optical Character Recognition (OCR)

Hasta este punto, lo que poseemos es una imagen segmentada de manera correcta, en la cual no solo poseemos la imagen de la placa segmentada, sino también poseemos a cada uno de los caracteres dentro de estas como elementos conectados independientes con los que se pueden trabajar tal como nos muestra la figura 11.

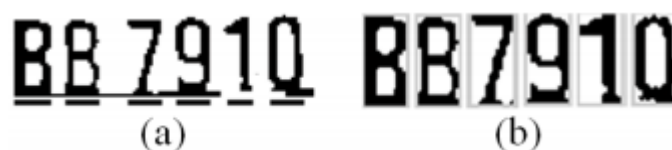


Figura 6 (a) Binarización matriz RV (b) Recorte Vertical. (Mundaca et Al, 2016)

La figura 6 (a) nos muestra como se encuentra la placa segmentada y la figura 6 (b) nos muestra como cada uno de los caracteres ya se encuentran recortados de manera correcta entre sí y pueden ser comparados a través de diferentes métricas para realmente saber que significa. A continuación, desarrollaremos la validación por medio de Optical Character Recognition (OCR), para conocer realmente que valor representa cada carácter.

Justificamos el uso del OCR, según lo dicho por Peng et al. (2013), quien señala que los sistemas OCR están diseñados para un script particular o idioma. Sin embargo, el enfoque ideal para un OCR debería ser el desarrollo de un sistema que a través del uso de datos de capacitación sea capaz de reorientar las modificaciones que puedan encontrarse en los caracteres. Esta última aproximación, refuerza la posibilidad de utilizar un OCR en el análisis de placas vehiculares pues cada una de estas imágenes de caracteres individuales pueden presentar ciertas variaciones leves que permitan diferenciarse unas de otras.

Asimismo, Mundaca et Al (2016) propone un reconocimiento óptico de caracteres – OCR. Este consiste en el reconocimiento digital de imágenes para la obtención de distintos caracteres, es decir si se tiene un texto en formato de imagen se convierte a texto en formato ASCII (código estándar para el intercambio de información)

Este método de reconocimiento se basa en plantilla que contiene los números y letras a utilizar para el reconocimiento. Esto se debe a que al tener segmentados los caracteres, se pueden comparar de manera sencilla dichas plantillas con estos caracteres segmentados. Para esto se utiliza una plantilla como la mostrada en la figura 7.

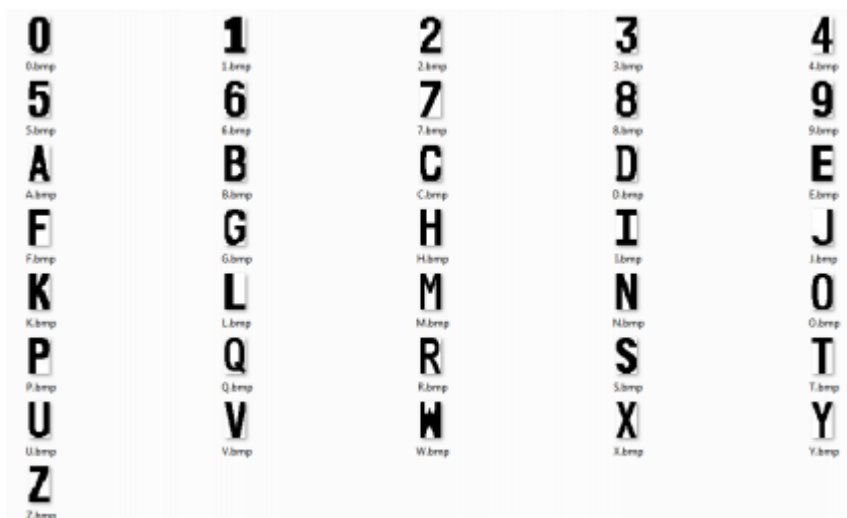


Figura 7 Plantilla de letras (Mundaca et Al, 2016)

Ahora, se compara la matriz de cada carácter obtenido de la placa con las plantillas creadas. Para esto se usó el método de correlación de Pearson que es una técnica para investigar la relación entre dos variables continuas y cualitativas. La ecuación 12 muestra la función a utilizar para la correlación de dichas matrices, donde \bar{A} y \bar{B} son los promedios de los valores de A y B respectivamente.

$$r = \frac{\sum_m \sum_n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{(\sum_m \sum_n (A_{mn} - \bar{A})^2) (\sum_m \sum_n (B_{mn} - \bar{B})^2)}} \quad (12)$$

Si el término r es igual a 1, significa que existe una relación directa entre A y B, es decir las matrices comparadas son iguales. Si el término es -1 significa que la relación es inversa. En cambio, si el resultado es 0, significa que no existe relación entre las matrices. Mientras mayor es r, significa que mayor es la igualdad de las matrices. Mientras mayor es r, significa que mayor es la igualdad de las matrices.

4. METODOLOGÍA

La presente sección pretende mostrar la manera en cómo hemos establecido el diseño de proceso del trabajo para identificar y analizar una placa vehicular. De esta manera, a continuación, se presentará el diagrama de bloque referente al proceso descrito anteriormente.

4.1 Diagrama de bloques

En la figura 8 se muestra la manera en cómo se desarrollarán cada una de las fases para realizar la identificación y análisis de la placa vehicular. Para esto, se ha considerado la literatura existente para desarrollar un proceso que esté acorde a las investigaciones relacionadas. (Gonzales y Woods, 1996) (Chaturvedi, 2013)

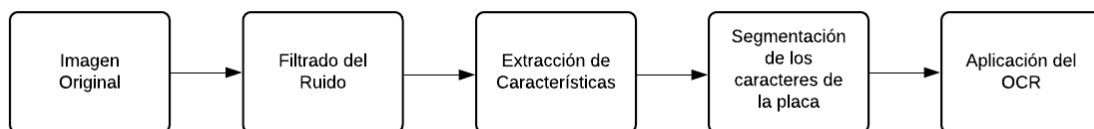


Figura 8. Diagrama de Bloques del proceso de identificación de la placa vehicular. Elaboración Propia

A continuación, se presenta una breve descripción de cada etapa del diagrama de bloques de la Figura 8, haciendo referencia a lo que sería el preprocesamiento de la imagen en forma general, llegando hasta el análisis de la placa vehicular para obtener un reconocimiento de sus caracteres a partir de la imagen en la matrícula del vehículo.

a. Filtrado del ruido

Este tipo de pre procesamiento consiste en la aplicación de los filtros adecuados para poder obtener una imagen limpia para su posterior procesamiento. Para ello se usan de filtros de suavizado de la imagen de segmentación por umbralización y filtros morfológicos. El filtro de suavizado consiste en un primer limpiado de todos aquellos pixeles que difuminen a la imagen y que causen un efecto de “blurring” que no sea propicio para la lectura y procesado de la imagen. A su vez, la umbralización de la imagen se encarga de intensificar el contraste de la imagen con su entorno para poder desde este punto realizar una pseudo segmentación en razón a las áreas que son de interés. Finalmente, filtros morfológicos sirven para obtener una imagen en una escala adecuada a lo que deseamos para que sea más sencillo identificar las regiones de interés.

b. Extracción de Características

La extracción de características de la imagen hace referencia a la detección y localización de lo que sería el área de la placa del vehículo en la que se encuentran principalmente los caracteres. De esta forma, se obtiene la región de interés de la imagen y se logra eliminar bordes indeseados en la placa que no brinden información relevante o que pudieran alterar el análisis de la imagen.

Luego se busca dejar a esta región de interés en condiciones óptimas para poder empezar segmentarla. Este procesamiento consiste en llevar la imagen a una escala de grises, así como la eliminación de distintos ruidos que puedan afectar a la imagen. Posteriormente, se aplicará un nivel de umbral adecuado a la imagen de modo que los ruidos que se hayan mantenido en la imagen y no alcancen el nivel adecuado de umbral sean eliminados. Finalmente, se procede a considerar todos aquellos elementos que permanecen dentro de la imagen como cuerpos conectados. A partir de ello, se establece un valor de mediana del área de los elementos conectados dentro de la imagen como límite a modo filtrar todos aquellos elementos conectados que se encuentren por debajo de esta mediana. En este caso, se utiliza el valor de la mediana, puesto que, al ser los caracteres de la imagen los elementos conectados con mayor representación o área en la imagen, se eliminarían los ruidos sin afectar a los caracteres de la placa.

c. Segmentación de los caracteres de la placa

La segmentación de la placa consiste en la eliminación de todos aquellos “ruidos” que se mantengan en la imagen y que realmente no sean relevantes para el análisis. Para eliminarlos, se realizan cortes a la imagen de manera que solo se mantengan en estas los caracteres de la placa. Para esto se establece un vector horizontal a partir de la mitad del alto de la imagen el cual cortará a la imagen horizontalmente y por cada pixel donde corte a la imagen almacenará su valor. Posteriormente se empezará a mover verticalmente el vector hasta lograr encontrar una posición donde los valores almacenados por este corte se asemejen a 0, lo cual representaría un corte adecuado en el cual no estamos cortando algún carácter importante. De manera similar se realiza un corte vertical, en esta ocasión utilizaremos un vector vertical que se ubicará en el inicio y final de la imagen y se moverá horizontalmente hasta encontrar la posición en la que la suma de los valores de los pixeles almacenados por la imagen se asemeje a 0 y se encuentre aproximadamente entre su posición de inicio y la posición donde existe un carácter relevante. Tras ello, se deben aislar cada uno de los caracteres que se encuentran en la imagen segmentada anteriormente. Para esto, se realizan cortes verticales pertinentes, basados en la lógica anterior, de modo que se logre tener 6 imágenes de cada uno de los caracteres que pertenecen a la placa.

d. Aplicación del OCR

Finalmente, cada una de estos caracteres aislados se comparan mediante el método de OCR (Optical Character Recognition) con una matriz que contiene las imágenes en formato PNG de todos los caracteres posibles que pueden presentarse en las imágenes, de modo que mediante esta comparación se identifique el carácter correcto correspondiente a cada una de las imágenes. Una vez se han reconocido los caracteres, se puede señalar que la concatenación de cada uno de estos nos brinda lo que sería el contenido o significativo de la placa. Asimismo, la obtención de la placa no solo se centra en este proceso, sino también en lo que representa la validación de los resultados obtenidos, ya que se podría obtener una placa que en su totalidad no representa el valor que posee en la imagen, pero de los 6 caracteres puede darse el caso que solo haya fallado en reconocer uno o más de ellos. Por lo que un análisis individual por carácter sería interesante para determinar el grado de precisión del proceso implementado.

4.2 Base de Datos

Para la construcción de la base de datos hemos obtenido imágenes de placas vehiculares a través de recortes de imágenes en páginas de internet, así como fotos que se han tomado a vehículos que nos lo han permitido. Asimismo, las imágenes de las placas han sido guardadas con el nombre que se forma de la unión de los caracteres de la placa en un formato JPEG. A continuación, en la figura 9 nos muestra cómo se encontraría una imagen rectificadas en nuestra base de datos.



Figura 9. Imagen de una placa rectificadas. Elaboración Propia

Consideramos que esta placa se encuentra rectificadas, puesto que en la imagen los caracteres se encuentran centrados y no podemos evidenciar algún tipo de orientación o inclinación de estos, así como de la imagen en general. Finalmente, tras establecer los parámetros con los que deben cumplir nuestras imágenes, es necesario señalar que se cuenta con 150 imágenes de placas vehiculares en un formato ideal para la realización de las pruebas de nuestro proceso sobre ellas. En este caso, se estimó un total de 150 imágenes, ya que es una cantidad media a las mostradas en investigaciones relacionadas al reconocimiento de caracteres utilizando OCR como la de Lee et al. (2018), se utilizaron alrededor de 50 imágenes digitalizadas para realizar el reconocimiento de caracteres o como la de Kabani et al. (2017) en la que se utilizaron alrededor de 300 imágenes de placas vehiculares.

4.3 Métricas de Evaluación

Una vez que se ha implementado el OCR, debemos proceder a la realización de pruebas para poder identificar si está funcionando de manera correcta o no. Asimismo, es necesario establecer métricas que nos permitan identificar con exactitud la precisión con la que está funcionando el programa. Por ello, se ha determinado que la mejor opción sería trabajar con una métrica basada en Overall Accuracy (OA).

Overall Accuracy es una métrica que se utiliza para la evaluación de modelos de clasificación. Exactamente, el overall accuracy hace referencia a la fracción de elementos que clasificamos de manera correcta en un modelo, y su expresión matemática es la mostrada en la ecuación 13.

$$OA = \frac{\text{Número de predicciones correctas}}{\text{Total Número de Predicciones realizadas}} \quad (13)$$

A partir de esto, entendemos que, para el caso desarrollado, sería necesario generar una estimación del Overall Accuracy tanto para las placas como para los caracteres que estos poseen.

En el caso de las placas, almacenamos todas nuestras imágenes en una matriz de manera que al leer cada una de ellas y obtener su placa, posteriormente se realice una comparación entre el nombre de la imagen y la placa obtenida por el proceso realizado, de modo que si son iguales entonces en efecto se ha identificado la placa de manera correcta, caso contrario, el proceso falló en alguna parte.

$$OA_{Placa} = \frac{\text{Placas Identificadas Correctamente}}{\text{Número Total De Placas}} \times 100\% \quad (14)$$

La ecuación 14 representa el porcentaje de precisión de placas identificadas de manera correcta respecto del total placas que han sido tomadas como muestra para la realización de las pruebas.

Por otro lado, si queremos identificar cuántos caracteres originaron una falla en la identificación de la placa, entonces sería necesario almacenar cada uno de los caracteres la imagen de la placa en una matriz diferente para luego proceder a su lectura y al compararlos con su respectiva asociación del nombre, se pueda identificar el o los caracteres que no se hayan identificado correctamente. Tal como señalábamos antes, la métrica para esta evaluación también será Overall Accuracy y podemos cuantificarla a través de la ecuación 15.

$$OA_{Caracteres} = \frac{\text{Número de Caracteres Identificados}}{\text{Número de Caracteres Totales}} \times 100\% \quad (15)$$

En esta ecuación 15, a través de la siguiente fórmula podemos conocer el porcentaje de precisión de cada uno de los caracteres que hemos identificado, sin necesidad de si la placa en efecto se reconoció de manera correcta en su totalidad. Además, el número de caracteres totales sería el número de placas que han sido procesadas por el valor de 6 que representa a la cantidad de caracteres que se presenta en cada una de las placas.

5. RESULTADOS

A continuación, se desarrolla una mención a la experimentación realizada, en la que se pueden evidenciar los resultados intermedios durante el proceso, así como algunos desafíos que se presentan en algunas placas y que pueden ser objeto de mejora en la metodología diseñada a futuro. Asimismo, se realiza un análisis cuantitativo de los resultados obtenidos en base a los métodos de evaluación de OA_{Placa} y $OA_{caracteres}$ establecidos anteriormente.

5.1 Experimentación

Para realizar el análisis de los resultados, se debe realizar una evaluación visual, donde en la fig. 10, se muestra el resultado de los diferentes procesos realizados a la imagen inicial hasta lograr que esta se encuentre en un formato adecuado para realizar su reconocimiento, donde en la figura 10 a), se observa las condiciones estándar en como ingresa una imagen en su formato original previo al procesamiento de esta en la figura 10 b), encontramos la aplicación de un primer filtro como sería la conversión de la imagen original a una que se encuentre en escala de grises. Posteriormente, la figura 10 c) nos muestra una placa tras la aplicación de un umbral a esta de modo que puedan ser observados con una mayor claridad los caracteres que se poseen en la imagen. Luego, la figura 10 d) muestra cómo actúa el proceso de binarización sobre la imagen y a la detección de todos los bounding boxes que podamos encontrar en ella para encerrar todas las áreas que podamos encontrar dentro de esta imagen. Finalmente, la figura 10 e) muestra la imagen de la placa tan solo OCR con las áreas de los caracteres a identificar, de modo que la imagen esta lista para ser validada a través del modelo OCR.

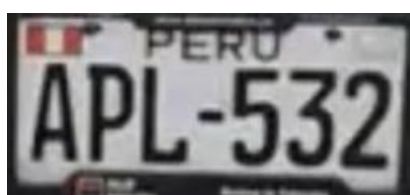




Figura 10. Proceso de Reconocimiento de Imágenes.

- a) Imagen Original
 - b) Imagen Escala de Grises
 - c) Imagen Umbralizada
 - d) Imagen Binarizada y Detección de BoundingBoxes
 - e) Imagen tras la eliminación de áreas no necesarias
- Elaboración Propia.

En resumen, en la figura anterior se observa que, en efecto, se ha realizada una secuencia de pasos sobre la imagen original a través de un preprocesamiento de esta, para luego establecer un umbral sobre esta, aplicar binarización en ella de modo que nos quedemos con los principales “ruidos” sobre ella, seamos capaces de detectar todos los ruidos posibles y finalmente terminemos con una segmentación que permita eliminar aquellos ruidos que no serán parte del análisis. Una vez se poseen los caracteres de la placa aislados, podemos identificar que estos se encuentran listos para ser procesados a través del método OCR de reconocimiento.

Sin embargo, este proceso no llega a ser exitoso en todos los casos ya que existen algunos resultados que aún presentan un reto a superar, como los presentados en la figura 11, donde se observan casos en los que el proceso realizado no es el más adecuado y no permite la identificación correcta de los caracteres por algunos motivos.

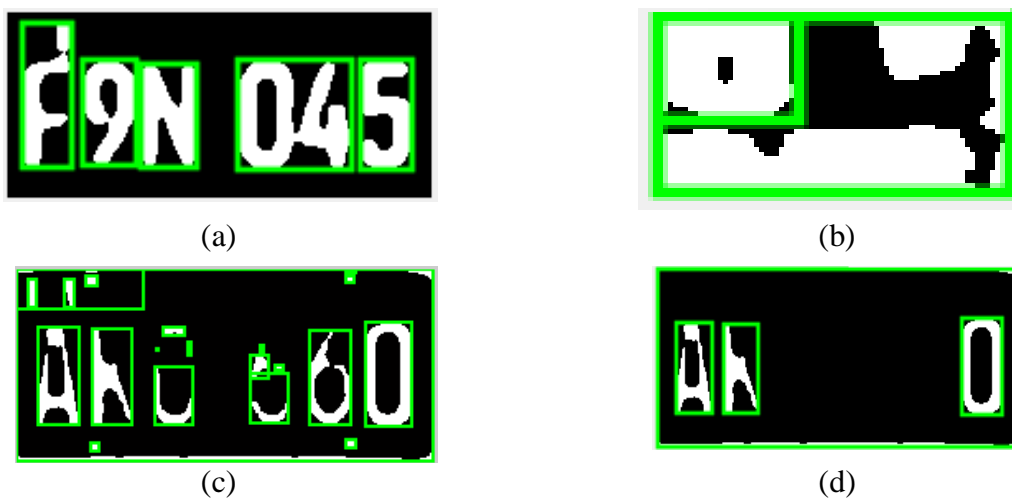


Figura 11. a) Imagen con un “ruido” conectado a un carácter principal b) Imagen con dimensiones inadecuadas tras la realización del proceso c) Desvanecimiento de ciertos caracteres tras la umbralización y binarización de la placa e) Eliminación de caracteres relevantes al estar disociados en la imagen. Elaboración Propia

En esta figura 11 a), observamos que en la parte superior del carácter representado por la letra F se encuentra un ligero ruido en la parte superior que no puede ser eliminado, puesto que está conectado al área que representa este carácter, lo cual conlleva a que este carácter sea posteriormente mal identificado y sea posteriormente reconocido como una M y no una F. Este caso es interesante de analizar, ya que tan simplemente con que un carácter esté mal identificado, la placa como tal ya no es correcta a pesar que el resto de caracteres si hayan sido identificados correctamente. En la figura 11 b), podemos observar cómo quedaría una imagen de dimensiones pequeñas si es que es procesada como si esta fuera una imagen estándar, como se puede observar, si se diera el caso entonces la imagen sería sobrepasada por la umbralización y binarización del proceso y haría irreconocible cualquier caracter. Por último, observamos la figura 11 c), en la que los caracteres de la imagen se encuentran borrosos y con zonas de estos perdidos, lo que conlleva que al analizar las áreas no relevantes se tenga una mayor cantidad de estas al eliminarlas, se estén

también eliminando partes importantes a la hora de la detección tal como lo muestra la figura 11 d) en la que tan solo se obtuvieron 3 caracteres para el posterior análisis.

5.2 Análisis Cuantitativo de Resultados

Desde el punto de vista de un análisis cuantitativo, tras la realización de pruebas se obtiene que, el resultado de precisión al obtener placas correctas de todo el proceso sobre la base de datos implementada es de 89.33%, valor que representa la precisión obtenida al identificar 134 placas de manera correcta de un total de 150 placas. Por otro lado, analizando la precisión por carácter encontramos que posee un valor de 96.11% de precisión, análisis que representa a 865 caracteres identificados de manera correcta de los 900 posibles. Esta estadística nos muestra que, si bien la placa final no se encontró de manera correcta, no significa que todos los caracteres encontrados se encuentren erróneos, sino que dicho error puede encontrarse en tan solo 1 o más. Para ello, la Tabla 1 nos muestra una tabla de frecuencias absolutas y relativas en razón a la cantidad de placas identificadas en base a los números de caracteres que fueron identificados del total de placas.

Tabla 1. Proporción de Placas según cantidad de caracteres identificados

Cantidad de Caracteres identificados correctamente por placa	Número de Placas según caracteres identificados correctamente (fi)	Número de Placas Acumuladas según caracteres identificados correctamente (Fi)	Proporción de las placas según caracteres identificados correctamente (hi)	Proporción de las placas acumuladas según caracteres identificados correctamente (Hi)
6 caracteres	134	134	89.33%	89.33%
5 caracteres	7	141	4.67%	94.00%
4 caracteres	5	146	3.33%	97.33%
3 caracteres	1	147	0.67%	98.00%
2 caracteres	0	147	0.00%	98.00%
1 carácter	3	150	2.00%	100.00%
0 caracteres	0	150	0.00%	100.00%

Elaboración Propia

La Tabla 1, nos muestra cómo se han distribuido las placas según la cantidad de caracteres identificados de manera correcta para cada uno de ellas, en las que se observa que el 89.33% logró identificar los 6 caracteres de manera correcta. Esta es una medida interesante, puesto que siempre que se logren identificar los seis caracteres, entonces podemos señalar que la identificación de la placa en su totalidad ha sido correcta. Asimismo, para el 4.67% de placas de la prueba se obtuvo el reconocimiento de 5 de sus caracteres, para un 3.33% de placas se obtuvieron 4 caracteres identificados de manera correcta, para el 0.67% de placas se obtuvo que 3 caracteres fueron identificados de manera correcta y 1 carácter fue identificado por el 2.00% de placas. Finalmente, se obtuvo 0% para la identificación de 2 y 0 caracteres por placas. Este último resultado es alentador, ya que podemos establecer de que al menos en cada una de las placas que se probaron se obtuvo como mínimo un caso de carácter identificado.

6. DISCUSIÓN

En relación al análisis de la placa vehicular, tras la realización del análisis cualitativo y cuantitativo, podemos pensar que es necesario continuar con la mejora del proceso a realizar previo a la identificación como tal de las placas puesto que aún existe un 10.67 % en el que todavía está pendiente la identificación correcta de todos sus caracteres. Sin embargo, este 10.67% no necesariamente representa una identificación incorrecta de todos los caracteres, puesto que, si un carácter no es reconocido o identificado, la identificación de la placa ya no será considerada como éxitos. Por esto, es necesario resaltar el reconocimiento e identificación de caracteres de manera individual en el que se obtuvo un 96.11% de éxito. Del 3.89% de caracteres que no fueron identificados de manera correcta, podemos identificar un caso común y es el problema de reconocimiento, puesto que se llegan a perder las imágenes de algunos de los caracteres de la placa por los motivos anteriormente explicados, lo cual nos brindaría incluso un porcentaje de éxito mayor si comparamos cuantos caracteres fueron identificados de manera correcta de aquellos que si fueron reconocidos.

Asimismo, el 89.33% de precisión al momento de identificar placas vehiculares es buen indicador, ya que la precisión obtenida es similar a la mostrada por el uso de otros algoritmos como el de SIFT propuesto por Zahedi y Mahdi (2010) en el que se obtuvo un 89% en el reconocimiento de los caracteres de la placa vehicular o el caso de Chaturvedi (2013) en el que utilizando el algoritmo de SURF logró un éxito del 90% al reconocer los caracteres de la

placa vehicular. Esto nos demuestra que en efecto el OCR tiene una precisión similar a otros descriptores utilizados para el análisis vehicular y muestra un buen desempeño frente a este problema. Además, logra una tasa similar a lo mostrado por Lee et al. (2018), quienes tras la realización de pruebas para identificación de caracteres utilizando métodos tradicionales lograron un 90.9% de precisión. Asimismo, nuestra precisión de 96.11% para los caracteres aislados es similar a la precisión mostrada por Lee et al. (2018) en la que consiguieron una tasa de éxito del 94% en su mejor prueba o los experimento realizados por Kakani et al. (2017) en los que se obtuvo un 92.2% de reconocimiento de caracteres en alrededor de 300 imágenes de placas vehiculares. Sin embargo, aun nos falta para asemejarnos al 98.2% conseguido por Jing et al. (2017), pero se explica en que ellos realizaron pruebas en caracteres ya aislados, a diferencia de esta investigación en la que se partió de una imagen de placa vehicular y en base a distintos procesos se logra obtener una imagen de carácter de placa vehicular.

7. CONCLUSIONES

En conclusión, en relación al análisis de la placa vehicular, tras la realización de las pruebas, hemos podido concluir que, si bien el número de placas analizadas correctamente no es el óptimo, no refleja que en efecto todos los caracteres en las placas no identificadas correctamente hayan sido erróneos, pues al realizar un análisis independiente por carácter, este nos brinda una mayor precisión al momento de identificarlos. Esta mayor precisión, hace referencia a que, en las placas no identificadas correctamente, se pueden identificar ciertos caracteres o no todos. Estos casos de error se pueden deber a diversos factores, como un procesamiento fallido en la imagen, la pérdida de caracteres en la imagen, dimensionamiento inadecuado de la imagen, identificación y comparación inadecuada de caracteres independientes, etc. Por estas razones, consideramos que se pueden seguir implementando otras métricas para mejorar este trabajo de procesamiento de imágenes tales como umbralización OTSU, transformaciones Bottom Hat, entre otros, de modo que podamos acaparar este 10.67% aproximado de imágenes en el que no todos los caracteres fueron identificados correctamente.

Asimismo, tal como ya señalábamos, se logró realmente un éxito del 96.11% en la identificación de caracteres que fueron procesados. Sin embargo, sería interesante establecer a futuro una métrica que nos señale realmente fueron identificados correctamente no en comparación de los 6 caracteres con los que inicia las imágenes sin procesar, sino respecto de la cantidad de caracteres que se reconocen correctamente se poseen tras la realización del procesamiento de la imagen previo a la validación por OCR.

Además, consideramos que se debe buscar la automatización de ciertos procesos tales como el establecimiento del umbral adecuado para cada imagen, puesto que actualmente se trabaja con un estándar que puede afectar el reconocimiento de los caracteres si es que no es adecuado. De la misma forma, se puede buscar un redimensionamiento de las imágenes pequeñas de modo que también puedan ser parte de este estudio a futuro.

8. TRABAJOS FUTUROS

Tal como ya se adelantaba en la sección de conclusiones, existen diversos pasos a seguir a futuro que podrían enriquecer el análisis de placa vehicular, como por ejemplo el trabajar con imágenes de placas que presenten cierto grado de rotación, de modo que se ampliaría el problema y no se trabajaría solo con imágenes centradas. De esta manera, se podrían utilizar métodos de visión computacional tales como la transformación proyectiva bidimensional para trabajar con imágenes orientadas.

Finalmente, este podría ser el principio de un sistema encargado de la verificación de posibles casos de suplantación vehicular partiendo de la base ya construida en relación al análisis de la placa vehicular y, añadiendo otros aspectos claves en este tema como el reconocimiento de la forma o modelo del vehículo.

REFERENCIAS

- Bay, H., Tuytelaars, T., & Gool, L. Van Gool. (2008). *SURF: Speeded-up robust features*, 1-14.
- Bu, Q., Lan, S., & Xu, P. (2017). *A CNN based car model recognition improvement*. (1), 86-89. <https://doi.org/10.1145/3094243.3094258>
- Cao, P., Rui, T., Zhang, J., & Zhou, Y. (2012). *An improved SIFT matching algorithm based on locality preserving projectioLPP*. 192. <https://doi.org/10.1145/2382336.2382391>
- Carrión, F. (2005). *La seguridad en América Latina y Europa*. *Quorum*, 12, 21-28.
- Chaturvedi A. (2013). *Automatic License Plate Recognition System using SURF Features and RBF Neural Network* 37-41
- Dwarakanath, D., Eichhorn, A., Halvorsen, P., & Griwodz, C. (2013). *Evaluating performance of feature extraction*

- methods for practical 3D imaging systems*. 250–255. <https://doi.org/10.1145/2425836.2425887>
- Gonzalez R., Woods R., (1996) *Digital Image Processing 2nd Edition*
- Jing Y., Youssefi, B., Mirhassani, M., Muscedere, R. (2017). *An efficient FPGA implementation of Optical Character Recognition for License Plate Recognition*. 2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE). doi:10.1109/ccece.2017.7946734
- Kakani, B., Gandhi, D., Jani, S. (2017). *Improved OCR based automatic vehicle number plate recognition using features trained neural network*. 2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT). doi:10.1109/icccnt.2017.8203916
- Khalifa A., Semary A., El-Sayed, M., Hadhoud, M. (2015) *A comparison of local detectors and descriptors for Multi-Object Applications*
- Lee, Y., Song, J., Won, Y. (2018). *Improving personal information detection using OCR feature recognition rate*. 1941-1952. <https://doi.org/10.1007/s11227-018-2444-0>
- Mundaca G., (2016) *Detección de Caracteres de placas de automóviles mediante técnicas de visión artificial*
- Paez C. Murillo, Peón I. Escalante, Ramirez Yesid. *Contexto de la seguridad ciudadana en América Latina y el Caribe: revisión de literatura (2007-2017)* 83-106. <http://dx.doi.org/10.21830/19006586.360>
- Patriau, C. (2016) *Implementación de helicópteros para la Seguridad Ciudadana en Lima Metropolitana; problemática y soluciones (2014-2015)*. Pontificia Universidad Católica del Perú (PUCP), Lima, Perú.
- Peng, X., Cao, H., Setlur, S, Govindaraju, V., Natarajan, P. (2013) *Multilingual OCR Research and Applications: An Overview* <https://doi.org/10.1145/2505377.2509977>
- Ruble E, Rabaud V., Konolige K., Bradski G. (2010) *ORB: an efficient alternative to SIFT or SURF*
- Zahedi M., Mahdi S. (2010) *License Plate Recognition System Based on SIFT Features* 998-1002 <https://doi.org/10.1016/j.procs.2010.12.164>
- Zapata, J. (2017). *Implantación de una norma jurídica e implementación de canales virtuales para enfrentar el problema de alta incidencia del delito contra la fe pública (falsedad genérica en la modalidad de suplantación de identidad vehicular – “clonación” de vehículos) en la transferencia de vehículos automotores por compraventa con placas duplicadas en la ciudad de Lima*. Pontificia Universidad Católica del Perú (PUCP), Lima, Perú.