



# **Reconocimiento Automático de Matrículas de Automóviles Particulares Mexicanos con Información del Color**

*que para obtener el grado de*  
Maestro en Ciencias de la Computación

*Presenta*

Ing. Nicolás Trejo de la Cruz

*Tutor Académico:*

Dr. en C. Comp. Farid García Lamont

Tutores Adjuntos:

Dr. en C. Jair Cervantes Canales

Mtra. en C. Yedid Erandini Niño Membrillo

Texcoco, Estado de México, diciembre del 2016





UAEM | Universidad Autónoma del Estado de México

CUTex

DICTÁMEN DE AUTORIZACIÓN Y OBTENCIÓN DE GRADO DE MAESTRÍA

Texcoco, Méx., a 4 de octubre del 2016



Título del proyecto:

Reconocimiento Automático de Matrículas de Automóviles Particulares Mexicanos con Información del Color

Tesista:

Nicolás Trejo de la Cruz

Dictamen:

No. de revisión: 1

- Rechazado
- Sujeto a modificaciones
- Aceptado, condicionado
- Aceptado

**POSGRADO  
TITULACION  
RECIBIDO**  
Por *Adriana Arevalo*  
Texcoco, Méx., a 4 de 10 del 2016

Observaciones generales:

Acceptado para la impresión

Acceptado para la defensa de grado

| Tutor Adjunto                     | Tutor Académico                    | Tutor Adjunto                              |
|-----------------------------------|------------------------------------|--|
|                                   |                                    |  |
| Dra. en C. Jair Cervantes Canales | Dr. en C. Com. Farid García Lamont | M. en C.C.A. Yedid Erandini Niño Membrillo |





# Resumen

---

El reconocimiento de las matrículas de vehículos ha sido investigado ampliamente en todo el mundo: Argentina [8], Bangladesh [28], China [33], Egipto [14], India [19], Japón [4], Malasia [2], entre otros. Normalmente los trabajos relacionados con ésta finalidad consisten en tres fases: 1) Localización de la placa dentro de la imagen, 2) extracción de los caracteres y 3) clasificación o reconocimiento de los caracteres. En el caso de México, el obstáculo principal se encuentra en la fase de extracción de caracteres, porque los algoritmos que existen en la literatura asumen que la placa no tiene patrones de textura en el fondo de la misma, debido a que en otros países por lo general el fondo es blanco y los caracteres negros. Sin embargo, en el caso de las placas mexicanas éstas tienen patrones de textura en el fondo produciendo que los algoritmos que funcionan bien para las placas sin patrones de textura no siempre funcionan correctamente. Por otro lado, cada entidad federativa y cada nuevo gobierno estatal puede diseñar su propio patrón de textura de fondo, esto implica que puedan existir por lo menos 32 clases de placas, número que se incrementa con los cambios en la administración gubernamental.

Es importante mencionar que, si bien cada entidad federativa puede diseñar el fondo de sus placas, las dimensiones de las placas y letras así como su estilo deben cumplir con las características que señala la Norma Oficial Mexicana **NOM-001-SCT-2-2000**; estas características son las que se emplean para reconocer la matrícula. De aquí que se propone crear un algoritmo que segmente los caracteres y los reconozca en función de sus características de color y de forma.

Para segmentar los caracteres de forma adecuada, se eliminó la mayor cantidad de los patrones de textura del fondo mediante el uso de un factor umbral, con el que se separaron los colores oscuros, que forman las letras; de los claros, que forman el fondo. Una vez filtrada la textura de fondo,

la imagen de la placa se binarizó y se obtuvieron los histogramas horizontal y vertical mediante la técnica de proyección de perfiles, con la finalidad de obtener las coordenadas de posición que se utilizaron para segmentar los caracteres.

Ya obtenidas las imágenes de los caracteres, se procedió a modelarlos y caracterizarlos mediante las técnicas de: momentos de Hu, Descriptores de Fourier y el Factor de Correlación Cruzada. Los datos obtenidos en esta etapa, se emplearon como alimentación en la etapa de clasificación.

Finalmente, en la etapa de clasificación, se utilizaron las técnicas de Plantillas, Clasificador Bayesiano y Redes Neuronales Artificiales. Los resultados obtenidos se discuten al final del trabajo.

---

# Agradecimientos

---

Agradezco al Dr. Farid García Lamont por su invaluable apoyo durante todos mis estudios de maestría, a mis asesores: Dr. Jair Cerantes Canales y M.en C. Yedid Erandini Niño Membrillo por ayudarme a mejorar mi trabajo a través de sus muy acertadas observaciones.

Al CONACyT por la beca que me otorgó, bajo el número de registro 561707.

Al Dr. José Ruiz-Shulcloper, al Dr. Eduardo Garea Llano, al M.en C. Reynolds León Guerra y a todos los miembros del equipo del Centro de Aplicaciones de Tecnologías de Avanzada (CENATAV), por haber fungido como anfitriones en la estancia académica realizada del 07 de junio al 10 de julio del 2015, en la Habana, Cuba.





## **Dedicado a...**

---

A mi padre amado, Yave...

Al León de la Tierra de Judá, Jesús...

Al Espíritu Santo..., que jamás me han abandonado.

A mis padres, Felipe y Florina, que me dieron vida y me formaron.

A mi hermosa esposa, Gisela y mis preciosas hijas, Mariel y Marlen.

A mis amigos, compañeros y a todos aquellos que siempre depositaron su confianza en mi...



# Contenido

---

|  |           |
|--|-----------|
| <b>Índice de Figuras</b>   | <b>IX</b> |
| <b>Índice de Tablas</b>  | <b>XI</b> |
| <b>1. Introducción</b>   | <b>2</b>  |
| 1.1. Planteamiento de problema . . . . .                         | 2         |
| 1.2. Justificación . . . . .                                     | 5         |
| 1.3. Objetivo General . . . . .                                  | 6         |
| 1.4. Objetivos Específicos . . . . .                             | 6         |
| 1.5. Hipótesis . . . . .   | 6         |
| <b>2. Antecedentes y trabajos previos</b>                        | <b>8</b>  |
| 2.1. Segmentación de caracteres . . . . .                        | 9         |
| 2.2. Reconocimiento de caracteres . . . . .                      | 10        |
| 2.3. Reconocimiento Automático de Matrículas en México . . . . . | 13        |
| <b>3. Propuesta de reconocimiento de matrículas de vehículos</b> | <b>14</b> |
| 3.1. Normalización y binarización de la placa . . . . .          | 14        |
| 3.2. Segmentación de caracteres . . . . .                        | 16        |
| 3.3. Modelado de los caracteres . . . . .                        | 17        |
| 3.3.1. Factor de Correlación . . . . .                           | 18        |
| 3.3.2. Momentos de Hu . . . . .                                  | 18        |

---

|  |           |
|--|-----------|
| 3.3.3. Descriptores de Fourier . . . . .                           | 19        |
| 3.4. Reconocimiento de los caracteres . . . . .                    | 21        |
| 3.4.1. Comparación de plantillas . . . . .                         | 21        |
| 3.4.2. Clasificador Bayesiano . . . . .                            | 22        |
| 3.4.3. Red Neuronal Artificial . . . . .                           | 22        |
| 3.5. Análisis de componentes principales . . . . .                 | 23        |
| <b>4. Experimentos, resultados y discusión</b>                     | <b>26</b> |
| 4.1. Experimentos . . . . .  | 26        |
| 4.1.1. Entrenamiento del reconocedor de letras y números . . . . . | 26        |
| 4.1.2. Reconocimiento de matrículas . . . . .                      | 31        |
| 4.1.3. Discusión . . . . .   | 31        |
| <b>5. Conclusiones</b>   | <b>34</b> |
| <b>Referencias</b>   | <b>36</b> |
| <b>A. Placas de circulación actual en México</b>                   | <b>42</b> |
| <b>B. Artículo</b>   | <b>44</b> |

---

# Índice de Figuras

---

|   |    |
|---|----|
| 1.1. Etapas de un sistema RAMV . . . . .  | 3  |
| 1.2. Ejemplos de placas de diferentes países. . . . .   | 4  |
| 1.3. Placas de algunos estados de México. . . . .   | 5  |
| 3.1. Posición de los caracteres dentro de la matrícula. . . . .   | 15 |
| 3.2. Espacio RGB y umbral de binarización. . . . .  | 16 |
| 3.3. Segmentación de caracteres. . . . .  | 16 |
| 3.4. Proyección horizontal y vertical de una imagen binarizada. . . . .   | 17 |
| 3.5. Proyección horizontal y vertical de una imagen binarizada. . . . .   | 17 |
| 3.6. Conjunto de puntos que forman el contorno de un caracter . . . . .   | 20 |
| 3.7. Similaridades para la imagen del caracter <b>G</b> . . . . .   | 21 |
| 4.1. <i>Resultados de clasificación de placas para los diferentes reconocedores implemen-</i><br><i>tados</i> . . . . . | 32 |



# Índice de Tablas

---

|   |    |
|---|----|
| 2.1. <i>Ventajas y desventajas de los métodos de extracción de caracteres</i> . . . . .   | 11 |
| 3.1. <i>Ejemplo de eliminación de secciones superiores e inferiores; (a) imagen original, (b) imagen obtenida tras eliminación de secciones</i> . . . . . | 15 |
| 3.2. <i>Región binarizada de los caracteres de un par de placas</i> . . . . .   | 16 |
| 4.1. Número de ejemplos empleados en el entrenamiento . . . . .   | 27 |
| 4.2. Resultados de clasificación de números con 8 características . . . . .   | 27 |
| 4.3. Resultados de clasificación de letras con 8 características . . . . .  | 28 |
| 4.4. Resultados de la clasificación de números utilizando el clasificador Bayesiano con descriptores de Fourier . . . . .                                 | 29 |
| 4.5. Resultados de la clasificación de letras utilizando el clasificador Bayesiano con descriptores de Fourier . . . . .                                  | 29 |
| 4.6. Resultados de clasificación de números utilizando el clasificador Bayesiano con Análisis de Componentes Principales . . . . .                        | 30 |
| 4.7. Resultados de clasificación de letras utilizando el clasificador Bayesiano con Análisis de Componentes Principales . . . . .                         | 30 |
| 4.8. Resumen de resultados generales para el clasificador Bayesiano . . . . .   | 30 |
| 4.9. Número de ejemplos de matrículas por estado . . . . .  | 31 |
| A.1. Placas de la República Mexicana . . . . .  | 42 |





# Introducción

---

## 1.1. Planteamiento de problema

El sistema Reconocedor Automático de Matrículas Vehiculares (RAMV) que extrae el número de matrícula de una imagen dada, puede ser compuesto por cuatro etapas. La Figura. 1.1 muestra las etapas de un RAMV.

1. Adquirir la imagen del automóvil utilizando una cámara.
2. Extraer la placa de la imagen.
3. Segmentar y extraer los caracteres de la placa.
4. Reconocer los caracteres extraídos.

Dentro de la segmentación y extracción de los caracteres de la placa, existen varios trabajos que abordan ésta tarea. Sin embargo, todos consideran o asumen que el fondo de las placas es liso o sin patrones de textura, por ejemplo, en la Figura. 1.2 se muestran imágenes de placas de algunos países latinoamericanos y europeos. Nótese que no tienen textura de fondo, sin embargo, en el caso de las placas de México ésto no es así.

En la Figura. 1.3 puede observarse que a pesar de ser el mismo país, las placas mexicanas tienen diferentes colores y patrones de textura en el fondo, ya que cada entidad federativa y cada nueva administración gubernamental puede diseñar sus placas, ello produce que existan más de 32



Figura 1.1: Etapas de un sistema RAMV

patrones de fondo, lo que genera un problema, debido a que con los algoritmos actuales no siempre es posible extraer los caracteres con éxito.

En el caso de éste proyecto, sólo se abordan las fases de extracción y reconocimiento de caracteres (etapas 3 y 4). Las principales dificultades que conllevan estas tareas son:

- Las placas pueden ser de diferentes tamaños, debido a la distancia de la cámara y al factor de acercamiento.
- Los colores de los caracteres y del fondo pueden ser oscuros de tal forma que al convertirlos a escala de grises no se pueda apreciar la diferencia entre carácter y fondo de la placa.
- Las imágenes de las placas pueden no ser totalmente claras debido a la suciedad que contengan.
- Las imágenes de las placas pueden ser adquiridas con un cierto ángulo de inclinación.
- Las imágenes adquiridas pueden tener diferentes tipos de iluminación.



Figura 1.2: Ejemplos de placas de diferentes países.

Para acotar el problema, las imágenes con que se trabajará; deberán cumplir las siguientes condiciones:

- Las fotografías serán previamente adquiridas de forma estática.
- Serán capturadas a una distancia entre 1 y 1.5 m.
- Se trabajará con la imagen de la placa previamente extraída de la fotografía original; por lo que, en el presente trabajo, no se abordará esta etapa de un sistema de RAMV
- Las imágenes de las placas no tendrán variaciones de iluminación
- Las fotografías tendrán ángulos de rotación muy pequeños (imágenes frontales o casi frontales)
- Sin oclusiones.



Figura 1.3: Placas de algunos estados de México.

## 1.2. Justificación

Las razones por la que se identifican los vehículos por su matrícula son varias, por ejemplo: control de tráfico, búsqueda de vehículos robados y/o sospechosos, registro estadístico del comportamiento vehicular, entre otros. Dada la cada vez mayor cantidad de automóviles que circulan en las carreteras y ciudades, es difícil o imposible que se destinen seres humanos para realizar eficientemente esta tarea las 24 horas del día. Lo anterior se debe a que cuando las personas realizan una misma actividad por un tiempo prolongado, con el paso del tiempo empiezan a sufrir algunos malestares, principalmente del tipo psicológico; por otro lado, el hecho de mantenerse atento frente a cámaras de video a una distancia relativamente corta y por un gran período de tiempo, genera problemas de visión. De aquí la necesidad de desarrollar sistemas inteligentes que de forma automática reconozcan las placas de los vehículos para alguno de los propósitos que se han mencionado anteriormente.

### 1.3. Objetivo General

Diseñar un sistema basado en visión artificial para el reconocimiento de placas mexicanas de vehículos particulares extrayendo características de color de baja intensidad y de forma.

### 1.4. Objetivos Específicos

- Desarrollar un algoritmo para la segmentación y extracción de caracteres (con base en sus características de color)
- Diseñar un algoritmo para el reconocimiento de caracteres (caracterización del carácter y desarrollo del clasificador)
- Realizar pruebas para validación de la propuesta

### 1.5. Hipótesis

Dado que las placas, independientemente de la entidad federativa a la que pertenezcan deben cumplir en su diseño con las características que señala la norma **NOM-001-SCT-2-2000** [1]. Los caracteres deben tener color obscuro o negro; por lo tanto, si de la imagen sólo se extraen los píxeles negros o de baja intensidad, entonces se estarán extrayendo los caracteres de la placa capturada.

---



# Antecedentes y trabajos previos

---

El Reconocimiento Automático de Matrículas es la extracción de la información de la placa de un vehículo de una imagen o una secuencia de imágenes. La información extraída puede ser usada, en muchas aplicaciones, entre las que se encuentran: sistemas de pago electrónico, pago de peaje, cobro por estacionamiento y monitoreo para la vigilancia del tráfico en autopistas y caminos arteriales. El RAMV, también es conocido como Identificación Automática de Vehículos, Reconocimiento de Placas Vehiculares, Reconocimiento Automático de Número de Placa y Reconocimiento Óptico de Caracteres para Vehículos [5].

Las razones por las que se desea identificar los vehículos por su matrícula pueden ser varias, por ejemplo: control de tráfico, detección de vehículos robados y registro estadístico del comportamiento de los conductores, por mencionar algunas.

De las cuatro fases de un sistema de RAMV, para el caso particular de México; las que plantean un mayor esfuerzo son las últimas dos, debido a que detrás de los caracteres existe un diseño colorido por estado, lo cual produce que existan por lo menos 32 diferentes fondos para las placas mexicanas; a diferencia de las placas de otros países que por lo general tienen un solo color de fondo; por lo anterior, el presente documento se centrará en analizar las etapas 3) Segmentación y Extracción y 4) Reconocimiento de caracteres. En las secciones siguientes se revisarán algunos trabajos que abordan éstas etapas.

## 2.1. Segmentación de caracteres

Una vez extraído el contorno de la placa de la imagen original, el siguiente paso es segmentarla para extraer solamente la parte necesaria que permita obtener el número de matrícula, tal parte, es el contorno que representa cada carácter. La tarea de extracción de los caracteres, es abordada mediante diferentes métodos, enlistados a continuación [5]:

- **Usando conectividad de píxeles:** Esta técnica consiste en etiquetar las regiones que forman los píxeles conectados en una imagen binaria de la placa. Dichas regiones son analizadas y aquellas que tienen igual proporción y tamaño son consideradas caracteres de la placa. Cuando los caracteres están juntos o rotos, éste enfoque falla.
  - **Usando proyección de perfiles:** Cuando las placas contienen colores de fondo, el uso de imágenes binarias, resulta inadecuado, debido a los valores que pueden tomar los píxeles de color, por ello se emplea el método de proyección de perfiles que utiliza el contraste en los colores de fondo y los caracteres de la imagen, algunos utilizan la imagen binaria y otros la información de color. Algunos métodos proyectan la imagen de manera vertical para determinar las posiciones iniciales y finales de los caracteres, para posteriormente proyectarlos de manera horizontal para extraer carácter por carácter.
  - **Usando conocimiento previo de caracteres:** En este enfoque, la imagen de la placa es explorada de forma horizontal, con la finalidad de identificar el inicio y el final de la región que contiene los caracteres que conforman la matrícula, lo anterior para determinar los puntos que se deben considerar al segmentar la parte que contiene cada símbolo; para ello, en algunos casos, las imágenes son redimensionadas a un valor determinado. El inconveniente de ésta técnica es que si existen variaciones en la placa, se extrae parte del fondo en lugar de los caracteres.
  - **Usando contorno de caracteres:** Este sistema trabaja en dos pasos: 1) La localización aproximada de cada carácter es encontrada mediante una técnica rápida de avance, combinada
-



con una función gradiente-dependencia y curvatura-dependencia. 2) Los límites exactos son obtenidos mediante un método especial de avance.

- **Usando combinación de características:** Para mejorar la segmentación de la placa, principalmente cuando ésta está seriamente dañada o incompleta, dos o más características de los caracteres, pueden ser utilizadas; algunas técnicas de este tipo, según su eficiencia son:
  - 1) Segmentación basada en Morfología Adaptativa, en los cuales un algoritmo basado en el histograma, detecta los fragmentos y posteriormente los mezcla para formar los caracteres.
  - 2) Algoritmo de engrosamiento morfológico, que se emplea para localizar líneas de referencia para separar caracteres sobrepuestos.
  - 3) Algoritmo de adelgazamiento morfológico, que junto con el cálculo del costo de segmentación, determina la línea principal para segmentar caracteres conectados. En ([26] se propone un método de segmentación placas indias, para la obtención de caracteres, que intenta superar la problemática de ruidos e iluminación heterogénea, considerando imágenes tomadas entre 1.5 y 2.5 metros, aplicándole un filtro Gaussiano, una transformación similar a la Laplaciana y un filtro medio. Una vez pre procesada la imagen, se emplean características propias de las letras Indias y otras relativas para segmentar la imagen, extraer sub imágenes (que contienen cada carácter) y posteriormente binarizarlas.

En la Tabla 2.1, se enlistan las ventajas y desventajas de los métodos de extracción de caracteres existentes enlistados anteriormente.

## 2.2. Reconocimiento de caracteres

Una vez segmentada la placa, los caracteres extraídos, son entonces reconocidos y la salida es el número de placa. Sin embargo, el reconocimiento de caracteres en los sistemas de RAMV tiene algunas dificultades, debido a:

- El factor de acercamiento de la cámara, los caracteres extraídos no tienen el mismo tamaño ni el mismo espesor.
-

Tabla 2.1: *Ventajas y desventajas de los métodos de extracción de caracteres*

| MÉTODOS                           | VENTAJAS  | DESVANTAJAS   |
|-----------------------------------|---|---|
| Conectividad de píxeles           | Simpleza y sencillez, robusto a la rotación de la placa                           | Falla en la extracción de todos los caracteres cuando los caracteres están unidos o incompletos.              |
| Proyección de Perfiles            | Independiente de la posición del carácter, capaz de tratar con algunas rotaciones | El ruido afecta el valor de la proyección, requiere conocimiento previo del número de caracteres de la placa. |
| Conocimiento previo de caracteres | Simple  | Limitado por el conocimiento previo, un cambio puede producir errores.  |
| Contorno de caracteres            | Puede obtener los límites exactos del carácter                                    | Es lento y puede producir contornos distorsionados o incompletos  |
| Características Combinadas        | Más confiable   | Computacionalmente complejo.  |

- Los caracteres extraídos pueden contener ruido o estar incompletos.
- Los caracteres también pueden estar rotados.

Los métodos de reconocimiento de caracteres más comunes son los siguientes:

- **Datos brutos de los caracteres:** La comparación de caracteres con plantillas mide la similitud entre un carácter y las plantillas. La plantilla más similar al carácter es clasificada como el objetivo. La comparación de plantillas es utilizado para reconocer caracteres con:
  - a) Fuente simple.
  - b) No rotados
  - c) No incompletos
  - d) Del mismo tamaño
- **Características extraídas:** Debido a que no todos los píxeles del carácter tienen la misma importancia en la distinción del mismo, una método que obtiene algunas características del carácter es una alternativa que complementa la técnica de comparación simple. Ello reduce el tiempo de procesamiento, debido a que no todos los píxeles son involucrados. También es

capaz de superar los problemas de comparación con plantillas si las características son suficientemente fuertes para distinguir los caracteres con alguna distorsión. Las características extraídas forman un vector que es comparado con otros vectores de características, previamente almacenados, para medir la similitud. En [5] el vector de características es generado mediante la transformada de Hotelling de cada carácter.

- **Métodos estadísticos y estructurales.** En [22] se propone un método de reconocimiento de caracteres que combina métodos estadísticos y estructurales para reconocer los caracteres. En esta técnica, primero se utilizan cuatro sub-clasificadores estadísticos independientes para reconocer el carácter de entrada; los resultados obtenidos del reconocimiento se combinan utilizando el método de Bayes. Si la salida de esta primera etapa contiene caracteres que pertenecen a los conjuntos de similitud prescritas, las imágenes son procesadas y clasificadas una vez más con el método de reconocimiento de la estructura para obtener sus características estructurales. Los datos obtenidos en la etapa anterior, se alimentan a un clasificador por árbol de decisión. Finalmente, se emplea un algoritmo genético para lograr parámetros óptimos del sistema, cuyo objetivo es lograr la optimización de parámetros adecuados para lograr una adecuada El propósito es lograr la optimización de parámetros adecuados para lograr un buen rendimiento en el reconocimiento de los caracteres.
  - **Técnicas combinadas:** Otra técnica de reconocimiento de caracteres, es propuesta en [35] en la que utilizan características combinadas para realizar tal fin; tales como la Transformada de Hough y el conocimiento previo de la segmentación horizontal y vertical. Este método utiliza la información de la intensidad y evita los cortes y el aglutinamiento de caracteres que es la desventaja de la binarización de imágenes. En [26] se propone un método reconocimiento de caracteres mediante el uso de imágenes binarizadas que son normalizadas de forma local mediante un filtro Gaussiano, para homogeneizar sus condiciones de iluminación. Para superar el problema de la diferencia de tamaño y espesor de los caracteres, las imágenes son estandarizadas a un tamaño de 40 x 40 píxeles, para alimentarlas a un sistema de redes neuronales en combinación con el principio de Análisis de Componentes Principales. En [8] se propone un método de reconocimiento de caracteres, utilizando la técnica de comparación
-

de plantillas en la que se emplean arboles de búsqueda que usan características particulares de los caracteres para disminuir el número de comparaciones necesarias, además, en lugar de utilizar plantillas del carácter completo, se emplean solamente sus esqueletos. Con el fin de definir los árboles de búsqueda, se identificaron las características más destacadas para las letras:

- a) Primero se analizó si tenían contornos interiores o no, para poder diferenciar letras como la A, D, P de aquellos que, como C, I o T. Entonces, en el caso haber contornos interiores, se analizó el tamaño de tales contornos de la letra en cuestión, con el fin de distinguir entre dos posibles grupos (A, P, R y D, O, Q). En éste grupo, existe una letra en particular (B) que contiene dos contornos interiores.
- b) Si no existen contornos interiores, se emplearon otras características particulares, mediante la generación de contornos con una línea horizontal o vertical en la matriz de pixeles del carácter y generando nuevos árboles de búsqueda con tales contornos.

### 2.3. Reconocimiento Automático de Matrículas en México

En México, se ha planteado el problema del reconocimiento de matrículas en [18], en éste trabajo se plantea un algoritmo de segmentación basado en el porcentaje del área que ocupan los caracteres dentro de la placa y en la baja intensidad de los caracteres que conforman el número de matrícula; según éste trabajo, el área que abarcan los caracteres es del 20 %.

Con el criterio del 20 %, las imagenes son binarizadas asignando un 1 a los píxeles cuya intensidad es menor que dicho valor, considerando las más bajas, tomadas del histograma acumulativo de intensidades; las imagenes resultantes son subdivididas horizontalmente en dos para eliminar los elementos no deseados de la parte superior e inferior de la placa, finalmente son nuevamente unidas.

De las imágenes binarizadas se extraen los caracteres, se modelan empleando momentos de Hu y descriptores de Fourier, para posteriormente ser alimentados a un Clasificador Bayesiano que funge como reconocedor.

---

# Propuesta de reconocimiento de matrículas de vehículos

---

A continuación se presenta el algoritmo desarrollado en este trabajo, el cual consiste en: 1) Normalización y binarización de la placa, 2) Segmentación de los caracteres, 3) Modelado de caracteres y 4) Reconocimiento de caracteres.

## 3.1. Normalización y binarización de la placa

La normalización de las imágenes se basa en la norma **NOM-001-SCT-2-2000**, en donde se establece que la proporción de dimensiones de las placas debe ser de 2 : 1, es decir 300 mm de ancho por 150 mm de alto. De aquí que, todas las imágenes se redimensionan a  $700 \times 350$  píxeles antes de ser binarizadas.

En la misma norma también se establece la posición en que deben colocarse los caracteres dentro de la placa, así como el tamaño y la distancia entre los mismos, como se muestra en la Figura 3.1

Con esta información se eliminan las secciones superior e inferior de la imagen original, con el objetivo de obtener una sección que contiene exclusivamente dichos caracteres. El resultado puede observarse en la Tabla 3.1

Debido a la gran variedad en los colores de fondo en las placas, antes de binarizarlas, fue necesario separar la textura del fondo de los caracteres que conforman la matrícula.

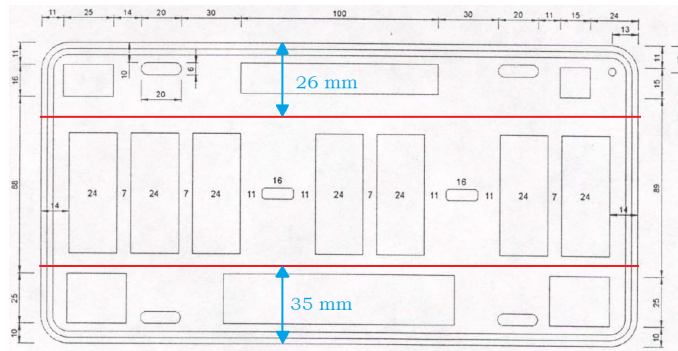


Figura 3.1: Posición de los caracteres dentro de la matrícula.

Tabla 3.1: Ejemplo de eliminación de secciones superiores e inferiores; (a) imagen original, (b) imagen obtenida tras eliminación de secciones



Es importante mencionar que, en este trabajo, los colores de la imagen están representados en el espacio RGB. En este espacio los colores son representados como una combinación lineal de los vectores base de rojo, verde y azul; es decir, el color de un pixel es representado como  $\phi = [r, g, b]$ . En la Figura 3.2 se muestra la forma del espacio RGB.

Para desarrollar el proceso de separación, se obtuvo de forma experimental un umbral de separación que consiste en la magnitud de un vector RGB, a partir del cual se realiza una comparación, para clasificar un pixel en una región u otra, ver Figura 3.2. El umbral ( $\delta$ ) obtenido es de 0.73.

Se puede ver en la Figura 3.2 la parte del espacio en que se encuentran los colores que forman parte de los caracteres. La separación se realizó utilizando la ecuación 3.1.

$$\Theta^* = \begin{cases} \vec{1}, & \|\phi\| \leq \delta \\ \vec{0}, & \|\phi\| > \delta \end{cases} \quad (3.1)$$

Donde  $\delta = 0.73$ ,  $\vec{1} = [1, 1, 1]$  y  $\vec{0} = [0, 0, 0]$ .

Con el método planteado, se logró diferenciar el fondo de los caracteres y con ello la posibilidad

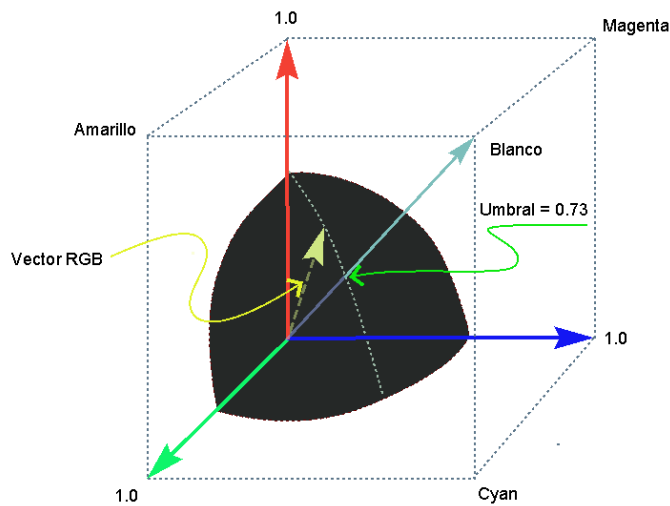


Figura 3.2: Espacio RGB y umbral de binarización.

de binarizar las imágenes obtenidas, para ello se utilizó el método de Otsu [21]. En la Tabla 3.2 se muestran dos ejemplos de imágenes binarizadas empleando el método de Otsu.

Tabla 3.2: *Región binarizada de los caracteres de un par de placas*



### 3.2. Segmentación de caracteres

Segmentar los caracteres consiste en extraer de la imagen original, las regiones que contienen de forma exclusiva, el símbolo de alguna letra o número. Para obtener dicha región, es necesario identificar las coordenadas de los límites que conforma la imagen del carácter, es decir, ubicar sus 4 esquinas, Figura 3.3



Figura 3.3: Segmentación de caracteres.

Para segmentar los caracteres se emplea el método de proyección horizontal y vertical [7], esta

técnica, ya ha sido empleada para fines similares en [4], [6], [23], [25] y [35]. Dada una imagen  $I(x, y)$  de ancho  $N$ , altura  $M$  y considerando que  $1 \leq x \leq N$  y  $1 \leq y \leq M$ , las proyecciones horizontal y vertical se definen como sigue:

$$P_{hor}(y_0) = \sum_{x=1}^N I(x, y_0), \forall y = 1, \dots, M \quad (3.2)$$

$$P_{ver}(x_0) = \sum_{y=1}^M I(x_0, y), \forall x = 1, \dots, N \quad (3.3)$$

Los histogramas obtenidos mediante las ecuaciones 3.2 y 3.3, permiten determinar las coordenadas de la región en que se encuentra cada carácter. La Figura 3.4 muestra un ejemplo de proyección horizontal y vertical de una imagen binarizada.

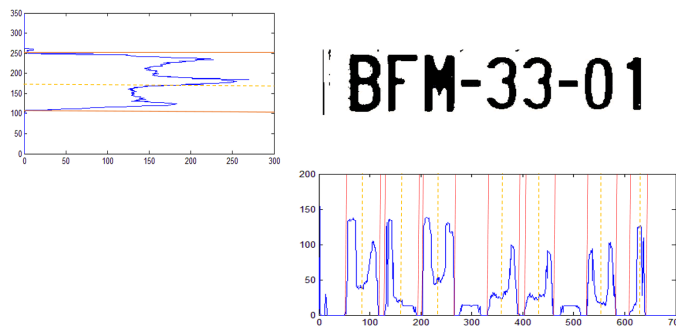


Figura 3.4: Proyección horizontal y vertical de una imagen binarizada.

En la Figura 3.5 se muestran algunos ejemplos de caracteres segmentados empleando las proyecciones horizontal y vertical.



Figura 3.5: Proyección horizontal y vertical de una imagen binarizada.

### 3.3. Modelado de los caracteres

Los caracteres segmentados son modelados utilizando el factor de correlación, los momentos de Hu y los descriptores de Fourier.



### 3.3.1. Factor de Correlación

La correlación es una técnica estadística que cuantifica la intensidad de la relación lineal entre dos variables. La cuantificación se realiza mediante el coeficiente de correlación lineal de Pearson [12], cuyo valor oscila entre -1 y 1. El coeficiente de correlación  $r$  se calcula como sigue:

$$r = \frac{\sum_{(x,y) \in O} (x - \bar{x})(y - \bar{y})}{\sqrt{\sum_{(x,y) \in O} (x - \bar{x})^2 \sum_{(x,y) \in O} (y - \bar{y})^2}} \quad (3.4)$$

Un valor cercano a  $-1$  indica una relación inversa o negativa entre las variables, un valor cercano a 0 indica poca o nula relación y finalmente, un valor cercano a 1 indica una relación directa entre las variables.

### 3.3.2. Momentos de Hu

Los momentos de Hu [10] han sido empleados para reconocer caracteres en [9] y [32]; en otros casos, para medir características geométricas como elipticidades [37] o circularidades [38]. Sea  $O = \{(x_1, y_1), \dots, (x_m, y_m)\}$  el conjunto de coordenadas de los píxeles que conforman un carácter extraído. Los momentos de Hu se calculan como sigue:

$$\begin{aligned} H_1 &= \eta_{20} + \eta_{02} \\ H_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\ H_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\ H_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} - \eta_{03})^2 \\ H_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ &\quad + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ H_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ &\quad + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\ H_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ &\quad - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned} \quad (3.5)$$

Donde  $\eta_{pq}$  es el momento centralizado y normalizado obtenido con:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{c+1}} \quad (3.6)$$

En la ecuación 3.6,  $c$  y  $\mu_{pq}$  se calculan de la forma siguiente:

$$c = \frac{(p+q)}{2}, p+q = 2, 3, \dots \quad (3.7)$$

$$\mu_{pq}(O) = \sum_{(x,y) \in O} (x - \bar{x})^p (y - \bar{y})^q \quad (3.8)$$

Finalmente,  $(\bar{x}, \bar{y})$  es la coordenada del centroide del objeto, que se calcula como sigue:

$$\bar{x} = \frac{1}{m} \sum_{(x,y) \in O} x \quad (3.9)$$

$$\bar{y} = \frac{1}{m} \sum_{(x,y) \in O} y \quad (3.10)$$

Los momentos de Hu son invariantes a la posición, rotación y escalado.

### 3.3.3. Descriptores de Fourier

La Figura 3.6 muestra el contorno de un carácter, que está definido por  $n$  puntos en el plano complejo. Sea  $\{(x_1, y_1), \dots, (x_n, y_n)\} \subset \mathbb{Z}^2$  el conjunto de puntos que forman el contorno, cada punto se representa como un número complejo:

$$z_n = x_n + j y_n \quad (3.11)$$

Donde:  $j = \sqrt{-1}$ .

La transformada discreta de Fourier [24] del conjunto de puntos que conforman el contorno del

caracter se puede calcular como sigue:

$$F(u) = \frac{1}{u\pi} \sum_{n=1}^m \varphi_n \left[ \cos\left(\frac{2\pi u s_n}{s_m}\right) - j \sin\left(\frac{2\pi u s_n}{s_m}\right) \right] \quad (3.12)$$

También se puede escribir como:

$$F(u) = \frac{1}{u\pi} \sum_{n=1}^m \varphi_n \exp\left(\frac{-j2\pi u s_n}{s_m}\right) \quad (3.13)$$

Finalmente se obtienen los descriptores de Fourier al calcular los valores absolutos de los números complejos:

$$f(u) = |F(u)| \quad (3.14)$$

Donde  $u = 1, 2, \dots, N$  y  $N$  es el total de descriptores a obtener,

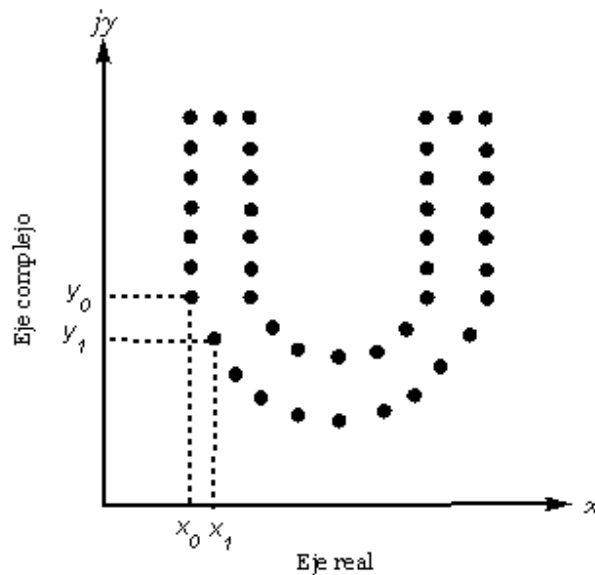


Figura 3.6: Conjunto de puntos que forman el contorno de un caracter

### 3.4. Reconocimiento de los caracteres

Para el reconocimiento de los caracteres, se utilizaron tres técnicas: 1) Comparación de plantillas, 2) Clasificador bayesiano y 3) Redes Neuronales Artificiales (RNA).

#### 3.4.1. Comparación de plantillas

La comparación de plantillas es una técnica que consiste en comparar la imagen del carácter que se desea reconocer, con una serie de plantillas conocidas para medir su similaridad y con ello identificar el símbolo que contiene la imagen a clasificar [5], ésta técnica ha sido empleada para el mismo fin en [8] y [33]. Cuando una imagen debe ser clasificada, se compara con todas las imágenes de las plantillas y se obtiene el máximo valor de similaridad ( $\Xi$ ):

$$\Xi_{\sigma} = \text{Max}(r_{\sigma}), \forall \sigma \in (0, 1, 2, \dots, 9, A, B, C, D, \dots, Z) \quad (3.15)$$

Con el valor  $\Xi$  obtenido se define qué símbolo es el contenido en la imagen. En la Figura 3.7 se muestran las diferentes similaridades obtenidas al procesar la imagen correspondiente al símbolo **G**.

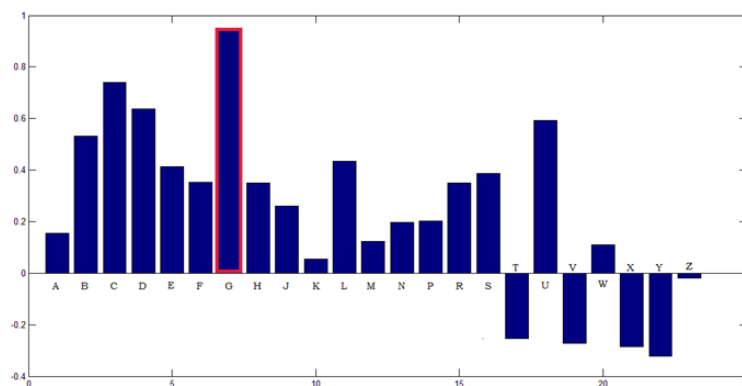


Figura 3.7: Similaridades para la imagen del caracter **G**

### 3.4.2. Clasificador Bayesiano

El clasificador Bayesiano se basa en el teorema de Bayes en donde se asume que el vector de características es una distribución Gaussiana multivariable [17]. Sea  $C = \{k_1, \dots, k_n\}$  el conjunto de  $n$  clases de objetos, la probabilidad de que un objeto  $\mathbf{A}$  sea de clase  $k_i$  se denota por  $p(k_i|\mathbf{A})$ . Del teorema de Bayes:

$$P(k_i|\mathbf{A}) = \frac{P(\mathbf{A}|k_i)P(k_i)}{P(\mathbf{A})} \quad (3.16)$$

Para la clasificación Bayesiana, se elige la clase  $k_i$  para la cual la  $P(\mathbf{A}|k_i)p(k_i)$  es mayor. De esta forma al objeto observado  $\mathbf{A}$  se le asigna la clase  $k_i$ . Asumimos que la distribución de probabilidad del vector de características del objeto observado  $\mathbf{A}$ , sea de clase  $k_j$ , es Gaussiana con media  $\mu_j$  y con matriz de covarianza  $\Omega_j$ :

$$P(\mathbf{A}|k_i) = \frac{1}{\Delta} \exp \left[ -\frac{1}{2}(\mathbf{A} - \mu_j)^T \Omega_j^{-1} (\mathbf{A} - \mu_j) \right] \quad (3.17)$$

Donde  $\Delta = (2\pi)^{m/2}(\det \Omega_j)^{1/2}$  y  $m$  es la dimensión de los vectores de características.

### 3.4.3. Red Neuronal Artificial

Una Red Neuronal Artificial (RNA) [16] es un modelo matemático que se inspira en la forma en que trabajan las neuronas biológicas; las cuales, a partir del conjunto de valores de entrada  $\{\phi_1, \phi_2, \dots, \phi_n\}$  de una instancia, se interconectan a una o más capas de neuronas que procesan los datos y cuyo objetivo es clasificar u obtener el valor objetivo de la instancia dada. Al igual que el Clasificador Bayesiano, la RNA es aplicada en el aprendizaje de tareas donde cada instancia  $\mathbf{A}$  es descrita por un conjunto de valores, que representan sus atributos y donde la función objetivo  $f(\mathbf{A})$  puede tomar cualquier valor de un conjunto finito  $V$ .

En el presente trabajo, se implementan dos RNA, una dedicada a la clasificación de los números y otra para las letras. Las RNA empleadas son del tipo Backpropagation cuya función de activación

es sigmoide.

### 3.5. Análisis de componentes principales

El análisis de componentes principales convierte un conjunto de observaciones de variables correlacionadas en otro de variables no correlacionadas con el objetivo de reducir su dimensionalidad.

Las componentes principales, se obtienen como se describe a continuación:

- Sea el conjunto de datos  $\{\phi_1, \phi_2, \dots, \phi_n\} \subset \mathbb{R}$ , se calcula la media:

$$\mu = \frac{1}{n} \sum_{i=1}^n \phi_i \quad (3.18)$$

- Se construye la matriz  $\Phi$  :

$$\Phi = [\phi_1 - \mu, \dots, \phi_n - \mu] \quad (3.19)$$

- Se obtiene la matriz de covarianza  $\Omega$  :

$$\Omega = \frac{1}{n} \Phi \Phi^T \quad (3.20)$$

- Se calculan los valores propios  $\{\lambda_1, \dots, \lambda_n\}$  y vectores propios  $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$  de la matriz de covarianza. Los vectores propios se ordenan de mayor a menor con respecto a su correspondiente valor propio y se colocan en la matriz  $\mathbf{V}$  como columnas.
- Finalmente se calculan las componentes principales proyectando los vectores iniciales  $\{\phi_1, \phi_2, \dots, \phi_n\}$  al espacio que forman las componentes principales:

$$\psi_i = \mathbf{V}^T(\phi_i - \mu), i = 1, 2, \dots, n \quad (3.21)$$

El análisis de componentes principales ha sido utilizado en [4], [26] y [34] para reducir la dimensión del espacio de los vectores de características empleados para realizar las clasificaciones. Se consideraron sólo las características cuya varianza acumulada es mayor o igual que el 95 % del total; bajo éste enfoque la precisión no se ve afectada considerablemente, sin embargo, los costos computacionales de procesamiento y memoria, si se ven disminuidos.

---





# Experimentos, resultados y discusión

---

En éste capítulo se explica la forma en que se desarrollaron los experimentos y se presentan los resultados obtenidos con los clasificadores implementados en cada una de las pruebas, también se realiza una discusión de los datos recabados.

## 4.1. Experimentos

Para realizar la experimentación se empleó una base de datos de 70 imágenes de placas de los 31 estados de la República Mexicana, excepto la Ciudad de México debido a que en ésta entidad las matrículas son diferentes (contienen 6 caracteres en lugar de 7). Esta base de datos se eligió debido a que se consideró el muestreo por conveniencia.

Las imágenes fueron capturadas a una distancia de 1.5 a 3 metros entre la cámara y la placa, procurando una iluminación uniforme y posteriormente normalizadas a  $700 \times 350$  pixeles.

La experimentación se realizó en dos fases:

- Entrenamiento del reconocedor de letras y números.
- Reconocimiento de matrículas.

### 4.1.1. Entrenamiento del reconocedor de letras y números

Para probar los reconocedores, se utilizaron 738 caracteres, 374 letras y 364 números, en la Tabla 4.1 se muestra el número de caracteres utilizado en cada caso, dicho número no es igual para

Tabla 4.1: Número de ejemplos empleados en el entrenamiento

| A  | B  | C  | D  | E  | F  | G  | H  | J  | K  | L  |
|----|----|----|----|----|----|----|----|----|----|----|
| 14 | 9  | 11 | 17 | 16 | 18 | 26 | 21 | 20 | 13 | 14 |
| M  | N  | P  | R  | S  | T  | U  | V  | W  | X  | Y  |
| 13 | 14 | 20 | 17 | 12 | 18 | 18 | 21 | 13 | 15 | 18 |
| Z  | 0  | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  |
| 16 | 27 | 27 | 40 | 32 | 41 | 50 | 43 | 34 | 33 | 37 |

Tabla 4.2: Resultados de clasificación de números con 8 características

| caracter     | Clasificador Bayesiano |              | Red Neuronal   |              |
|--------------|------------------------|--------------|----------------|--------------|
|              | Aciertos               | %            | Aciertos       | %            |
| <b>0</b>     | 8/27                   | 29.62        | 23/27          | 85.18        |
| <b>1</b>     | 27/27                  | 100.00       | 25/27          | 92.53        |
| <b>2</b>     | 28/40                  | 70.00        | 40/40          | 100.00       |
| <b>3</b>     | 29/32                  | 90.62        | 32/32          | 100.00       |
| <b>4</b>     | 15/41                  | 36.58        | 39/41          | 95.12        |
| <b>5</b>     | 35/50                  | 70.00        | 50/50          | 100.00       |
| <b>6</b>     | 9/43                   | 20.93        | 26/43          | 60.46        |
| <b>7</b>     | 32/34                  | 94.11        | 32/34          | 94.11        |
| <b>8</b>     | 30/33                  | 90.90        | 33/33          | 100.00       |
| <b>9</b>     | 2/37                   | 5.40         | 33/37          | 89.18        |
| <b>Total</b> | <b>246/364</b>         | <b>67.58</b> | <b>333/364</b> | <b>91.48</b> |

todos los símbolos debido a que no aparecen con la misma frecuencia en las matrículas.

Se implementaron tres tipos de reconocedores de caracteres: 1) Coeficiente de correlación, 2) Clasificador Bayesiano y 3) Redes Neuronales Artificiales (RNA), en éstas últimas, se implementaron 2 redes; una para los números y otra para las letras.

Para el caso del Clasificador Bayesiano y las RNA, primero se probaron con vectores de características de dimensión 8, formados por los 7 momentos de Hu y el Factor de Correlación, los resultados al clasificar los números y las letras se muestran en las Tablas 4.2 y 4.3. En este experimento, las redes son de 2 capas con la siguiente estructura: [8, 10] para los números y [8, 20] para las letras.

Posteriormente se añadieron a los vectores de características: 30 y 60 descriptores de Fourier, formando vectores de dimensión 38: factor de correlación, 7 momentos de Hu y 30 descriptores de Fourier; y dimensión 68: factor de correlación, 7 momentos de Hu y 60 descriptores de Fourier. Los resultados obtenidos en las pruebas con el clasificador Bayesiano pueden observarse en las

Tabla 4.3: Resultados de clasificacion de letras con 8 características

| Clasificador | Bayesiano    |                | Red Neuronal |                | caracter     | Bayesiano |          | Red Neuronal |          |        |
|--------------|--------------|----------------|--------------|----------------|--------------|-----------|----------|--------------|----------|--------|
|              | caracter     | Aciertos       | %            | Aciertos       |              | %         | Aciertos | %            | Aciertos | %      |
|              | A            | 12/14          | 85.71        | 10/14          | 71.42        | N         | 14/14    | 100.00       | 9/14     | 64.28  |
|              | B            | 0/9            | 0.00         | 1/9            | 11.11        | P         | 20/20    | 100.00       | 11/20    | 55.00  |
|              | C            | 9/11           | 81.81        | 11/11          | 100.00       | R         | 17/17    | 100.00       | 17/17    | 100.00 |
|              | D            | 10/17          | 58.82        | 7/17           | 41.17        | S         | 12/12    | 100.00       | 12/12    | 100.00 |
|              | E            | 16/16          | 100.00       | 16/16          | 100.00       | T         | 18/18    | 100.00       | 18/18    | 100.00 |
|              | F            | 18/18          | 100.00       | 0/18           | 0.00         | U         | 15/18    | 83.33        | 15/18    | 83.33  |
|              | G            | 15/26          | 57.69        | 7/26           | 26.92        | V         | 18/21    | 85.71        | 15/21    | 71.42  |
|              | H            | 12/21          | 57.14        | 18/21          | 85.71        | W         | 11/13    | 84.61        | 9/13     | 69.23  |
|              | J            | 6/20           | 30.00        | 17/20          | 85           | X         | 13/15    | 86.66        | 15/15    | 100.00 |
|              | K            | 9/13           | 69.23        | 13/13          | 100.00       | Y         | 18/18    | 100.00       | 18/18    | 100.00 |
|              | L            | 12/14          | 85.71        | 12/14          | 85.71        | Z         | 12/16    | 75.00        | 16/16    | 100.00 |
|              | M            | 7/13           | 53.84        | 6/13           | 46.15        |           |          |              |          |        |
|              | <b>Total</b> | <b>294/374</b> | <b>78.60</b> | <b>273/374</b> | <b>72.99</b> |           |          |              |          |        |

tablas 4.4 para los números y 4.5 para las letras. En éste experimento también se implementaron las redes neuronales siguientes: para el vector de 38 características: Letras = [38, 76, 23] Números = [38, 38, 10], en el caso del vector de 68 elementos, las redes implementadas tienen la siguiente estructura: Letras = [68, 68, 23] Números = [68, 34, 10], el número de elementos por capa fue obtenido de forma experimental. Ambas redes clasificaron correctamente el 100 % de los ejemplos.

Finalmente, se realizó el Análisis de Componentes Principales con lo cual se logró reducir la dimensión de los vectores de características de la siguiente forma:

- Vector de 38 características: el vector de números disminuyó a 9 componentes, mientras que el vector de letras pasó de 38 a 8 elementos. Las redes neuronales implementadas en este caso, tienen la siguiente composición: Números: [9, 27, 36, 10], Letras: [8, 16, 32, 23].
- Vector de 68 características: , el vector de números disminuyó a 13 componentes, mientras que el vector de letras pasó a sólo 11. Las redes neuronales implementadas en este caso, tienen la siguiente composición: Números: [13, 26, 13, 10], Letras: [11, 22, 46, 23].

Los resultados obtenidos se muestran en la Tabla 4.6 para los números y la Tabla 4.7 para las letras.

En la Tabla 4.8 pueden observarse los resultados totales para el caso del clasificador Bayesiano.

Tabla 4.4: Resultados de la clasificación de numeros utilizando el clasificador Bayesiano con descriptores de Fourier

| caracter     | 38 Características |              | 68 Características |              |
|--------------|--------------------|--------------|--------------------|--------------|
|              | Aciertos           | %            | Aciertos           | %            |
| <b>0</b>     | 20/27              | 74.07        | 25/27              | 92.6         |
| <b>1</b>     | 27/27              | 100.00       | 27/27              | 100.00       |
| <b>2</b>     | 39/40              | 97.50        | 39/40              | 97.5         |
| <b>3</b>     | 31/32              | 96.88        | 31/32              | 96.9         |
| <b>4</b>     | 39/41              | 95.12        | 37/41              | 90.2         |
| <b>5</b>     | 43/50              | 86.00        | 43/50              | 86.0         |
| <b>6</b>     | 38/43              | 88.37        | 40/43              | 93.0         |
| <b>7</b>     | 34/34              | 100.00       | 34/34              | 100.0        |
| <b>8</b>     | 31/33              | 93.94        | 32/33              | 97.0         |
| <b>9</b>     | 23/37              | 62.16        | 23/37              | 62.2         |
| <b>Total</b> | <b>325/364</b>     | <b>89.40</b> | <b>331/364</b>     | <b>91.50</b> |

Tabla 4.5: Resultados de la clasificación de letras utilizando el clasificador Bayesiano con descriptores de Fourier

| caracter | 38 Características |        | 68 Características |        | caracter     | 38 Características |             | 68 Características |             |
|----------|--------------------|--------|--------------------|--------|--------------|--------------------|-------------|--------------------|-------------|
|          | Aciertos           | %      | Aciertos           | %      |              | Aciertos           | %           | Aciertos           | %           |
| A        | 14/14              | 100.00 | 14/14              | 100.00 | N            | 14/14              | 100.00      | 14/14              | 100.00      |
| B        | 9/9                | 100.00 | 9/9                | 100.0  | P            | 19/20              | 95.00       | 19/20              | 95.0        |
| C        | 11/11              | 100.00 | 11/11              | 100.00 | R            | 17/17              | 100.00      | 17/17              | 100.00      |
| D        | 17/17              | 100.00 | 17/17              | 100.00 | S            | 12/12              | 100.00      | 12/12              | 100.00      |
| E        | 16/16              | 100.00 | 16/16              | 100.00 | T            | 17/18              | 94.4        | 17/18              | 94.4        |
| F        | 18/18              | 100.00 | 18/18              | 100.00 | U            | 17/18              | 94.4        | 17/18              | 94.4        |
| G        | 26/26              | 100.00 | 26/26              | 100.00 | V            | 16/21              | 76.2        | 21/21              | 100.00      |
| H        | 21/21              | 100.00 | 21/21              | 100.00 | W            | 11/13              | 84.6        | 13/13              | 100.00      |
| J        | 20/20              | 100.00 | 20/20              | 100.00 | X            | 14/15              | 93.3        | 15/15              | 100.00      |
| K        | 13/13              | 100.00 | 13/13              | 85.71  | Y            | 18/18              | 100.00      | 18/18              | 100.00      |
| L        | 12/14              | 85.7   | 13/13              | 100.00 | Z            | 16/16              | 100.00      | 16/16              | 100.00      |
| M        | 12/13              | 92.3   | 13/13              | 100.00 |              |                    |             |                    |             |
|          |                    |        |                    |        | <b>Total</b> | <b>360/374</b>     | <b>96.4</b> | <b>369/374</b>     | <b>98.7</b> |

Tabla 4.6: Resultados de clasificación de números utilizando el clasificador Bayesiano con Análisis de Componentes Principales

| caracter     | 8 Características |             | 13 Características |             |
|--------------|-------------------|-------------|--------------------|-------------|
|              | Aciertos          | %           | Aciertos           | %           |
| <b>0</b>     | 25/27             | 92.6        | 25/27              | 92.6        |
| <b>1</b>     | 26/27             | 96.3        | 26/27              | 96.3        |
| <b>2</b>     | 35/40             | 87.5        | 34/40              | 85.0        |
| <b>3</b>     | 31/32             | 96.9        | 31/32              | 96.9        |
| <b>4</b>     | 35/41             | 85.4        | 35/41              | 85.4        |
| <b>5</b>     | 41/50             | 82.0        | 42/50              | 84.0        |
| <b>6</b>     | 13/43             | 30.2        | 14/43              | 32.6        |
| <b>7</b>     | 32/34             | 94.1        | 33/34              | 97.1        |
| <b>8</b>     | 25/33             | 75.8        | 26/33              | 78.8        |
| <b>9</b>     | 28/37             | 75.7        | 27/37              | 73.0        |
| <b>Total</b> | <b>291/364</b>    | <b>81.6</b> | <b>293/364</b>     | <b>82.2</b> |

Tabla 4.7: Resultados de clasificación de letras utilizando el clasificador Bayesiano con Análisis de Componentes Principales

| caracter<br>% | 8 Características |        | 13 Características |        | caracter     | 8 Características |             | 13 Características |             |
|---------------|-------------------|--------|--------------------|--------|--------------|-------------------|-------------|--------------------|-------------|
|               | Aciertos          | %      | Aciertos           | %      |              | caracter          | Aciertos    | %                  | Aciertos    |
| A             | 13/14             | 92.9   | 13/14              | 92.9   | N            | 14/14             | 100.00      | 14/14              | 100.00      |
| B             | 8/9               | 88.9   | 8/9                | 88.9   | P            | 19/20             | 95.00       | 19/20              | 95.0        |
| C             | 10/11             | 90.9   | 10/11              | 90.9   | R            | 15/17             | 88.2        | 17/17              | 100.00      |
| D             | 16/17             | 94.1   | 16/17              | 94.1   | S            | 12/12             | 100.00      | 12/12              | 100.00      |
| E             | 16/16             | 100.00 | 16/16              | 100.00 | T            | 16/18             | 88.9        | 16/18              | 88.9        |
| F             | 18/18             | 100.00 | 18/18              | 100.00 | U            | 17/18             | 94.4        | 17/18              | 94.4        |
| G             | 24/26             | 92.3   | 26/26              | 100.00 | V            | 18/21             | 85.7        | 17/21              | 81.00       |
| H             | 20/21             | 95.2   | 21/21              | 100.00 | W            | 9/13              | 69.2        | 9/13               | 69.2        |
| J             | 16/20             | 80.00  | 17/20              | 85.00  | X            | 14/15             | 93.3        | 14/15              | 93.3        |
| K             | 13/13             | 100.00 | 13/13              | 100.00 | Y            | 17/18             | 94.4        | 16/18              | 88.9        |
| L             | 14/14             | 100.00 | 13/14              | 92.9   | Z            | 14/16             | 87.5        | 14/16              | 87.5        |
| M             | 13/13             | 100.00 | 12/13              | 92.3   |              |                   |             |                    |             |
|               |                   |        |                    |        | <b>Total</b> | <b>346/374</b>    | <b>92.7</b> | <b>348/374</b>     | <b>92.8</b> |

Tabla 4.8: Resumen de resultados generales para el clasificador Bayesiano

|                | Dim. 8       | Dim. 38      | Dim. 68      | PCA dim. 38 | PCA dim. 68 |
|----------------|--------------|--------------|--------------|-------------|-------------|
| <b>Numeros</b> | <b>91.48</b> | <b>89.40</b> | <b>91.50</b> | <b>81.6</b> | <b>82.2</b> |
| <b>Letras</b>  | <b>72.99</b> | <b>96.4</b>  | <b>98.7</b>  | <b>92.7</b> | <b>92.8</b> |

Tabla 4.9: Número de ejemplos de matrículas por estado

|                |                 |                     |          |                  |
|----------------|-----------------|---------------------|----------|------------------|
| Aguascalientes | Baja California | Baja California Sur | Campeche | Chihuahua        |
| 2              | 1               | 0                   | 2        | 2                |
| Colima         | Coahuila        | Chiapas             | Durango  | Estado de México |
| 1              | 2               | 2                   | 1        | 3                |
| Guerrero       | Guanajuato      | Hidalgo             | Jalisco  | Michoacán        |
| 3              | 4               | 3                   | 4        | 4                |
| Morelos        | Nuevo León      | Nayarit             | Oaxaca   | Puebla           |
| 2              | 2               | 3                   | 2        | 2                |
| Querétaro      | Quintana Roo    | San Luis Potosí     | Sinaloa  | Sonora           |
| 4              | 1               | 4                   | 3        | 1                |
| Tabasco        | Tamaulipas      | Tlaxcala            | Veracruz | Yucatán          |
| 2              | 2               | 1                   | 2        | 2                |
| Zacatecas      | <b>Total</b>    |                     |          |                  |
| 3              | <b>70</b>       |                     |          |                  |

#### 4.1.2. Reconocimiento de matrículas

En la Tabla 4.9 se muestra el número de imágenes de placas utilizadas por estado. Los resultados obtenidos al probar los reconocedores desarrollados se muestran en la Figura 4.1. Para considerar el reconocimiento de la matrícula como acierto, se tiene el siguiente criterio: se dice que una matrícula es reconocida exitosamente si todos sus caracteres son clasificados correctamente; esto implica que con tener un carácter erróneamente clasificado significa que la matrícula no es reconocida exitosamente.

Los resultados obtenidos en la clasificación de matrículas, para todos los reconocedores de caracteres implementados, se pueden observar en la Figura 4.1

#### 4.1.3. Discusión

En el desarrollo de ésta propuesta se implementaron 3 tipos de reconocedores: 1) Factor de Correlación, 2) Clasificador Bayesiano y 3) Red Neuronal Artificial, de éstas técnicas, el Factor de correlación tuvo el mejor desempeño como puede observarse en la Figura 4.1 la principal dificultad de esta técnica radica en elegir adecuadamente las plantillas a comparar, sin embargo, las ocasiones

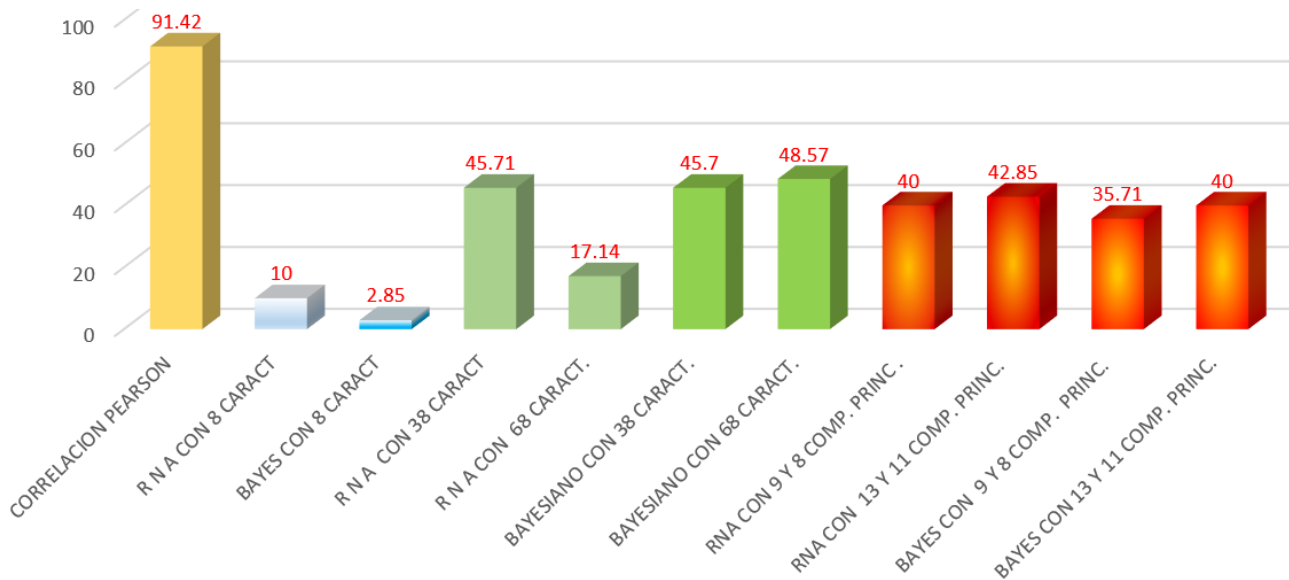


Figura 4.1: Resultados de clasificación de placas para los diferentes reconocedores implementados

en las que falla, es debido a que los caracteres segmentados tienen un exceso de ruido o bien la forma y el tamaño general del carácter por reconocer es muy aproximado al del ejemplo de la plantilla, esto ocurre por ejemplo con los números 1 y 7.

Utilizando el clasificador Bayesiano, cuando se utilizan solamente 8 características, el porcentaje de aciertos al clasificar las matrículas es de sólo 2.85 %, sin embargo, cuando se incrementó el número de características con los descriptores de Fourier, el porcentaje mejoró considerablemente, pasando a 45.7 % con 38 características y 48.57 % con 68 características, siendo éste el mejor valor alcanzado. Cuando se implementó el Análisis de Componentes Principales, el porcentaje disminuyó en promedio un 8 %, lo cual muestra que el criterio de utilizar las Componentes Principales cuya varianza acumulada es de al menos 95 % es adecuado, ya que se logró disminuir el vector de características hasta un 88 %.

En el caso de las redes neuronales, con 38 características se logró un 45.71 %, mientras que sólo alcanzó un 17.14 % con 68 características; sin embargo, al realizar el Análisis de las Componentes Principales, al reducir el tamaño del vector de 38 características, el desempeño disminuyó sólo un 5 %, por otro lado, al disminuir el vector de 68 características, el porcentaje alcanzado se incrementó al 43.85 %, lo cual muestra que incrementar el número de elementos en el modelado de los caracteres, no siempre es la mejor alternativa.

Los errores de clasificación tanto en el reconocedor Bayesiano se presentan de la siguiente forma: a) En los números el 9 que se clasifica como 6 y 4, el 1 se reconoce como 7 y el 8 que se clasifica como 0. b) En el caso de las letras: J y T se clasifican como L y H, respectivamente; y letras X y M son reconocidas como W.

Los errores de clasificación en las Redes Neuronales Artificiales, se dan como sigue: a) Para los números el 9 se identifica como 6 y viceversa, 5 como 3 y en algunos casos el 4 como 9. b) Para las letras la R se clasifica como K; la L es reconocida como J; la M es clasificada como W y viceversa; la V es reconocida como A.

En ambos clasificadores, otro motivo de error en el reconocimiento de los caracteres se da en aquellos cuya apariencia general o su forma son parecidas como el 5 y el 8, la H y la K, por otro lado, es importante recordar que los momentos de Hu son invariantes a la rotación, por lo que en varios errores de clasificación se presentan en imágenes que son iguales o casi iguales si están rotadas. Por ejemplo, la M y la W, el 6 y el 9, la V y la A.

---



# Conclusiones

---

Se ha presentado una propuesta para el reconocimiento de matrículas de vehículos mexicanos que logra diferenciar de la imagen, los caracteres que la conforman de los colores de textura del fondo, mediante el uso de un umbral, los resultados obtenidos muestran que la técnica propuesta es robusta a ligeras variaciones de iluminación, marcas de agua y al color de los caracteres; siempre y cuando éstos conserven una baja intensidad, debido a que en las pruebas de reconocimiento, aún cuando algunas de las técnicas empleadas sólo alcanzaron como máximo un 48 %, con el Factor de Correlación de Pearson se alcanzó un 91.42 %, lo cual muestra que la segmentación de los caracteres fue adecuada.

Ello es relevante debido a que en los trabajos relacionados que tratan este tema diseñan sus algoritmos asumiendo que las matrículas están formadas con caracteres negros sobre fondo blanco, mientras que en el caso de las matrículas mexicanas tienen diferentes patrones de textura y de colores en el fondo; por lo que al aplicar dichos algoritmos, su desempeño se ve considerablemente afectado. Porque al momento de segmentar los caracteres, usualmente éstos son extraídos con parte de la textura del fondo o son extraídos incompletos.

En el caso de la tesis “Reconocimiento Automático de Matrículas de Automóviles Particulares Mexicanos”, el autor concluye que su algoritmo tiene problemas cuando las imágenes tienen las siguientes características: 1) Cuando la iluminación no es uniforme, 2) Cuando los caracteres tienen marcas de agua y 3) cuando el color de los caracteres no es negro; sin embargo, como se mencionó al principio, el algoritmo presentado en este trabajo, no tiene problemas ante esas condiciones. Por otro lado, el autor también concluye que no es recomendable emplear la técnica del Análisis de

Componentes Principales para reducir la dimensionalidad del vector de características, no obstante, como puede observarse en la Figura 4.1, los resultados obtenidos al aplicar dicha técnica son similares a aquellos cuando no se aplica, más aún, en el caso de la RNA con 68 características, el rendimiento mejoró. Finalmente, el autor concluye que sus resultados son adecuados con 110 descriptores de Fourier, pero se comprobó experimentalmente que una gran cantidad de descriptores, no necesariamente significa un impacto considerable en el rendimiento del sistema. En este trabajo se obtuvieron resultados similares con sólo 60 descriptores, lo cuál además sólo incrementó en un 3 % el rendimiento obtenido con 30 descriptores, ver Figura 4.1.

Como se muestra en la Tabla 4.8 y la Figura 4.1, se obtiene un reconocimiento exitoso del 90 % de los caracteres extraídos de las imágenes utilizadas en los experimentos; por lo tanto, se considera que la hipótesis planteada al inicio se cumple.

Nuestra propuesta se basa en las características de diseño que deben ser cumplidas de acuerdo a la norma oficial mexicana NOM-001-SCT-2-2000 en donde se establecen las dimensiones de la placa así como el tamaño y ubicación de los caracteres dentro de la misma, ésta información facilitó la segmentación. Por otra parte, se obtuvo de forma experimental un umbral de intensidad de los píxeles que corresponden a los caracteres, lo que permite separarlos del fondo con mayor precisión.

Finalmente, aunque el reconocimiento de placas es bajo, se debe considerar que la clasificación errónea de un solo carácter implica que la placa no sea reconocida exitosamente; siendo esta situación la que ocurrió en la mayoría de los experimentos. La caracterización empleando el Coeficiente de Correlación, los momentos de Hu y los Descriptores de Fourier ha permitido reconocer los caracteres exitosamente, pues aún cuando los resultados obtenidos en la fase de reconocimiento de matrículas son bajos, de forma individual se lograron buenos resultados. Sin embargo, es necesario establecer un mecanismo que permita reducir los efectos negativos de la invarianza a la rotación de los momentos de Hu, ya que varios caracteres son similares si éstos son rotados y por lo tanto son reconocidos erróneamente; por ejemplo el 6 y el 9 o la W y la M, lo cual produce la disminución del porcentaje de aciertos al intentar reconocer la serie que conforma la matrícula.

---

## Referencias

---

- [1] Norma oficial mexicana nom-001-sct-2-2000, placas metálicas, calcomanías de identificación y tarjetas de circulación empleadas en automóviles, autobuses, camiones, minibuses, motocicletas y remolques matriculados en la república mexicana, licencia federal de conductor y calcomania de verificación físico-mecánica-especificaciones y métodos de prueba., 2001.
  
- [2] Wisam Al Faqheri and Syamsiah Mashohor. A real-time malaysian automatic license plate recognition (m-alpr) using hybrid fuzzy. *IJCSNS International Journal of Computer Science and Network Security*, 9(2):333–340, 2009.
  
- [3] Christos Nikolaos E Anagnostopoulos, Ioannis E Anagnostopoulos, Vassilis Loumos, and Eleftherios Kayafas. A license plate-recognition algorithm for intelligent transportation system applications. *IEEE Transactions on Intelligent transportation systems*, 7(3):377–392, 2006.
  
- [4] Yue Cheng, Jiauning Lu, and Takashi Yahagi. Car license plate recognition based on the combination of principal components analysis and radial basis function networks. In *Signal Processing, 2004. Proceedings. ICSP'04. 2004 7th International Conference on*, volume 2, pages 1455–1458. IEEE, 2004.
  
- [5] Shan Du, Mahmoud Ibrahim, Mohamed Shehata, and Wael Badawy. Automatic license plate recognition (alpr): A state-of-the-art review. *IEEE Transactions on circuits and systems for video technology*, 23(2):311–325, 2013.

- 
- [6] Tran Duc Duan, TL Hong Du, Tran Vinh Phuoc, and Nguyen Viet Hoang. Building an automatic vehicle license plate recognition system. In *Proc. Int. Conf. Comput. Sci. RIVF*, pages 59–63. Citeseer, 2005.
- [7] Marco Pérez Erick Cuevas, Daniel Zaldívar. *Procesamiento digital de imagenes con MATLAB y Simulink*. Alfaomega Ra-Ma, primera edición edition, 2010.
- [8] Nicolás Fernando Gazcón, Carlos Iván Chesñevar, and Silvia Mabel Castro. Automatic vehicle identification for argentinean license plates using intelligent template matching. *Pattern Recognition Letters*, 33(9):1066–1074, 2012.
- [9] Sebastián Gómez González and Santiago Gutiérrez Alzate. Desarrollo de un sistema prototipo de reconocimiento de dígitos usando momentos invariantes. 2011.
- [10] Ming-Kuei Hu. Visual pattern recognition by moment invariants. *IRE transactions on information theory*, 8(2):179–187, 1962.
- [11] Jianbin Jiao, Qixiang Ye, and Qingming Huang. A configurable method for multi-style license plate recognition. *Pattern Recognition*, 42(3):358–369, 2009.
- [12] Douglas A Marchal Lind, William G Wathen, Samuel A Douglas A Lind, William G Marchal, and Samuel A Wathen. *Estadística aplicada a los negocios y la economía*. McGraw-Hill,, 2012.
- [13] Yawei Ma, Jieru Chi, Ran Hu, and Guowei Yang. A new algorithm for characters segmentation of license plate based on variance projection and mean filter. In *2011 IEEE 5th International Conference on Cybernetics and Intelligent Systems (CIS)*, pages 132–135. IEEE, 2011.
- [14] M. Gergais R. Bakhit M.A. Massoud, M. Sabee. Automated new license plate recognition in egypt. *Alexandria Engineering Journal*, 2013.
-

- 
- [15] VinhDu Mai, Duoqian Miao, Ruizhi Wang, and Hongyun . An improved method for vietnam license plate location. In *Multimedia Technology (ICMT), 2011 International Conference on*, pages 2942–2946. IEEE, 2011.
- [16] Mark Hudson Beale Orlando de JesÃ³s Martin T. Hagan, Howard B. Demuth. *Neural Network Design*. 2014.
- [17] Tom M. Mitchell. *Machine Learning*. McGraw-Hill Science/Engineering/Math, March 1, 1997.
- [18] Erick Josu  Espinoza Morales. Reconocimiento autom tico de matr culas de autom viles particulares mexicanos. Master’s thesis, Universidad Aut noma del Estado de M xico, 2014.
- [19] Satadal Saha & Subhadip Basu & Mita Nasipuri. ilpr: an indian license plate recognition system. *Multimedia Tools and Applications*, 2014.
- [20] Martinsk  Ondrej, V Zboril Frantisek, and Drahansk  Martin. Algorithmic and mathematical principles of automatic number plate recognition systems. *Brno University of technology*, page 10, 2007.
- [21] Nobuyuki Otsu. A threshold selection method from gray-level histograms. *Automatica*, 11(285-296):23–27, 1975.
- [22] Xiang Pan, Xiuzi Ye, and Sanyuan Zhang. A hybrid method for robust car plate character recognition. *Engineering Applications of Artificial Intelligence*, 18(8):963–972, 2005.
- [23] Zhong Qin, Shengli Shi, Jianmin Xu, and Hui Fu. Method of license plate location based on corner feature. In *2006 6th World Congress on Intelligent Control and Automation*, volume 2, pages 8645–8649. IEEE, 2006.
- [24] Steven L. Eddins Rafael C. Gonzalez, Richard E. Woods. *Digital Image Processing Using MATLAB*. Pearson Prentice Hall, 2004.
-

- 
- [25] Zhang Sanyuan, Zhang Mingli, and Ye Xiuzi. Car plate character extraction under complicated environment. In *Systems, Man and Cybernetics, 2004 IEEE International Conference on*, volume 5, pages 4722–4726 vol.5, 2004.
- [26] Amir Sedighi and Mansur Vafadust. A new and robust method for character segmentation and recognition in license plate images. *Expert Systems with Applications*, 38(11):13497–13504, 2011.
- [27] Vladimir Shapiro, Georgi Gluhchev, and Dimo Dimov. Towards a multinational car license plate recognition system. *Machine Vision and Applications*, 17(3):173–183, 2006.
- [28] Nahian Alam Siddique, Asif Iqbal, Fahim Mahmud, and Md Saifur Rahman. Development of an automatic vehicle license plate detection and recognition system for bangladesh. In *Informatics, Electronics & Vision (ICIEV), 2012 International Conference on*, pages 688–693. IEEE, 2012.
- [29] Young Sung Soh, Byung Tae Chun, and Ho Sub Yoon. Design of real time vehicle identification system. In *Systems, Man, and Cybernetics, 1994. Humans, Information and Technology, 1994 IEEE International Conference on*, volume 3, pages 2147–2152. IEEE, 1994.
- [30] Yushuang Tian, Kim-Hui Yap, and Yu He. Vehicle license plate super-resolution using soft learning prior. *Multimedia Tools and Applications*, 60(3):519–535, 2012.
- [31] Feng Wang, Lichun Man, Bangping Wang, Yijun Xiao, Wei Pan, and Xiaochun Lu. Fuzzy-based algorithm for color recognition of license plates. *Pattern Recognition Letters*, 29(7):1007–1020, 2008.
- [32] Wai-Hong Wong, Wan-Chi Siu, and Kin-Man Lam. Generation of moment invariants and their uses for character recognition. *Pattern Recognition Letters*, 16(2):115–123, 1995.
- [33] B. Hu X. Yang X. Ban Y. Wang, J. Chen. License plate recognition based on sift feature. *Optik - International Journal for Light and Electron Optics*, 2015.
-

- [34] Genlin Ji Preetha Phillips Yudong Zhang, Shuihua Wang. Fruit classification using computer vision and feedforward neural network. *Journal of Food Engineering*, 2014.
- [35] Yungang Zhang and Changshui Zhang. A new algorithm for character segmentation of license plate. In *Intelligent Vehicles Symposium, 2003. Proceedings. IEEE*, pages 106–109. IEEE, 2003.
- [36] Lihong Zheng, Xiangjian He, Bijan Samali, and Laurence T Yang. An algorithm for accuracy enhancement of license plate recognition. *Journal of computer and system sciences*, 79(2):245–255, 2013.
- [37] Dragiša Žunić and Joviša Žunić. Shape ellipticity from hu moment invariants. *Applied Mathematics and Computation*, 226:406–414, 2014.
- [38] Joviša Žunić, Kaoru Hirota, and Paul L Rosin. A hu moment invariant as a shape circularity measure. *Pattern Recognition*, 43(1):47–57, 2010.
-





## Apéndice A

# Placas de circulación actual en México

En la siguiente tabla se muestran imágenes de las placas de los diferentes estados de la República Mexicana empleadas en el presente trabajo.

|   |   |   |  |   |
|---|---|---|--|---|
|    |    |    |    |    |
| Aguascalientes  | Baja California   | Baja California Sur   | Campeche   | Chiapas   |
|  |  |  |  |  |
| Ciudad de México  | Chihuahua   | Coahuila  | Colima   | Durango   |
|  |  |  |  |  |
| Estado de México  | Guanajuato  | Guerrero  | Hidalgo  | Jalisco   |
|  |  |  |  |  |
| Michoacán   | Morelos   | Nayarit   | Nuevo León   | Oaxaca  |
|  |  |  |  |  |
| Puebla  | Querétaro   | Quintana Roo  | San Luis Potosí  | Sinaloa   |
|  |  |  |  |  |
| Sonora  | Tabasco   | Tamaulipas  | Tlaxcala   | Veracruz  |
|  |  |   |  |   |
| Yucatán   | Zacatecas   |   |  |   |

Tabla A.1: Placas de la República Mexicana



## Apéndice B

# Artículo

Derivado del presente trabajo, se escribió un artículo denominado: “License Plates Recognition of Mexican Private Vehicles Using Artificial Vision”, que fue enviado a 6th International Conference on Pattern Recognition Applications and Methods :

<http://www.icpram.org/>

The screenshot displays the Scimago Journal & Country Rank website for the ICPRAM 2017 conference. The page features a search bar at the top right and a navigation menu. The main content area shows the conference title, a large '2' indicating its ranking, and details such as 'Country: Portugal', 'Subject Area and Category: Computer Science', and 'H Index'. Below this is a banner for 'ICPRAM 2017' with the subtitle '6th International Conference on Pattern Recognition Applications and Methods' and the location 'Porto, Portugal' and dates '24 - 26 February, 2017'. A 'Call for Papers' section is visible, including a list of actions like 'On-line Registration' and 'Submit Paper', and a 'SCOPE' section describing the conference's focus on pattern recognition applications.

**Cuyo acuse de recibo se muestra a continuación:**

## ICPRAM 2017 - Paper Submission



PRIMORIS Information System

Hoy, 08:26 p.m.

Usted ▾



Responder | ▾

Dear Mr. NICOLAS TREJO,

Thank you for your submission. We would like to inform you that your paper has been successfully submitted to ICPRAM:

Title: License Plates Recognition of Mexican Private Vehicles Using Artificial Vision

Paper number: 113

Paper Area: Theory and Methods

Uploaded on Date/Time: 2016-11-17 8:24:00 PM

Submission type: Position Paper

Co-Authors: Farid García Lamont and Carlos Hiram Moreno

To access the ICPRAM conference management system, please go to the following address: <http://www.insticc.org/Primoris/> and enter your username and password.

We recommend that you use this possibility now, to check that all submitted information is correct. By using your author's area in the conference management system you have access to several options including re-submitting the paper or updating the list of co-authors. All co-authors will have access to all relevant information concerning the paper.

Please note that all papers submitted to ICPRAM will be peer-reviewed by at least two reviewers from the international program committee, who are independent of the conference editorial board.

Best regards,  
Marina Carvalho  
ICPRAM Secretariat

# *License Plates Recognition of Mexican Private Vehicles Using Artificial Vision*

**Keywords:** License plate recognition, Bayesian Classifier, Artificial Neural Network, Correlation Factor, Principal Component Analysis, Hu's Moments.

**Abstract:** In most of the works with the recognition of license plates from different countries assume that have white background, without texture patterns, with black characters. Mexican's vehicles plates are different because they have different texture's patterns and colors in the background, depending on the State of the Republic. That's why the recognition of these plates is not always successful, using the algorithms of work. This article proposes an algorithm for the recognition of vehicle registration of Mexican vehicles consisting of three parts: a) background - Character clearance: This phase is achieved using a threshold factor, which separate dark colors that form letters; in the clearings that are at the bottom. B) Characterization and modeling of symbols by technics of: Hu moments, Fourier's descriptors and correlation cross factor and C) classification, where were used comparison's techniques of template matching, Bayesian classifier and Artificial Neural Networks to process images of plates from different States. Results obtained are discussed at the end of the present work.

## 1 INTRODUCTION

Recognition of license plates of vehicles has been investigated in the world. Argentina (Nicolás Fernando Gazcón, 2011), Bangladesh (Siddique, 2012), China (Y. Wang, License Plate Recognition Based on SIFT Feature, 2015), Egypt (M.A. Massoud, 2013), India (Nasipuri, 2014), Japan (Cheng, 2004), Malaysia (Al Faqheri, 2009), among others. Normally works consist in four phases: 1) To get the image of the automobile 2) The plaque's location inside the image, 3) Characters extraction and 4) Classification or recognition of characters. In Mexican Republic, the main obstacle is located in the phase of extraction of characters because algorithms that exist in the literature assume that the plate has not texture patterns. The background usually is white and the characters black. However, in the case of Mexican plates have patterns of texture in the background producing algorithms work well for plates without patterns of texture are not always work correctly. On the other hand, each State and each new State Government can design an own pattern of background texture, this implies that there may be at least 32 kinds of plates, a number that increases with changes in government administration.

It is important to mention that, although each State can design the background of its plates, the dimensions of them plates and letters, as well as its style must meet with them features that designate the official Mexican standard, **NOM-001-SCT-2-2000**;

these characteristics are used to recognize the registration. This work proposes create an algorithm that segments characters and recognize depending of their characteristics of color and form. To segment characters properly, many of the patterns were removed of the background texture through the use of a threshold factor, which separate dark colors, forming the letters; clears, which form the background. Once filtered the background texture, the image of the plate was binarized and vertical and horizontal histograms were gotten using the technique of projection of profiles, just to obtain the coordinates of position used to segment characters. When the images of the characters were obtained, we proceeded to model them and characterize them using the techniques of: Hu moments, Fourier's Descriptors, and Cross-correlation factor.

Data obtained at this stage, were used as complementary in the classification phase, finally, in the stage of classification, techniques of templates, Bayesian classifier and artificial networks neural were used. Results obtained are discussed at the end of the present work.

The article is organized in the following form: in section 2 is presented the proposal of recognition of license plates. Experiments conducted in the section are presented on section 3. A discussion of the results obtained is made in section 4 and the article ends with conclusions in section 5.

## 2 PROPOSAL OF VEHICLE LICENSE PLATE RECOGNITION

Automatic Number Plate Recognition (ANPR) consists essentially of four stages: 1) To get the image of the car using a camera, 2) Extract the image plate, 3) Segment and remove the plate characters and 4) Recognize the characters extracted. Figure 1.

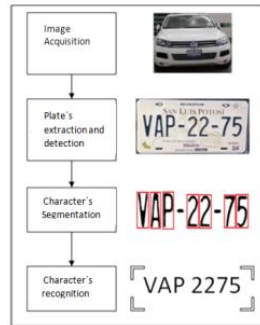


Figure 1: Stages of a system ANPR.

In the present work only discusses the segmentation and recognition of characters; steps (3) and (4). Images employed contain exclusively the region that conforms to the plate, and have the following characteristics:

- Images of the plates must be get between 1 and 1.5 meters of distance between camera and the registration.
- They must not have lighting variations.
- Images are frontal or near frontal, we mean, images must have very small rotation angles.
- No occlusions or considerable physical damages.

Now, the developed algorithm in this work is presented and consists in: 1) Standardization and binarization of the plate, 2) Segmentation of the characters, 3) Character modeling and 4) Recognition of characters.

### 2.1 Normalization and Binarization of the plate

Normalization of the images is based on the standard **NOM-001-SCT-2-2000**, where it is established that dimension's proportion of the plates

must be of 2: 1, so, all images are resized at  $700 \times 350$  pixels of being binaryzed.

The same standard also establishes the position in which the characters have to be collocated inside the plate, as well as the size and the distance between them, as it shows on the Figure 2.

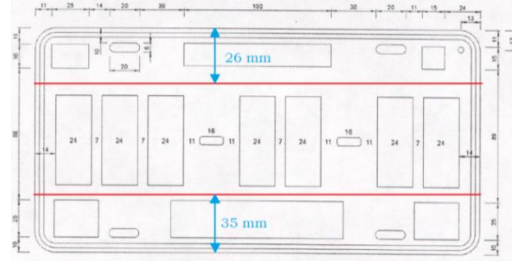


Figure 2: Position of the characters inside the registration.

With this information is eliminated the upper and lower sections of the original image, with the objective to get a section that contains exclusively characters. The result is shown on Figure 3.



Figure 3: Example of sections elimination upper and lower; (a) original image, (b) image obtained after sections elimination.

Due to the great variety of background colors in plates, before being binaryzed, was necessary separate the texture of the background of characters that comply the registration.

It is important to mention that, in the present work, the image's colors are represented in the space RGB. In this space the colors are represented as a linear combination of the vectors base of red, green and blue; the color of a pixel is represented as:  $\phi = [r, g, b]$ . Figure 4 shows the form of the RGB space.

To develop the separation's process, is obtained from an experimental form a threshold of separation that is in the magnitude of a vector RGB, starting from which is performed a comparison; to classify a pixel in a region or another, see Figure 4. The threshold ( $\delta$ ) gotten is of 0.73.

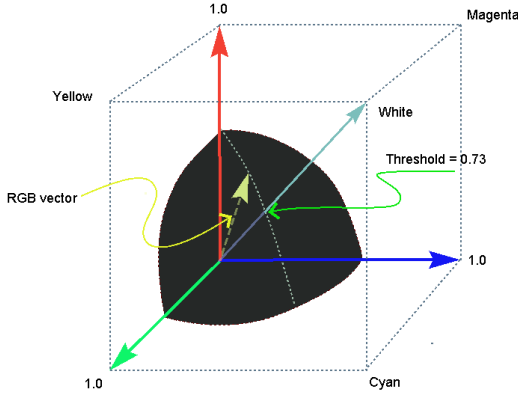


Figure 4: RGB Space and threshold of intensity.

On the figure 4, it is showed the part of the space where are found colors from the characters. Separation was made using the equation (1).

$$\theta^* = \begin{cases} \vec{1}, & \|\phi\| \leq \delta \\ \vec{0}, & \|\phi\| > \delta \end{cases} \quad (1)$$

Where  $\delta = 0.73$ ,  $\vec{1} = [1,1,1]$  y  $\vec{0} = [0,0,0]$ . With the proposed method, managed to differentiate the characters background was gotten and the possibility of binarize the images obtained. The Otsu method was used (Otsu, 1979). Figure 5 shows two examples of binarized images using the Otsu method.



Figure 5: Example of two images of plates binarized using our approach.

## 2.2 Characters Segmentation

The horizontal and vertical projection method is used to segment characters (Erick Cuevas, 2010), this method has been used for similar purposes in (Cheng, 2004),(Duan, 2005) and (Qin, 2006). Given an image  $I(x, y)$  wide  $N$ , height  $M$  and considering that  $1 \leq x \leq N$  and  $1 \leq y \leq M$  horizontal and vertical projections are defined as:

$$P_{hor}(y_0) = \sum_{x=1}^N I(x, y_0), \forall y = 1, \dots, M \quad (2)$$

$$P_{ver}(x_0) = \sum_{y=1}^M I(x_0, y), \forall x = 1, \dots, N \quad (3)$$

The histograms obtained by equations (2) and (3), allow determine the coordinates of the region which is each character. Figure 6 shows an example of horizontal and vertical projection of a binarized image.

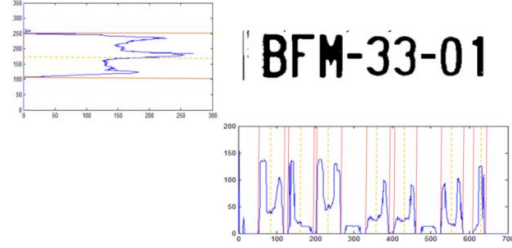


Figure 6: Horizontal and vertical projection of a binarized image.

Figure 7 shows some examples of segmented characters used-do horizontal and vertical projections.



Figure 7: Examples of characters segmented using horizontal and vertical projections.

## 2.3 Modelling of the characters

Segmented characters are modeled using the Correlation factor, Hu moments and Fourier's descriptors.

### 2.3.1 Correlation factor

Correlation is a statistical technique that quantifies the strength of the linear relationship between two variables. The quantification was performed using the coefficient of Pearson's correlation linear (A. Lind Douglas, 2012), whose value ranges between  $-1$  and  $1$ . The correlation coefficient  $r$  is calculated as:

$$r = \frac{\sum_{(x,y) \in O} (x - \bar{x})(y - \bar{y})}{\sqrt{\sum_{(x,y) \in O} (x - \bar{x})^2 \sum_{(x,y) \in O} (y - \bar{y})^2}} \quad (4)$$

### 2.3.2 Hu moments

Have been employees to recognize characters in (Wong, 1995) and (Gómez González, 2011), to measure features geometric as ellipticities (Zunic D. a., 2014) or circularities (Zunic J. a., 2010)

It's  $O = \{(x_1, y_1), \dots, (x_m, y_m)\}$  the set of pixels coordinates that have a character extracted. Hu moments are calculated as:

$$\begin{aligned}
 H_1 &= \eta_{20} + \eta_{02} \\
 H_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\
 H_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\
 H_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\
 H_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 \\
 &\quad - 3(\eta_{21} + \eta_{03})^2] \\
 &\quad + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} \\
 &\quad + \eta_{12})^2 \\
 &\quad - (\eta_{21} + \eta_{03})^2] \\
 H_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 \\
 &\quad - (\eta_{21} + \eta_{03})^2] \\
 &\quad + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\
 H_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 \\
 &\quad - 3(\eta_{21} + \eta_{03})^2] \\
 &\quad - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} \\
 &\quad + \eta_{12})^2 \\
 &\quad - (\eta_{21} + \eta_{03})^2]
 \end{aligned}$$

Where  $\eta_{pq}$  is the centralized and standardized moment gotten with:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{c+1}} \quad (6)$$

The equation (6),  $c$  and  $\mu_{pq}$  is calculated as:

$$c = \frac{p+q}{2}, \quad p+q = 2,3, \dots \quad (7)$$

$$\mu_{pq}(O) = \sum_{(x,y) \in O} (x - \bar{x})^p (y - \bar{y})^q \quad (8)$$

Where  $(\bar{x}, \bar{y})$  is the coordinate of the centroid of the objects, calculated as:

$$\bar{x} = \frac{1}{m} \sum_{(x,y) \in O} x \quad (9)$$

$$\bar{y} = \frac{1}{m} \sum_{(x,y) \in O} y \quad (10)$$

It's important to mention that Hu's moments are invariant to the position, rotation and scaling.

### 2.3.3 Fourier's Descriptors

It's  $\{(x_1, y_1), \dots, (x_n, y_n)\} \subset \mathbb{Z}^2$  the set of points that form the outline of a closed figure, each point is represented as a complex number.

$$z_n = x_n + jy_n, \quad \text{where } j = \sqrt{-1}. \quad (11)$$

The discrete transformed of Fourier (Steven L. Eddins Rafael C. Gonzalez, 2004.), of the set of points that has character's outline can be calculated as follows:

$$F(u) = \frac{1}{u\pi} \sum_{n=1}^m \varphi_n \left[ \cos\left(\frac{2\pi u s_n}{s_m}\right) - j \sin\left(\frac{2\pi u s_n}{s_m}\right) \right] \quad (12)$$

It can be written as:

$$F(u) = \frac{1}{u\pi} \sum_{n=1}^m \varphi_n \exp\left(\frac{-2j\pi u s_n}{s_m}\right) \quad (13)$$

Finally Fourier's descriptors are gotten to transform to calculate the values absolute of the complex numbers.

$$f(u) = |F(u)| \quad (14)$$

Where  $u = 1, 2, \dots, N$  and  $N$  is the total number of descriptors to obtain.

## 2.4 Characters Recognition

For characters recognition, are used three techniques: 1) Template matching, 2) Bayesian's classifier and 3) Artificial Neural Networks (ANN).

### 2.4.1 The template matching

The template matching is a technique which consists in comparing the image of the character that we want to recognize with a series of known templates to measure its similarity; of the results obtained, one that has the maximum values the symbol that contains the image to classify (Du, 2013). This technique has been used for the same purpose in (Gazcón, 2012) and (Y. Wang, 2015).

### 2.4.2 Bayesian classifier

The Bayesian classifier (Tom, 1997) is based on the Bayes theorem where is assumed that the vector of characteristics is a multivariate Gaussian distribution. It's  $C = \{k_1, \dots, k_n\}$  the set of  $n$  objects kinds, the probability that an object  $A$  is a kind  $k_i$  is denoted by  $p(k_i|A)$ . The Bayes theorem:

$$p(k_i|A) = \frac{p(A|k_i)p(k_i)}{p(A)} \quad (15)$$



For Bayesian classification, it's chosen the class  $ki$  which  $p(A|k_i)p(k_i)$  is bigger. This way the observed object  $\mathbf{A}$  is assigned the class  $ki$ . We think that the probability distribution of the characteristics of the observed object  $\mathbf{A}$  vector, is  $k_w$  kind, It's Gaussian with mean  $\mu_w$  and  $\Omega_w$  covariance matrix:

$$p(\mathbf{A}|K_w) = \frac{1}{\Delta} \exp \left[ -\frac{1}{2} (\mathbf{A} - \mu_w)^T \Omega_w^{-1} (\mathbf{A} - \mu_w) \right] \quad (16)$$

Where  $\Delta = (2\pi)^{m/2} (\det \Omega_w)^{1/2}$  and  $m$  is the vector's dimension of characteristics.

### 2.4.3 Artificial Neural Network

It is a mathematical model that is inspired by the way biological neurons work. The ANN (T. Hagan Martín, 2014) is applied in the learning of tasks where each instance to is described by a set of values, that represent its Special features and where the function objective  $f(\mathbf{A})$  can take any value from a finite set  $\mathbf{V}$ . In the present work, are implemented two ANN, a dedicated to the classification of numbers and another for letters.

The ANN deployed is of type Backpropagation, activated by the function sigmoid.

### 2.4.4 Principal Component Analysis

The Principal Components Analysis (PCA) is a technical in which a set of variables correlated are transformed to another set of variables not correlated that is a linear combination of the original variables in which most of these variables can be removed with minimal loss of the original information; This characteristic allows the PCA to be employed to reduce the dimensionality of a large set of data losing the least of information.

## 3 EXPERIMENTS

To perform the experimentation was employed a database of 70 images of plates from the 31 States of the Mexican Republic, except the Mexico's city, because this entity license plates are different (contain 6 characters instead 7). This database was chosen because it was considered the sampling by convenience. The images were captured at a distance of 1.5 to 3 meters between the camera and the plate, seeking a lighting uniform and subsequently normalized to  $700 \times 350$  pixels.

Experimentation was made in two phases:

- (a) Training of the recognizer's letters and numbers.
- (b) License plate recognition.

### 3.1 Training of the recognizer's letters and numbers

To test the recognizers, 738 characters, 374 letters and 364 numbers were used. On table 1 it showed the number of characters used in each case. That number isn't equal for all symbols since they do not appear with the same frequency in the license plate.

Three types of characters recognizers: 1) coefficient of correlation 2) Bayesian's classifier and 3) Artificial Neural Network (ANN) were implemented. These last, 2 networks were implemented; one for numbers and another for letters.

Table 1: Number of examples used during the training.

|    |    |    |    |    |    |    |    |    |    |    |
|----|----|----|----|----|----|----|----|----|----|----|
| A  | B  | C  | D  | E  | F  | G  | H  | J  | K  | L  |
| 14 | 9  | 11 | 17 | 16 | 18 | 26 | 21 | 20 | 13 | 14 |
| M  | N  | R  | P  | S  | T  | U  | V  | W  | X  | Y  |
| 13 | 14 | 20 | 17 | 12 | 18 | 18 | 21 | 13 | 15 | 18 |
| Z  | 0  | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  |
| 16 | 27 | 27 | 40 | 32 | 41 | 50 | 43 | 34 | 33 | 37 |

In the case of the Bayesian classifier and ANN, vector of dimension's characteristics 8 were used at first, formed by 7 Hu moments and correlation Factor, the results to classify the numbers and letters are showed on Table 2. In the present experiment, the networks are two layers with the following structure: [8, 10] for numbers and [8, 20] for letters.

Table 2: Classification's results with 7 Hu moments and the correlation Factor.

| Character | Bayesian classifier |       | ANN     |       |
|-----------|---------------------|-------|---------|-------|
|           | hits                | %     | hits    | %     |
| Numbers   | 246/364             | 67.58 | 333/364 | 91.48 |
| Letters   | 294/374             | 78.60 | 273/374 | 72.99 |

Were subsequently added vectors with characteristics 30 and 60 Fourier descriptors, forming vectors of dimension 38: 7 Hu moments, 30 Fourier descriptors and correlation factor, and dimension 68: correlation factor, 7 Hu moments and 60 Fourier's descriptors. Results obtained in the tests with the Bayesian's classifier can be observed

on table 3. In this experiment were also implemented following neural networks: for 38 features vector networks have the following components: letters = [38,76,23], numbers = [38,38,10], in the case of the 68 items vector, the networks implemented have the following structure: letters = [68,68,23], numbers = [68,34,10], the element's number by layer was retrieved from an experimental way. Both networks were classified correctly 100% of the examples.

Table 3: Bayesian classifier's results with 38 and 68 features.

| Character | Bayesian classifier |       | ANN     |       |
|-----------|---------------------|-------|---------|-------|
|           | hits                | %     | hits    | %     |
| Numbers   | 325/364             | 89.40 | 331/364 | 91.50 |
| Letters   | 360/374             | 96.4  | 369/374 | 98.7  |

Finally, the Principal Components Analysis was made which was achieved to reduce the dimension of the vector of characteristics in the following way:

- **38 features vector:** the numbers vector decreased to 9 components, while the letters vector went from 38 to 8 items. Neural networks implemented in this case, have the following composition: *numbers*: [9, 27, 36, 10], *Letters*: [8, 16, 32, 23].

- **68 features vector:** The numbers vector decreased to 13 components, while the letters vector went to only 11. Neural networks implemented in this case, have the following composition: *Numbers*: [13,26,13,10], *Letters*: [11,22,46,23]. The results obtained are shown on table 4.

Table 4: Bayesian classifier results with mean component analysis.

| Character | 38 Features |       | 68 Features |       |
|-----------|-------------|-------|-------------|-------|
|           | hits        | %     | hits        | %     |
| Numbers   | 325/364     | 89.40 | 331/364     | 91.50 |
| Letters   | 360/374     | 96.40 | 369/374     | 98.7  |

**Recognition of Vehicle's registration.** On the table 5 is shown the number of images of plates used by State. The results obtained on the test, the recognizers developed are shown on the figure 4.1. To consider recognition of the registration as a success, it has the following criteria: It says that a registration is successfully recognized if all their characters are classified correctly. This implies that having a character wrongly classified means that registrations aren't recognized successfully.

Table 5. Examples number of registrations by State.

|                |                 |                     |          |                  |
|----------------|-----------------|---------------------|----------|------------------|
| Aguascalientes | Baja California | Baja California Sur | Campeche | Chihuahua        |
| 2              | 1               | 0                   | 2        | 2                |
| Colima         | Coahuila        | Chiapas             | Durango  | Estado de México |
| 1              | 2               | 2                   | 1        | 3                |
| Guerrero       | Guanajuato      | Hidalgo             | Jalisco  | Michoacán        |
| 3              | 4               | 3                   | 4        | 4                |
| Morelos        | Nuevo León      | Nayarit             | Oaxaca   | Puebla           |
| 2              | 2               | 3                   | 2        | 2                |
| Querétaro      | Quintana Roo    | San Luis Potosí     | Sinaloa  | Sonora           |
| 4              | 1               | 4                   | 3        | 1                |
| Tabasco        | Tamaulipas      | Tlaxcala            | Veracruz | Yucatán          |
| 2              | 2               | 1                   | 2        | 2                |
| Zacatecas      | <b>Total</b>    |                     |          |                  |
| 3              | <b>70</b>       |                     |          |                  |

Results obtained in the classification of registrations, for all recognizers of characters implemented, can be observed on Figure 8.

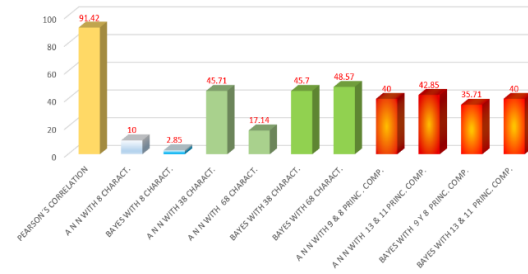


Figure 8: Classification's results of plates to different tree organizers implemented.

## 4 DISCUSSION

In the development of this proposal were implemented 3 types of recognizers: 1) correlation's factor, 2) Bayesian's classifier and 3) Artificial neuronal network, of these techniques, the correlation's Factor had the best performance as you can see on Figure 1, the main difficulty of this technique is to properly choose templates to compare, however, the occasions in which fails, is due to the segmented characters have an excess of noise or the shape and the overall size of the character to recognize is so approximate to the template's example. It happens with the number 1 and 7 for example. Using the Bayesian's classifier, when only 8 features, the percentage of hits are used to classify the vehicle's registration is only 2.85%, however, when it increased the number of features with Fourier's descriptors, the percentage has improved significantly, going to 45.7% with 38

features and 48.57% with 68 characteristics, being this the best value achieved.

When the main components analysis were deployed, the percentage declined on average by 8%, which shows that the criterion to use main components which accumulated variance is at least 95% which was achieved to decrease the characteristic's vector up to 88%.

In the case of neural networks, with 38 features was of 45.71%, whereas only reached 17.14% with 68 features, however, the analysis of the main components, to reduce the size of the vector of 38 features, performance decreased only 5%, on the other hand, to decrease the vector of 68 features, the percentage reached was increased to the 43.85%, which shows that increase the number of elements in the modeling of the characters, it isn't the best alternative.

Classification's Mistakes in the Bayesian's recognizer is presented in the following form: (a) Number 9 is classified as 6 and 4, 1 is recognized as 7 and 8 is classified as 0. (b) Letters: J and T are classified as L and H, respectively; and letters X and M are recognized as W.

The classification's mistakes in networks neural artificial, are: (a) Number 9 is recognized as 6 and vice versa, 5 as 3 and in some cases 4 as 9. (b) Letter R is classified as K; L is recognized as J; M is rated as W and vice versa; V is recognized as A.

In both classifiers, another mistake's source in the character recognition occurs in those whose general appearance or its form is similar as 5 and 8, H and K. on the other hand, it's important to remember that Hu's moment are invariant to rotation, because in several classification's mistakes are presented in images that are the same. For example: M and W, 6 and 9, V and A.

## 5 CONCLUSIONS

In the development of this proposal were implemented 3 types of recognizers: 1) correlation's factor, 2) Bayesian's classifier and 3) Artificial neuronal network, of these techniques, the correlation's Factor had the best performance as you can see on Figure 1, the main difficulty of this technique is to properly choose templates to compare, however, the occasions in which fails, is due to the segmented characters have an excess of noise or the shape and the overall size of the character to recognize is so approximate to the template's example. It happens with the number 1 and 7 for example. Using the Bayesian's classifier, when only 8 features, the percentage of hits are used to classify the vehicle's registration is only 2.85%,

however, when it increased the number of features with Fourier's descriptors, the percentage has improved significantly, going to 45.7% with 38 features and 48.57% with 68 characteristics, being this the best value achieved.

When the main components analysis were deployed, the percentage declined on average by 8%, which shows that the criterion to use main components which accumulated variance is at least 95% which was achieved to decrease the characteristic's vector up to 88%.

In the case of neural networks, with 38 features was of 45.71%, whereas only reached 17.14% with 68 features, however, the analysis of the main components, to reduce the size of the vector of 38 features, performance decreased only 5%, on the other hand, to decrease the vector of 68 features, the percentage reached was increased to the 43.85%, which shows that increase the number of elements in the modeling of the characters, it isn't the best alternative.

Classification's Mistakes in the Bayesian's recognizer is presented in the following form: (a) Number 9 is classified as 6 and 4, 1 is recognized as 7 and 8 is classified as 0. (b) Letters: J and T are classified as L and H, respectively; and letters X and M are recognized as W.

The classification's mistakes in networks neural artificial, are: (a) Number 9 is recognized as 6 and vice versa, 5 as 3 and in some cases 4 as 9. (b) Letter R is classified as K; L is recognized as J; M is rated as W and vice versa; V is recognized as A.

In both classifiers, another mistake's source in the character recognition occurs in those whose general appearance or its form is similar as 5 and 8, H and K. on the other hand, it's important to remember that Hu's moment are invariant to rotation, because in several classification's mistakes are presented in images that are the same. For example: M and W, 6 and 9, V and A.

## ACKNOWLEDGEMENTS

Thanks to CONACyT by the scholarship to get studies of mastery. We also thank to Dr. Eduardo Garea Llano to be a host his academic stay in the Advanced Technologies Applications Center (CENATAV), in Havana, Cuba. Thank To Liliana Sandoval Rios, for her collaboration to translate the present work.

## REFERENCES

- A. Lind Douglas, A. W. (2012). *Estadística Aplicada a los Negocios y la Economía*. Mexico D. F.: McGraw Hill.
- Al Faqheri, W. a. (2009). A real-time Malaysian automatic license plate recognition (M-ALPR) using hybrid fuzzy. *IJCSNS International Journal of Computer Science and Network Security* , 9 (2), 333-340.
- Cheng, Y. a. (2004). Car license plate recognition based on the combination of principal components analysis and radial basis function networks. *7th International Conference on Signal Processing* , 2, 1455--1458.
- Du, S. a. (2013). Automatic license plate recognition (ALPR): A state-of-the-art review. *IEEE Transactions on circuits and systems for video technology* , 23 (2), 311--325.
- Duan, T. D. (2005). Building an automatic vehicle license plate recognition system. *Proc. Int. Conf. Comput. Sci. RIVF* , 59--63.
- Erick Cuevas, D. Z. (2010). *Procesamiento digital de imagenes con MATLAB y Simulink*. México: Alfaomega Ra-Ma.
- Gazcón, N. F. (2012). Automatic vehicle identification for Argentinean license plates using intelligent template matching. *Pattern Recognition Letters* , 33 (9), 1066--1074.
- Gómez González, S. a. (2011). Desarrollo de un sistema prototipo de reconocimiento de dígitos usando momentos invariantes. Universidad Tecnológica de Pereira.
- Hu, M.-K. (1962). Visual Pattern Recognition by Moment Invariants. *IRE Transactions on Information Theory* , 179-187.
- M.A. Massoud, M. S. (2013). Automated new license plate recognition in Egypt. *Alexandria Engineering Journal* .
- Nasipuri, S. S. (2014). iLPR: an Indian license plate recognition system. *Multimedia Tools and Applications* .
- Nicolás Fernando Gazcón, C. I. (2011). Automatic vehicle identification for Argentinean license plates using intelligent template matching. *Elsevier* , 1066-1074.
- Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man and Cybernetics* , 62-66.
- Qin, Z. a. (2006). Method of license plate location based on corner feature. *6th World Congress on Intelligent Control and Automation* , 2, 8645--8649.
- Siddique, N. A. (2012). Development of an automatic vehicle license plate detection and recognition system for Bangladesh. *Electronics & Vision (ICIEV), 2012 International Conference on Informatics* , 688--693.
- T. Hagan Martín, B. D. (2014). *Neural Network Design*.
- Tom, M. M. (1997). *Machine Learning*. Ithaca, N. Y.: McGraw-Hill.
- Wai-Hong Wong, W.-C. S.-M. (1995). Generation of moment invariants and their uses for character recognition. *Elsevier* , 115-123.
- Wong, W.-H. a.-C.-M. (1995). Generation of moment invariants and their uses for character recognition. *Pattern Recognition Letters* , 16, 115--123.
- Y. Wang, J. C. (2015). License Plate Recognition Based on SIFT Feature. *Optik - International Journal for Light and Electron Optics* .
- Y. Wang, J. C. (2015). License Plate Recognition Based on SIFT Feature. *Optik - International Journal for Light and Electron Optics* .
- Zunic, D. a. (2014). Shape ellipticity from Hu moment invariants. *Applied Mathematics and Computation* , 226, 406--414.
- Zunic, J. a. (2010). A Hu moment invariant as a shape circularity measure. *Pattern Recognition* , 43 (1), 47--57.