

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ
FACULTAD DE CIENCIAS E INGENIERÍA



LOCALIZACIÓN Y RECONOCIMIENTO AUTOMÁTICO
DEL NÚMERO DE LA PLACA DE UN AUTOMÓVIL

Tesis para Optar el Título de:
INGENIERO ELECTRÓNICO

Presentado por:
KRISTIANS EDGARDO DIAZ ROJAS

LIMA – PERÚ
2006

RESUMEN

La presente monografía propone un sistema de localización e identificación automática de placas de vehículos, a partir del reconocimiento de los caracteres (letras y números) presentes en los mismos. A diferencia de otros sistemas existentes, esta propuesta de procesamiento digital de imágenes se basaría en las características de los vehículos existentes en el Perú, los cuales cuentan con dos formatos de placas: tres o dos letras, un guión, y tres o cuatro dígitos, respectivamente.

Este trabajo está dividido en cuatro capítulos. A manera de introducción se describen los requerimientos que justifican el desarrollo del sistema, en función al uso específico de diferentes aplicaciones prácticas. En el primer capítulo se señalan las características del sistema propuesto y las consideraciones que se toman en cuenta para delimitar los márgenes de su aplicación. En el segundo capítulo se describe detalladamente el desarrollo del sistema planteado a partir de dos procesos: entrenamiento y reconocimiento. Se indica que el sistema se basa en las características geométricas de los vehículos con la finalidad de determinar la localización central de la placa a partir del empleo de técnicas de procesamiento de imágenes y redes neuronales en el reconocimiento de los caracteres. Para ello, se incluye la descripción del proceso de entrenamiento inicial de las redes neuronales que permiten encontrar los parámetros calibradores del sistema y se señalan los reajustes realizados al sistema inicial con el fin de obtener mejores resultados. En el tercer capítulo se analizan los datos de las pruebas experimentales, mostrándose que no toda la tarea de reconocimiento se debe dejar a las redes neuronales y se señala la necesidad de adición de un bloque de confirmación de salida que permita mejorar la eficiencia del sistema. Por último, en el cuarto capítulo se exponen las recomendaciones y se termina con un apartado para las conclusiones en las que se incluye la verificación del funcionamiento óptimo del sistema propuesto.

ÍNDICE

DESCRIPCIÓN DEL REQUERIMIENTO

<u>CAPÍTULO 1: SISTEMA PROPUESTO</u>	1
1.1 Consideraciones	3
1.2 Localización: Primera intención.....	4
1.3 Reconocimiento de los caracteres del primer sistema.....	7
<u>CAPÍTULO 2: DESARROLLO DEL SISTEMA PLANTEADO</u>	12
2.1 Características de la imagen adquirida.....	13
2.2 Proceso de entrenamiento.....	14
2.3 Localización de la placa.....	15
2.3.1 Centro.....	16
2.3.2 Primera estimación horizontal.....	17
2.3.3 Estimación vertical.....	21
2.3.4 Segunda estimación horizontal.....	22
2.4 Umbralización.....	24
2.5 Ubicación de los caracteres.....	25
2.6 Segmentación.....	28
2.7 Reconocimiento de los caracteres.....	32
2.7.1 Creación de patrones.....	33
2.7.1.1 Normalización.....	33
2.7.2 Procesamiento de los patrones.....	34
2.8 Redes neuronales.....	35

2.9 Problemas en el desarrollo del sistema.....	39
2.9.1 Localización de la placa.....	40
2.9.1.1 Error en la etapa de estimación vertical.....	40
2.9.1.2 Error en la segmentación.....	41
2.9.1.2.1 Mejoramiento en la parte inferior de los caracteres.....	47
2.9.1.2.2 Mejoramiento del tamaño de los caracteres.....	48
2.9.1.2.3 Mejoramiento en los cortes extremos de la placa.....	49
2.9.1.2.4 Mejoramiento en los caracteres que rodean al guión..	50
2.9.2 Error en el reconocimiento.....	51
2.10 Diagrama de bloques del sistema propuesto.....	52
<u>CAPÍTULO 3: ANÁLISIS DE LOS RESULTADOS</u>	53
3.1 Análisis a los números.....	55
3.2 Análisis a las letras.....	61
3.3 Diagrama de bloques final del sistema propuesto.....	75
<u>CAPÍTULO 4: RECOMENDACIONES</u>	76
<u>CONCLUSIONES</u>	80
<u>BIBLIOGRAFÍA</u>	82
<u>ANEXO</u>	
Programas en MATLAB (CD)	

AGRADECIMIENTOS Y DEDICATORIA

Agradezco la paciencia y comprensión de familiares y amigos a quienes resté merecida atención durante las horas dedicadas a este trabajo de investigación. Espero que en sus resultados se vean justificadas y perdonadas mis ausencias. También quiero señalar mi deuda con el ingeniero Donato Andrés Flores Espinoza. Este trabajo compromete preocupaciones vinculadas al tema básico sugerido por él, desde aquí hago presente mi reconocimiento y agradecimiento por su interés, guía y apoyo iniciales sin los cuales no habría sido posible su logro.

El desarrollo de esta monografía ha sido posible gracias al estímulo constante brindado por mis padres, Richard y Elizabeth, y mi hermano Richard. A ellos, por el apoyo que me ofrecieron a lo largo de toda mi carrera universitaria y por el que me ofrecen a cada momento, va dedicada la presente investigación.

DESCRIPCIÓN DEL REQUERIMIENTO

El sistema vehicular actual requiere varias de las aplicaciones prácticas utilizadas por los sistemas de identificación digital de placas de rodaje. Muestra de ello es la localización y el reconocimiento de placas vehiculares utilizadas en situaciones como: el parqueo inteligente, el monitoreo de violaciones de tráfico, el control en el acceso de vehículos en áreas restringidas y en la detección de vehículos robados. Dichas aplicaciones se explican a continuación.

Parqueo inteligente. Existen más de 600 millones de vehículos¹ que transitan en todo el mundo y hacen necesaria la identificación de lugares para su estacionamiento. The International Parking Institute² estima que hay más de 105'200,000 espacios donde parquear en los USA, un número que cambia cada día y que demanda el uso de un sistema de parqueo inteligente. El problema principal es el tiempo que se pierde en la ubicación de un lugar para el parqueo, así como para recordar la ubicación donde se

¹ Cifra señalada por el estudio realizado en el año 2001 por Marina Stasenko. Se manifiesta la existencia de millones de vehículos que transitan en todo el mundo a la que se aúna el movimiento creciente de la creación de automóviles a gran escala en, por lo menos, una docena de países a nivel mundial, incluyendo a Estados Unidos de Norteamérica. En la realidad del parque automotor estudiada por dicha organización, se demuestra que el incremento del número de vehículos va en aumento creciente a nivel planetario. La información puede ubicarse en: Marina Stasenko – 2001, Number of Cars. The Physics Factbook. 29 de agosto del 2003. [Consulta: 25 de febrero del 2004]. <http://hypertextbook.com/facts/2001/MarinaStasenko.shtml>.

² International Parking Institute. Frequently Asked Questions. International Parking Institute. 2002. [Consulta: 18 de enero del 2004]. <http://www.parking.org/Value/FAQ/Default.aspx>

realizó. Para enfrentar este problema, se ha creado el sistema de parqueo inteligente, el cual consiste en la colocación de una pantalla en la puerta de ingreso de los vehículos, con el fin de visualizar los lugares libres u ocupados para el estacionamiento. En cada entrada de parqueo se tiene un tablero donde el conductor puede teclear la placa de su vehículo y obtener, automáticamente, la localización del mismo. Para ello, se dispone en cada espacio de parqueo con una cámara especial que reconoce la placa y la presencia del vehículo.

El objetivo de dicho sistema es agilizar y obtener una más rápida selección del sitio de parqueo, para así evitar conflictos internos y lograr un ahorro de tiempo en la búsqueda. Su importancia, radica justamente en que su utilización incrementa la eficiencia del servicio de parqueo al reducir el tiempo de búsqueda.

Monitoreo de violaciones de tránsito. Durante el siglo XXI la velocidad se ha convertido en un peligro socialmente aceptado y subestimado. En efecto, cualquiera acepta que “30 kilómetros por hora” es la velocidad a la que toca suelo quien se precipita peligrosamente desde 9 metros de altura, vale decir, de quien cae desde el techo de un tercer piso y muere. Sin embargo, pese al reconocimiento de ese peligro, se conduce sin temor a velocidades de 90 Km/hora. Producto de esa temeridad se obtienen como resultado miles de accidentes cada año. Una cifra que ejemplifica esta situación es la de los 25 mil 312 accidentes³ que se registraron en el 2004 como consecuencia del exceso de velocidad, un 33% del total de los accidentes de tránsito.

Ha señalado Martín Solari de la Fuente que la única solución viable frente a esta situación es la educación vial, especialmente, en la formación profesional de los

³ Solari De La Fuente, Martín. “Realidad Actual y Estrategias para la Prevención de los Accidentes de Tránsito”. En: Seguridad y educación vial. Touring Perú. 2004. [Consulta: 23 de enero del 2005]. <http://www.touringperu.com.pe/gxpfiles/sitio/ztema10.htm>

chóferes de los vehículos de servicio público de pasajeros, con el fin de que cumplan eficientemente su función. Mientras eso se cumple, se ha ido dando otras alternativas. Con la finalidad de disminuir la cantidad de accidentes de tránsito se han utilizado, por ejemplo, una serie de dispositivos que sirven para reducir la velocidad de los vehículos. Estos dispositivos son, generalmente, los rompe muelles, tachones (ojo de gato) y los sonorizadores. Lamentablemente, la realidad muestra que su colocación no ha sido el resultado de un estudio de tránsito, lo cual ha originado que no sólo persista el problema, sino que se produzcan nuevas dificultades.

En esa línea, con el fin de colaborar en la solución del problema, se da la utilización de cámaras de video bajo un nuevo enfoque: detectar los vehículos que no cumplen con las leyes de tránsito, principalmente las reglamentaciones sobre la velocidad y el respeto a las señales del semáforo como luz roja. Su uso efectivo, evidencia la falta cometida por el chofer al cual se le entrega la siguiente información: la identificación de la placa de su vehículo y la identificación de la hora, fecha y lugar donde cometió la falta. La visualización de un hecho hipotético como el señalado, se muestra en la figura I.



Figura I Detección de violación de la señal de tránsito

Todo lo anterior señala la urgencia de crear un sistema de localización y detección de la placa de un vehículo, con el fin de monitorear las violaciones de tránsito que acontecen día a día.

Control en el acceso a áreas restringidas. Existe un enorme incremento de la violencia, robos, terrorismo e inseguridad a nivel mundial. Frente a la falta de seguridad reinante, se hace necesaria la creación y establecimiento de un sistema de control que colabore con la erradicación de estos problemas socioculturales en ciertas áreas delimitadas. Un aporte de la tecnología electrónica con respecto al tema, es el que se establece por medio de los sistemas de control en el acceso a áreas restringidas. La utilización de estos sistemas supone la aplicación ordenadora de un filtro de control de ingreso o egreso de las personas a dichas áreas, lo cual implica el desarrollo de un sistema de monitoreo y seguridad que permite salvaguardar los peligros a los que podría estar expuesta una zona determinada. En esa medida, se ejercita un sistema de automatización para la libre circulación de peatones en un área específica con el fin de contrarrestar los peligros de seguridad a los que pudiera estar expuesta, como ocurre en el caso de las instituciones públicas o privadas.

En la actualidad existe una serie de dispositivos de control de acceso a áreas. Entre ellos tenemos: los postes de concreto (bollards), sistemas de uso frecuente que permiten limitar el acceso a locaciones específicas. Estos, al reconocer un vehículo autorizado, descienden automáticamente permitiendo su paso. También se encuentran las puertas (gates), dispositivos que utilizan un brazo que se eleva permitiendo el paso de los vehículos; además de las puertas especiales (speciality gates) y los bloqueadores de vehículos, ambos son utilizados para un alto grado de seguridad en instituciones del gobierno o instalaciones militares. Todos estos dispositivos pueden utilizar un sistema de localización y reconocimiento de la placa de un vehículo con el fin de agilizar el acceso. Una muestra gráfica de estos tipos de mecanismos de control de acceso a áreas se incluyen en la figura II.



Bollards

Energy Company, Northern California



Puerta

Pharmaceutical Corp.



Puertas especiales

BMW, Germany

Bloqueador de vehículos

Air National Guard Base, Ohio



Figura II Diferentes tipos de control de acceso a áreas restringidas

Detección de vehículos robados. Debido a la gran cantidad de vehículos robados que se presentan en diferentes ciudades, se tiene su detección como la aplicación más importante entre las que puede emplearse un sistema de localización y reconocimiento de placas de vehículos. Para ilustrar la importancia de la realización de esta tesis, en las figura III y IV se muestra la cantidad de vehículos robados a nivel de Lima y otras provincias peruanas. Los datos señalados están de acuerdo con los informes elaborados por la División de Prevención de Robo Vehicular (*DIPROVE-PNP*) en las fechas que van del 28 de agosto de 1987 a febrero del 2005.

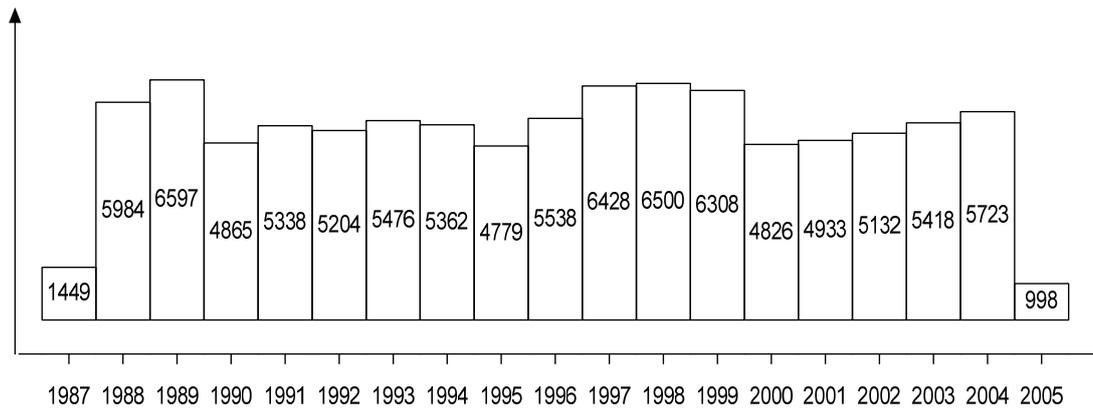


Figura III Vehículos recuperados en Lima del 9/87 al 2/05

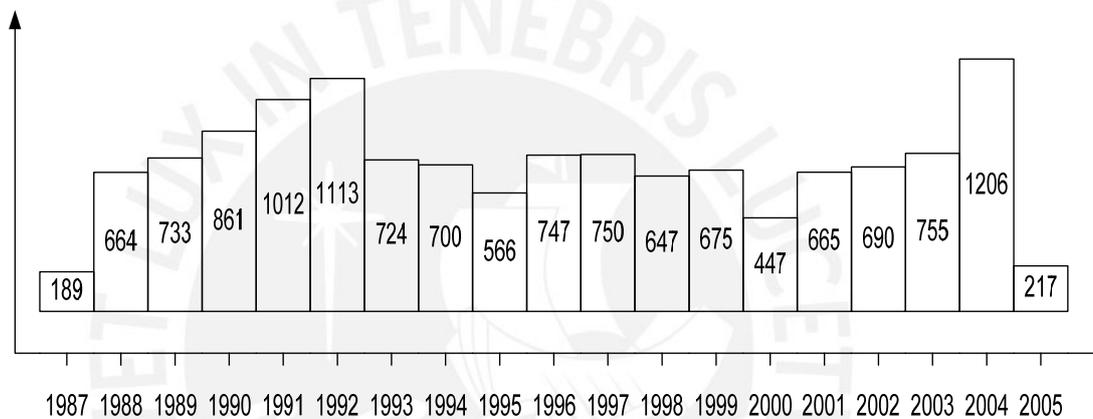


Figura IV Vehículos recuperados en provincias del 9/87 al 2/05

En las figuras V, VI y VII, que se muestran a continuación, se señalan el número de vehículos recuperados a nivel nacional y el porcentaje correspondiente por año, según la *DIPROVE-PNP*.

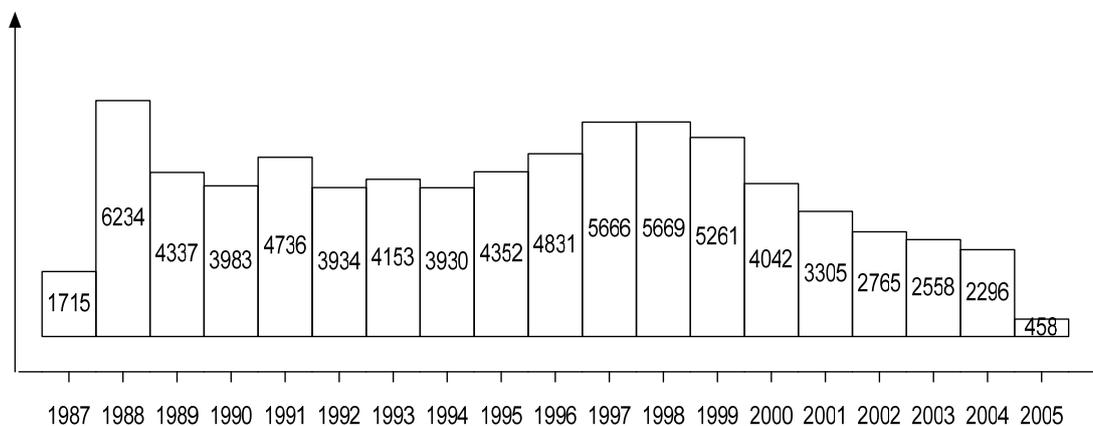


Figura V Vehículos recuperados en Lima del 9/87 al 2/05

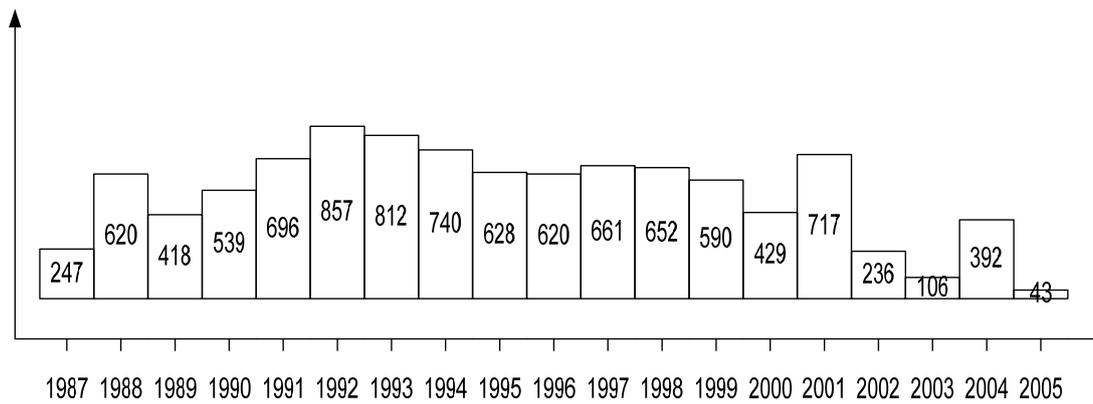


Figura VI Vehículos recuperados en provincias del 9/87 al 2/05

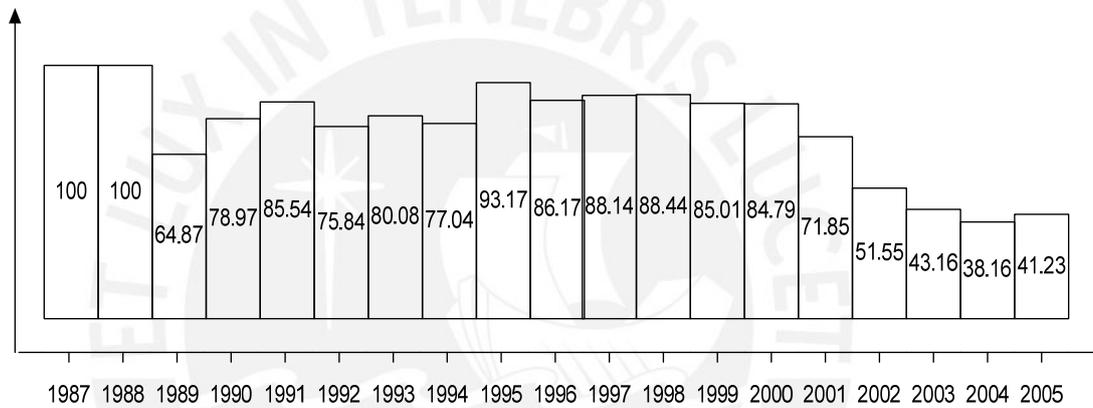


Figura VII Porcentaje de vehículos recuperados a nivel nacional del 9/87 al 2/05

En las gráficas anteriores se observa que el porcentaje de vehículos recuperados a nivel nacional en los últimos diez años ha tenido una baja significativa, con una disminución más notoria en los últimos cinco años (2002- 2005). Según la gráfica, durante ellos el número de automóviles robados en promedio es de 14 por día, mientras que el número de autos recuperados por día es de 8. Como se observa en cada una de las barras estadísticas, estas cifras tienden a aumentar y disminuir respectivamente.

Según el análisis realizado por la *DIPROVE-PNP*, la dificultad para recuperar vehículos robados se debería a que el mercado automotriz ha aumentado notablemente originando mayor dificultad en la recuperación de los vehículos robados. Por este motivo se hace necesario elaborar un sistema que ayude en la recuperación

de vehículos robados a partir de una rápida identificación de sus placas vehiculares. En esa línea, la presente tesis propone la creación de un sistema de localización y reconocimiento automático del número de la placa de un automóvil adaptado a los requerimientos de las placas vehiculares peruanas.

La identificación automática de vehículos robados, usando lógica difusa y redes neuronales basados en la identificación de sistemas [1]-[2], ha sido una principal fuente de investigación en los últimos años. Lo anterior se debe, principalmente, a dos importantes razones: el tiempo computacional y la disminución del error [3]-[4]. En la actualidad hay pocos sistemas basados en la localización de la placa y los que existen centralizan su actuación en el proceso de reconocimiento, sin enfatizar en el primer problema que es su localización, umbralización y segmentación (que son los pasos previos para el reconocimiento). Habría que señalar que también existen sistemas encargados de la ubicación de la placa, pero estos utilizan la transformada de Hough [5]-[6] y la Cuantización Vectorial [7], procedimientos matemáticos que no logran, en muchos casos, un buen nivel de éxito.

Tomando en cuenta los sistemas existentes en la actualidad, en la propuesta de este trabajo se procura revertir el fracaso de los sistemas precedentes, con el fin de mejorar el nivel de éxito y el óptimo funcionamiento de un sistema novedoso de localización y reconocimiento automático del número de la placa de un automóvil. Esta tarea implica la tarea de desarrollar un sistema capaz de diferenciar entre placas de tres letras y tres dígitos, y placas de dos letras y cuatro dígitos. Se debe tomar en cuenta, también, que actualmente las placas de automóviles que circulan en el Perú cuentan con 26 letras y 10 dígitos siendo estos: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z, 0, 1, 2, 3, 4, 5, 6, 7, 8 y 9. La identificación de estas placas, al estar localizadas en la parte inferior del vehículo, no es fácil, sobre todo en los casos en los que algunos de sus dueños utilizan una máscara de protección en los bordes. Dicha

acción dificulta la detección de los caracteres al producir una unión de los mismos, en algunos casos, y, en otros, una especie de sombras en la parte inferior de los caracteres. Además, en el caso de no contar con estas máscaras, algunas placas presentan deterioro en los bordes. Otro punto a tomar en cuenta en el desarrollo del nuevo sistema, es el hecho de que en el Perú la mayoría de los vehículos no presentan un correcto pintado en sus placas. Todo lo anterior, hace necesario advertir que cualquier sistema planteado debe tomar en cuenta estas características con el fin de establecer los dispositivos internos que le permita ser capaz de localizar y reconocer correctamente, no sólo las placas en buen estado, sino también las placas que no lo están, lo cual ocurre en un buen porcentaje de vehículos que transitan por el Perú.

En esa medida, el presente trabajo desarrolla un nuevo método de localización y reconocimiento de placas adaptado a las características de las placas de vehículos existentes en el Perú. La propuesta incluye la obtención de imágenes del vehículo en movimiento a partir de una cámara de video, la misma que luego es pasada a imágenes gráficas digitalizadas que permitirán analizar e identificar los caracteres de la placa. En este proceso, el movimiento de los automóviles se refleja en la imagen como ruido (la imagen presenta una apariencia de borrosidad), el cual es descartado durante el análisis digital que permitirá extraer los caracteres de la placa. Tanto las letras como los dígitos cuentan con varios factores que permiten o dificultan la captación de la imagen, estos son: la localización de la placa dentro de la imagen, la detección de la luminosidad, la distancia entre la cámara y el vehículo, la corrección de pequeñas inclinaciones de la placa, así como los pequeños giros, la verificación de las condiciones climatológicas, entre otros.

La idea es establecer un sistema capaz de solucionar los problemas de la captación de la imagen. Por ello, para reducir los efectos de los factores señalados

anteriormente, se extraen de cada carácter dos patrones que son procesados con el fin de obtener un tamaño previamente determinado; así como para corregir otros defectos de la imagen, como pequeñas inclinaciones o falta de pintura en la placa. Estos patrones serán utilizados para entrenar una red neuronal backpropagation o multicapa [8], la cual cuenta con: una capa de entrada, una sola capa oculta y una capa de salida. A la capa oculta se le han realizado pequeñas variaciones en lo que respecta a las funciones de activación, en vista de que se cuenta con demasiada información por cada carácter y es necesario evitar la saturación de la red. Para dicho efecto, se agregó a la función de activación dos factores más: los de desplazamiento e inclinación. Estos factores también son corregidos por la red al igual que los pesos.

Por último, cabe señalar que este sistema puede ser utilizado para varias aplicaciones, tales como: la seguridad y control de vehículos en un área restringida, la detección automática de vehículos que circulan a exceso de velocidad en las carreteras, la identificación de vehículos en áreas prohibidas, el parqueo inteligente, entre otros.

CAPÍTULO 1

SISTEMA PROPUESTO

En los últimos años, el campo del procesamiento digital de imágenes ha evolucionado enormemente. Dichos cambios se deben, principalmente, a la presencia de dispositivos capaces de capturar y digitalizar imágenes de una manera sencilla, y a la aparición de computadoras constituidas por microprocesadores cada vez más veloces que han ayudado a manipular mayor cantidad de información en menor tiempo. Esta evolución es de suma importancia para esta tesis, dada la gran cantidad de datos encontrados en la imagen digitalizada de una placa vehicular que necesita ser procesada al interior de un sistema de localización y reconocimiento de caracteres como el que se pretende proponer.

Los avances tecnológicos dados, también han permitido consolidar un nuevo campo dentro de las ciencias de la computación. Este progreso abarca a todo un conjunto de metodologías caracterizadas por su inspiración en los sistemas biológicos que sirven para resolver problemas relacionados con el mundo real (reconocimiento de formas, toma de decisiones, etc.), ofreciendo soluciones robustas y de fácil implementación. Este conjunto de metodologías emergentes comprende: la lógica borrosa, las redes neuronales, el razonamiento aproximado, los algoritmos genéticos, la teoría del caos y la teoría del aprendizaje. De entre todas estas metodologías, las Redes Neuronales Artificiales son las que actualmente están causando un mayor impacto, debido a su extraordinaria aplicabilidad práctica. Recientemente, dicha tecnología ha captado la

atención de los profesionales dedicados a la estadística y al análisis de datos, quienes comienzan a incorporar las redes neuronales al conjunto de herramientas estadísticas orientadas a la clasificación de patrones y la estimación de variables continuas.

Dados los anteriores progresos, se tienen ya las principales herramientas para el desarrollo de la presente tesis, la cual esta dedicada al procesamiento de imágenes mediante la utilización de redes neuronales. La diferencia con los sistemas preexistentes radicaría en que la propuesta necesita basarse en las características de los vehículos existentes en el Perú, ya mencionadas anteriormente, lo cual, pese a la tecnología, no es una tarea tan fácil. Los problemas principales para desarrollar un sistema como el que se propone son: localizar la placa dentro de la imagen, dado que no tiene un tamaño ni posición fija dentro de la imagen captada; y establecer la arquitectura de una red neuronal eficiente, ya que no existe ninguna receta para determinar el número de neuronas en la capa oculta, lo cual sólo puede determinarse en base a una serie de ensayos de prueba y error.

Cabe señalar que resulta más viable realizar cualquier tipo de tratamiento sobre la imagen (como la identificación, el reconocimiento, filtraje, etc.), únicamente cuando ésta ha sido sometida a una previa operación de pre-procesamiento. En particular, una operación de pre-procesamiento ayudará a la identificación de ciertos patrones (formas), que a su vez servirán de datos de entrada para automatizar un sistema con un fin determinado. En nuestro caso: con la finalidad del reconocimiento de caracteres. El objetivo final es que el sistema propuesto sea lo suficientemente robusto para que funcione correctamente ante los problemas que se presentan en las placas de rodaje, sea en el caso de su deterioro u otros problemas existentes en gran parte de las placas de vehículos existentes en el Perú.

1.1 Consideraciones

Se hace necesario tomar ciertas consideraciones para delimitar los márgenes de aplicación del sistema propuesto. Las que se toman en cuenta en la presente investigación son:

- Imágenes de color.
- Placa con ocho caracteres, siendo este el formato establecido en el Perú con tres o dos letras, el guión, y tres o cuatro números.
- Placa localizada en el centro de la parte delantera del vehículo.
- Fondo claro y caracteres en color oscuro. Esta consideración se toma en cuenta en el caso de las placas de color azul, amarillo, y blanco; ya que éstas cuentan con caracteres de color negro, mientras que las placas rojas utilizan caracteres blancos.
- Imagen frontal al vehículo.
- Caracteres identificables a simple vista. Lo anterior implica no tomar en cuenta placas cuyo reconocimiento sea difícil para el ser humano.

Como el sistema debe funcionar para casi todo tipo de placa, en las consideraciones señalamos que no se toma el hecho de que si las placas deben de tener los caracteres bien definidos, ni el tamaño de los mismos. Otro punto que no debe ser una limitación para el sistema es que existen placas que cuentan con un borde negro, ya sea por seguridad o por adorno. Tampoco se hace mención a la iluminación que debe tener esta área. Todo ello, dado que el sistema debe funcionar para casi todo tipo de placas menos para las que no cumplen con las consideraciones impuestas.

El sistema se subdivide en dos procesos: reconocimiento y localización. Inicialmente se propuso realizar la localización de la placa, para ello se ensayaron varios tipos de

sistemas en base a una serie de pruebas que fueron modificando la propuesta inicial hasta perfeccionarla. La base para la definición de la propuesta final es la que se presenta a continuación.

1.2 Localización: primera intención

La idea original era encontrar y extraer el área donde se ubican los caracteres de la placa de un vehículo dentro de una imagen capturada, con el fin de su posterior análisis (reconocimiento). Para ello, inicialmente se pensó en desarrollar un sistema capaz de detectar los bordes de la imagen, es decir, de encontrar las líneas que delimitan la placa de un vehículo. Con esto se pretendía resolver el problema de la ubicación de la placa, siempre y cuando todas tuvieran el mismo tamaño. En este proceso se empleó la transformada de Hough⁴ encontrándose todas las líneas verticales como horizontales que se tienen en la imagen procesada.

Para lograr lo anterior, todas las imágenes capturadas debían ser tomadas a una misma distancia, para así poder calcular de antemano el tamaño de la placa y realizar su más fácil ubicación. En la práctica este hecho no se da, debido a que no se puede predecir la forma de ingreso de un vehículo que va a ser captado por una cámara. En esa línea, tampoco se puede prever la distancia adecuada que debería tener la cámara con respecto al vehículo en análisis. Pese a la dificultad para determinar la distancia entre la cámara y el vehículo, consideramos que este problema no debía ser un límite para el sistema. No obstante, el problema se presenta como restricción para la ejecución del sistema y frente al mismo surgen otros. Es el caso de los faros de los vehículos que también tienen una distancia entre sus bordes similar a la de la placa, lo

⁴ La Transformada de Hough es un algoritmo empleado en reconocimiento de patrones en imágenes que permite encontrar ciertas formas dentro de una imagen, como líneas, círculos, etc. La versión más simple consiste en encontrar líneas.

cual origina un nuevo problema ya que dificulta la ubicación de la placa. Otro problema es el de la demasiada información que se tendría que analizar, ya que la transformada de Hough detecta bordes y las imágenes de los autos en su parte delantera esta muy recargada de estas características. Todos los problemas señalados llevaron a la necesidad de replantear las intenciones de la propuesta original y a pensar en otra alternativa que permitiera solucionar el problema inicial del sistema: la ubicación de la placa.

La otra alternativa de solución fue partir del centro del vehículo, dado que es una condición colocar la placa en dicho lugar central. Partiendo de la localización del centro del vehículo se procedió a realizar una estimación horizontal y vertical en base a las características propias de los mismos, obteniéndose un resultado positivo debido a la simetría que todos los vehículos presentan. Dicha simetría también favorece la detección del centro al producir una mayor variación de tonalidades en la placa (ya que los números son oscuros (negro) y el fondo claro). Lo anterior permite que al emplear una herramienta para la detección de bordes, el área que cuente con mayor grado de información sea la placa. Ver la ilustración de este proceso en la figura 1.1.

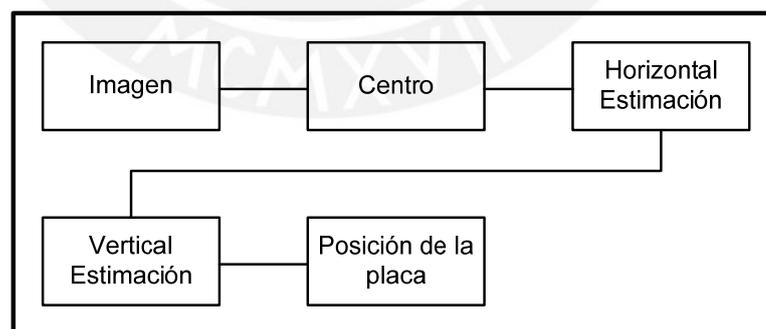


Figura 1.1 Diagrama de bloques de la ubicación de la placa

Las ventajas vislumbradas anteriormente, señalan a la ubicación del centro del vehículo como la base para la propuesta final del nuevo sistema. La explicación

detallada del mismo, junto a las ligeras variaciones realizadas, se describen en el capítulo correspondiente al desarrollo del sistema.

Otro punto importante en el sistema que se propone, es la segmentación de los caracteres. Inicialmente se propuso un sistema de segmentación basado principalmente en dos vectores y etiquetas. Dicho procedimiento se empleó al inicio de la propuesta de segmentación en base a pocas salidas de información; pero luego se cambió, dado que era necesario conocer el ancho, la altura, así como la ubicación de los caracteres de la placa dentro del área seleccionada. El proceso de segmentación y los ligeros cambios que permitieron ampliar el sistema y arrojar más salidas de información (alto, ancho y ubicación de la placa), también se describen en el capítulo correspondiente.

Antes de realizar la segmentación, se hace necesaria la implementación de una etapa previa: la umbralización de la imagen capturada. Para ello, se requiere que el sistema encuentre por sí solo el valor de umbral óptimo para su desarrollo, es decir, que en el histograma de la imagen obtenida (ver figura 1.2) se encuentre un valor apropiado que permita dividir los dos principales modos que corresponden a los diferentes niveles de gris.

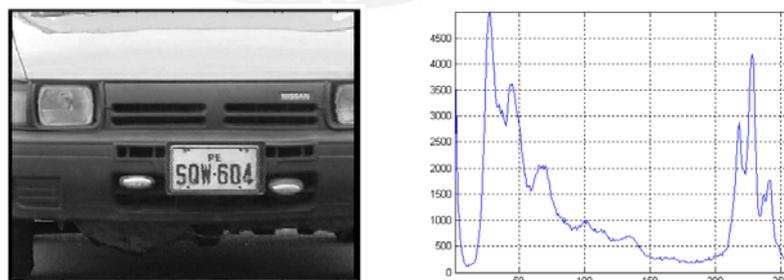


Figura 1.2 Histograma

La dificultad presentada consistía en que las imágenes no tenían por qué estar en las mismas condiciones de iluminación, ya que no es una condición la distribución homogénea de luz sobre el vehículo. Dado que no siempre se cumplen las

condiciones deseables de iluminación del vehículo, se tiene que la umbralización no puede ser un valor fijo sino uno adaptable a los diferentes histogramas.

Una solución planteada fue la de dividir la imagen en pequeñas regiones y luego analizar sus histogramas para así obtener el valor de umbral de cada región. Según este procedimiento, se originan diferentes valores de umbral para diferentes regiones con la finalidad de obtener una distribución homogénea. El inconveniente que se produce son los límites en las regiones, ya que puede darse el caso de que el salto de valor umbral de una región a otra sea muy grande y se provoque con ello una mala umbralización global de toda la imagen. Al final, la solución de la propuesta fue tomar el promedio de todas las regiones y mantener el resultado como el nuevo valor de umbral. Dicho procedimiento se describe a detalle en el siguiente capítulo.

1.3 Reconocimiento de los caracteres: primer sistema

Inicialmente se propuso un sistema que permitiera encontrar los caracteres según el error cuadrático medio. Con ese fin se desarrolló un sistema que cuenta con los bloques de morfología y comparación. A continuación se grafica la distribución de los mismos.

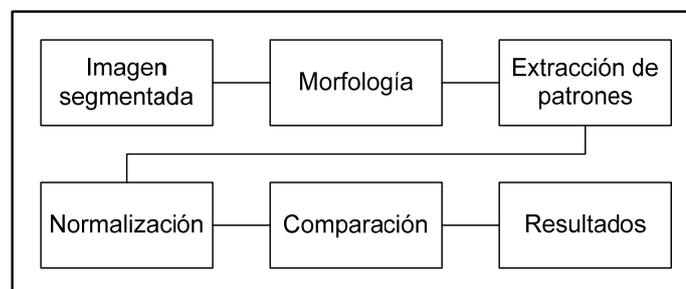


Figura 1.3 Diagrama de bloques del reconocimiento de los caracteres del primer sistema

El bloque de morfología tiene como misión corregir los desperfectos que se puedan tener como la falta de información en los caracteres debido a la falta de pintura. Para

ello se utilizó la dilatación, así como la erosión de la imagen. Se utiliza la dilatación debido a que expande una imagen permitiendo que los bordes de los caracteres tengan una forma más pareja; y la erosión, en la medida que permite contraer la imagen y regresarla a sus dimensiones iniciales.

El bloque de comparación se realiza como su nombre lo indica: comparando los patrones del carácter a reconocer con los patrones de los caracteres conocidos. Para ello se crea una base de datos formada por cuatro matrices: dos matrices para el reconocimiento de números y dos para el reconocimiento de letras. Los patrones se crean en función a la suma de las columnas y la suma de las filas como se muestra en la figura 1.4

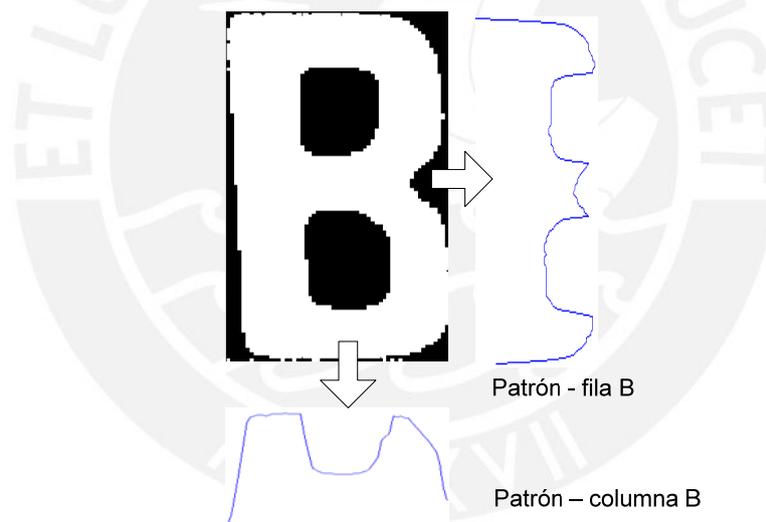


Figura 1.4 Creación de patrones

A cada carácter se le extraen dos patrones: columnas y filas. Como resultado se obtienen dos matrices para los números al igual que para las letras.

En este proceso se crea cada matriz ordenando todos los patrones obtenidos. Hay que tomar en cuenta que son 26 (de la A a la Z) para las letras y 10 (de 0 a 9) para los números. Estos patrones son tomados de los mismos caracteres que se encuentran en la placa, por este motivo la elección de los caracteres para formar la base de datos

del sistema es la tarea más delicada en este análisis, principalmente porque no todos los caracteres tienen el mismo tamaño. En esa línea es que se vuelve necesario el empleo del proceso de normalización.

La comparación se realiza uniendo los dos patrones que se extraen del carácter a reconocer y se compara con las matrices. El menor error (e_{menor}) nos indicará a que carácter corresponde. Se necesita señalar que no se juntan los números y las letras en una misma matriz, dado que en las placas ambos grupos se encuentran separados mediante un guión. Por ello, la presencia del guión indica que antes de él se debe utilizar el juego de matrices para las letras y, después, el otro juego de matrices destinado a los números.

$$e_0 = \sum |Mf(1) - f| + \sum |Mc(1) - g|, \quad e_{menor} = 1;$$

Para $j = 2 \text{ a } L$

$$e = \sum |Mf(j) - f| + \sum |Mc(j) - g|$$

If $e_0 < e$

$$e_{menor} = e_{menor}$$

else

$$e_{menor} = j$$

$$e_0 = e$$

End

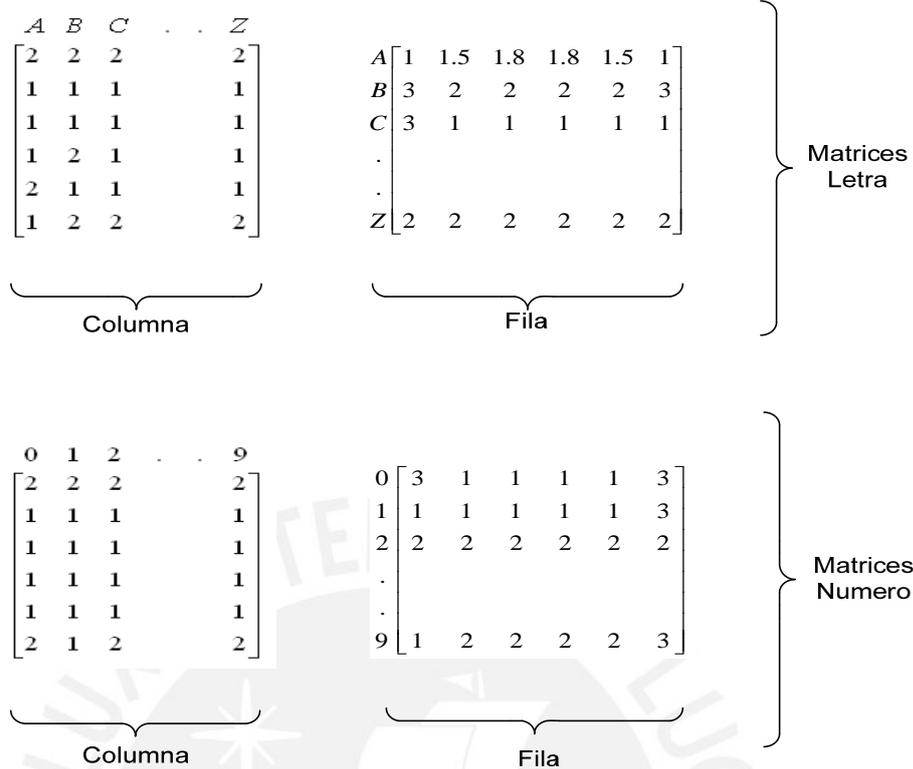
Donde:

f y g son los patrones que se extraen del carácter a reconocer,

Mf: matriz de fila

Mc matriz de columna

L: toma el valor de 26 ó 10 según se trate de una letra o número a reconocer



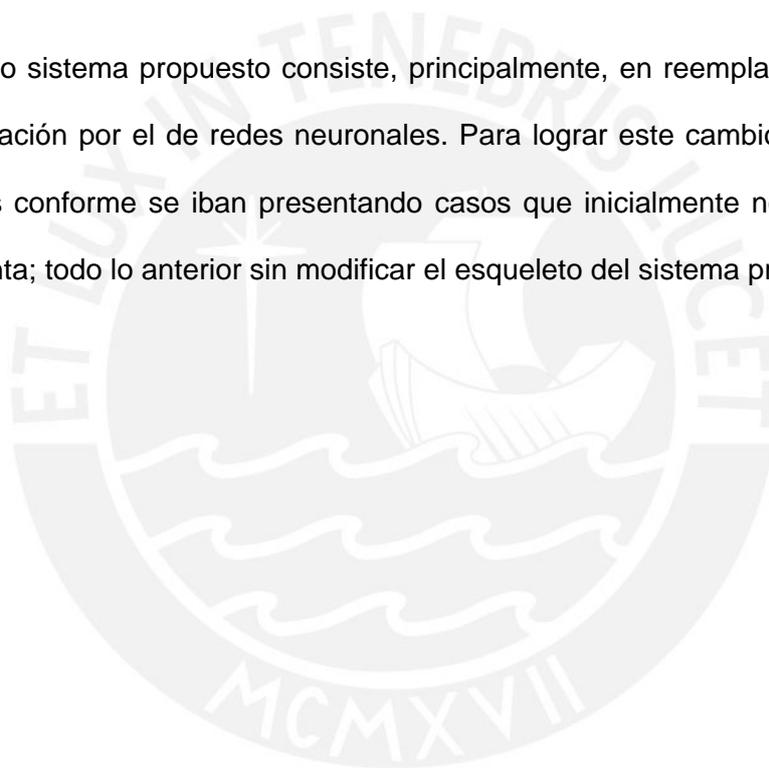
Los inconvenientes que surgen al emplear este procedimiento son:

- Inicialmente, al emplear morfología matemática, ocurre el problema de que al no saber el tamaño de los caracteres la dilatación puede ser muy potente y que al realizarse, en lugar de mejorar la imagen se obtenga una mancha imposible de identificar. Dicho problema surge especialmente en los caracteres 9, 6 y 8.
- Otro problema son los patrones que se deben considerar para la formación de las matrices, ya que estos servirán como guía para la comparación y posterior identificación del carácter. Lo anterior demanda tener cuidado y precisión en el criterio de selección de patrones.
- El error cuadrático medio se hace más notorio cuando los caracteres no tienen el mismo tamaño. Lo anterior se debe a que en el momento del proceso de normalización se pierde información del patrón, lo cual incrementa el porcentaje de error.

Al realizarse las pruebas utilizando este método y seleccionando los mejores patrones, se obtuvo un porcentaje de éxito igual a 49%. Dicho promedio es muy bajo en la

medida que manifiesta un alto porcentaje de error. El éxito no alcanzado implica que de dos caracteres de la placa uno se visualizaría mal. Por los motivos anteriormente expuestos y frente al bajo nivel de efectividad se propuso un nuevo método de reconocimiento de caracteres. Para lograr este objetivo se emplearon las redes neuronales de backpropagation con el fin de obtener mejores resultados y una mejor flexibilidad en variaciones posteriores. En suma, con la finalidad de mejorar los niveles de aciertos encontrados.

El nuevo sistema propuesto consiste, principalmente, en reemplazar un bloque: el de comparación por el de redes neuronales. Para lograr este cambio, se insertaron más bloques conforme se iban presentando casos que inicialmente no se habían tomado en cuenta; todo lo anterior sin modificar el esqueleto del sistema propuesto.



CAPÍTULO 2

DESARROLLO DEL SISTEMA PLANTEADO

El sistema planteado está compuesto por dos procesos principales: el proceso de entrenamiento y el proceso de reconocimiento. El proceso de entrenamiento consiste, primordialmente, en el desarrollo de una base de datos formada por patrones extraídos de los caracteres de la placa y en el entrenamiento de las redes neuronales multicapa utilizados en la identificación. Por otro lado, el proceso de reconocimiento está conformado por tres etapas: la primera consiste en la localización de la placa del vehículo dentro de la imagen capturada; la segunda en la umbralización de la placa de rodaje; y la tercera en la segmentación de los caracteres y el reconocimiento de los mismos a partir del uso de las redes neuronales multicapa cuyos parámetros son determinados durante el proceso de entrenamiento.

A lo largo de todo el presente capítulo se realiza el desarrollo del sistema planteado y el modo de funcionamiento de cada subproceso incluido en el mismo. Para empezar se detallan las características de las imágenes obtenidas y a continuación se explican los procesos de entrenamiento y reconocimiento señalados anteriormente. Después se desarrolla un apartado sobre las redes neuronales y se explican las características de los tipos de red utilizados en esta investigación. A continuación se incluye un apartado en que se señalan los problemas encontrados en el desarrollo del sistema, tanto lo que respecta a la identificación de sus errores como a la aplicación de las correcciones consecuentes. Por último se incluye un diagrama de bloques del sistema propuesto.

2.1 Características de la imagen adquirida

Las imágenes adquiridas pertenecen a vehículos que ingresan a la PUCP. Para la obtención de dichas imágenes se utilizó una cámara de video analógica mediante la que se realizaron grabaciones en cintas de video. Una vez realizadas las grabaciones, se pasó a realizar el cambio a formato digital mediante una tarjeta de captura de video y el programa controlador de la misma, logrando obtener una amplia secuencia de imágenes. Del grupo de imágenes obtenidas que contenían al mismo vehículo, se seleccionó de manera manual las mejores con el fin de su ulterior procesamiento. Habría que hacer hincapié en que se realizó la selección manual debido a que el alcance de esta tesis es la localización y detección de los caracteres, por lo que se deja para un proyecto posterior la realización de un módulo de captura y selección automática.

Las dimensiones de las imágenes adquiridas son: 480 píxeles de alto por 640 píxeles de ancho. El controlador de captura inserta un borde negro que no es de interés para el análisis por lo cual se elimina, originando una nueva imagen de 476 píxeles de alto por 623 píxeles de ancho. Habría que señalar que la imagen está representada por tres matrices (R, G y B) necesarias para otorgar el color correspondiente a un pixel. En el análisis se toma solamente una matriz, la matriz R, que es representada en diferentes tonos grises. Como resultado, la cantidad de información obtenida es menor con relación a una imagen digitalizada en formato de color.

La longitud de la placa dentro de la imagen debe comprender entre el 15% y el 40% de su longitud total. Lo anterior con el fin de que los caracteres sean completamente legibles y de que en la imagen se incluya una parte del vehículo, ya que esto ayuda a

la localización del centro, y en consecuencia a la posterior localización de la placa de rodaje.

2.2 Proceso de entrenamiento

El proceso de entrenamiento tiene como finalidad establecer el porcentaje de error de la red neuronal, así como encontrar los pesos de las neuronas del sistema implementado. Durante este proceso se elabora una base de datos conformada por patrones producidos por los caracteres de la placa. Para ello se crean tres bases de datos: una para el reconocimiento de los números, el cual cuenta con 10 elementos; y dos bases de datos para las letras, las cuales cuentan con 26 elementos posibles tomados del alfabeto.

En el proceso de entrenamiento se emplearon 80 patrones para los números, incluyendo ocho patrones de entrenamiento por cada número. En el caso de las letras se emplearon 180 patrones por cada una, dado que ya no se tiene la misma homogeneidad como sí ocurre en el caso de los números. La asignación de patrones a cada elemento se realizó según la dificultad de su detección. Al respecto, las letras O, Q, y B son las que cuentan con mayor cantidad de patrones, debido a la no tan marcada diferencia entre las mismas con relación al resto de los caracteres alfabéticos. Se tomaron 90 patrones para el entrenamiento de una primera red y los otros 90 para el entrenamiento de una segunda red. La explicación de este procedimiento se mostrará en la sección de reconocimiento de caracteres.

Habría que anticipar que cada patrón está formado por la sumatoria de elementos verticales como horizontales de los caracteres. La explicación detallada de este procedimiento, así como de los cambios realizados para la mejora de los patrones

también se realizará en la sección correspondiente al reconocimiento de los caracteres.

Con los patrones asignados se entrenaron las diferentes redes neuronales utilizando la configuración de la red backpropagation con una sola capa oculta [8]. Las entradas están compuestas por 110 neuronas, se utilizan 200 neuronas en la capa oculta y 26 ó 10 neuronas en la capa de salida para las letras y los números, respectivamente. El tamaño de las neuronas de entrada como las de salida son determinadas mediante el sistema, es decir, se cuenta con 110 datos de entrada, 26 neuronas de salida para la cantidad de letras a reconocer, y con 10 neuronas para los números del 0 al 9. Habría que señalar que no existe una técnica para determinar el número de neuronas en la capa oculta, este dato sólo se obtiene de forma experimental en base a ensayos de prueba y error.

2.3 Localización de la placa

El proceso de localización es una operación importante, debido a que si la posición de la placa no puede ser exactamente determinada durante el proceso, el reconocimiento de caracteres se vuelve más difícil y menos exacto de determinar.

La mayoría de los vehículos en el Perú localizan sus placas en la parte central e inferior de los mismos. En ese sentido, si la distancia entre la cámara y el vehículo varía, la ubicación central de la placa vehicular juega un rol importante para el proceso de localización del sistema. Tomando en cuenta estas referencias, esta tesis propone un método de localización de la placa basado en las características de la imagen capturada [9]. En la figura 2.1 se muestra dicho proceso, el cual consta de tres etapas: la detección del centro del vehículo, la estimación vertical y la estimación horizontal.

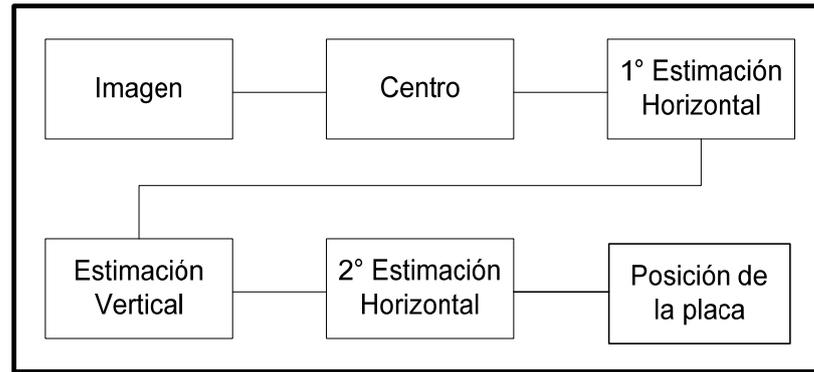


Figura 2.1 Proceso de localización de la placa

El desarrollo de este proceso y sus tres etapas se basa en las propiedades de la imagen capturada. A continuación se describe cada proceso:

2.3.1 Centro

Al no saber dónde se encuentra localizada la placa dentro de la imagen un buen punto de partida es encontrar el centro del vehículo, aprovechando la simetría de los vehículos y el hecho de que la placa se localiza en el centro de los mismos. Para lograrlo basta con encontrar la ubicación horizontal donde se produce la simetría.

Para encontrar el centro se emplea la siguiente fórmula:

$$centro = \min \left(\sum_{i=1}^l \sum_{j=1}^n \frac{|ima(i, j) - ima(x-i, j)|}{x} \right)$$

Donde:

n: altura de la imagen

l: longitud de la imagen

x: desplazamiento horizontal

j: coordenada vertical de la imagen

i: coordenada horizontal de la imagen

ima: la imagen capturada,

Para el desarrollo de la fórmula primero debe formarse la función $ima(x-i,j)$. Esta operación consiste simplemente en doblar $ima(i,j)$ alrededor del origen obteniendo $ima(-i,j)$ y luego desplazar esta función a un valor de x . A continuación, para cada valor de x se sustrae $ima(i,j)$ con el correspondiente $ima(x-i,j)$, se toma el valor absoluto y se le aplica la sumatoria de todos los elementos obteniéndose una nueva función. Como resultado el centro se localiza en el valor mínimo de la función. En la práctica no es necesario tomar todos los valores de j .

En la figura 2.2 se muestra como es posible dar la imagen de un vehículo completo gracias a la localización del centro y de la simetría de los vehículos.

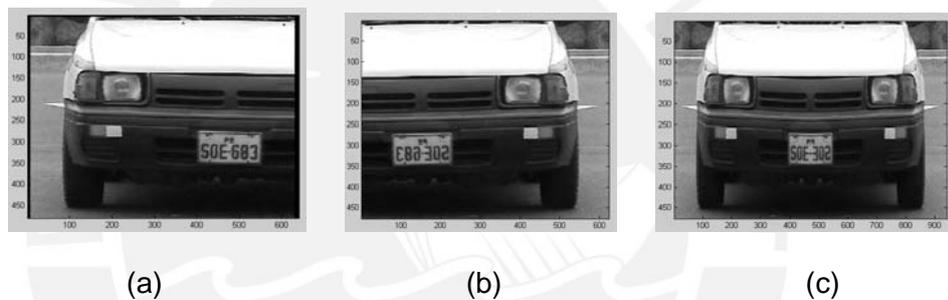


Figura 2.2 (a) Imagen original (b) Imagen invertida (c) Imagen formada al unir (a) y (b) por el valor centro

2.3.2 Primera estimación horizontal

Para la estimación horizontal se toman en cuenta las propiedades del vehículo. Una vez localizado el centro se procede a tomar una región de la imagen. Por lo general la imagen tiene una medida de 150 píxeles o menos a la derecha, y 150 píxeles o menos a la izquierda; lo anterior en el caso de que la placa se encuentre en los extremos de la imagen. Se toma como referencia 150 píxeles debido a que la mayor longitud de la placa será del 40%, es decir 250 píxeles de longitud. En total se toma 300 píxeles para estimar la ubicación de la placa (150 a cada lado) dando un margen de 50 píxeles en errar la ubicación del centro. Habría que señalar que la región de la imagen donde se

localiza la placa es la que tiene la mayor cantidad de transiciones de intensidad, ya que los caracteres son oscuros (negro) y el fondo de color claro.

El operador laplaciano es útil para delimitar los bordes de una imagen cuando se utilizan sus propiedades de paso por cero. Para ello se realiza la convolución de la región con el laplaciano de una función gaussiana bidimensional (ver figura 2.3 (a)) de la forma:

$$h(i,j) = \exp\left(-\frac{i^2 + j^2}{2\sigma^2}\right)$$

Sea $r^2 = i^2 + j^2$. El laplaciano de h es:

$$\nabla^2 h = \left(\frac{r^2 - \sigma^2}{\sigma^4}\right) \exp\left(-\frac{r^2}{2\sigma^2}\right)$$

Donde σ es la desviación estándar

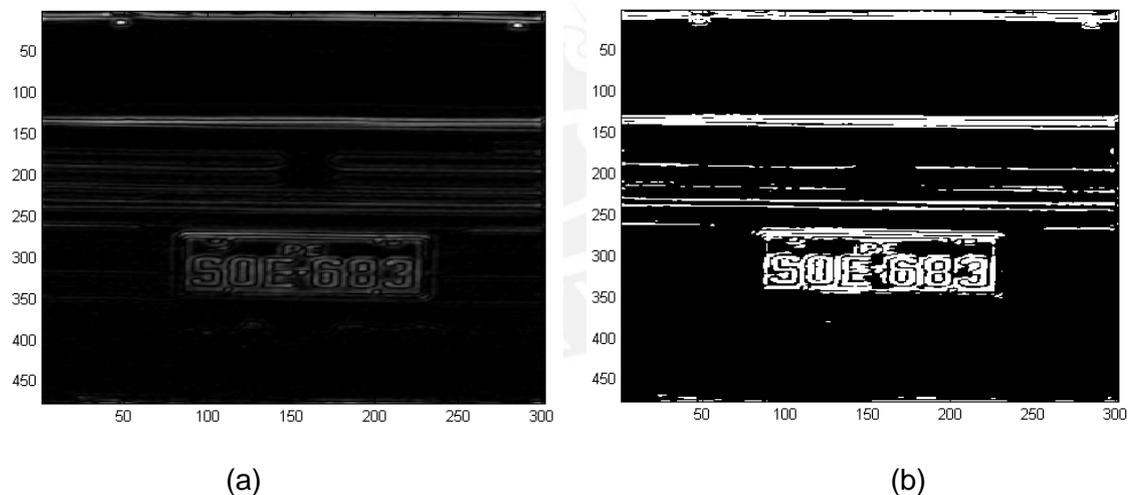


Figura 2.3 (a) Convolución entre la imagen y el Laplaciano;
(b) resultado de volver binario a (a)

Una vez que se tiene la imagen en binario a partir del uso de la umbralización como se muestra en la figura 2.3 (b) (se explicará más adelante la realización de este paso) se procede a acotar horizontalmente la imagen y se construyen las siguientes funciones:

Para $i=1, \dots, d$

$$j=1, \dots, 476$$

$$x(j) = |imd(j, c-d) - imd(j, c-d+i)|$$

$$f_D(i) = \sum_{j=1}^{476} x(j)$$

Para $i=1, \dots, a$

$$j=1, \dots, 476$$

$$x(j) = |imd(j, a) - imd(j, a-i+1)|$$

$$f_I(i) = \sum_{j=1}^{476} x(j)$$

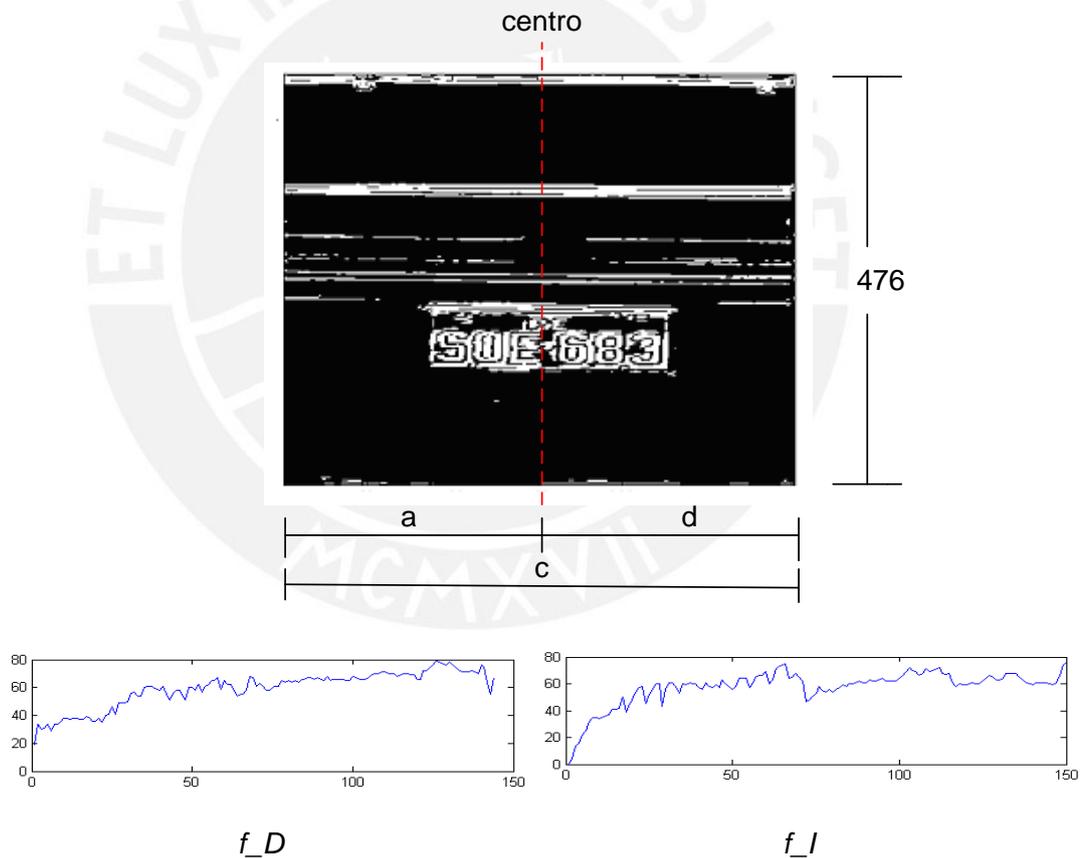


Figura 2.4 f_D Función para delimitar el margen derecho
 f_I Función para delimitar el margen izquierdo

De las gráficas presentadas en la figura 2.4 no se puede realizar ninguna estimación, lo que hace necesario realizar una etapa previa. Se procede a dilatar la imagen

utilizando para ello morfología matemática [11], con la finalidad de obtener el bloque que representa la placa (ver figura 2.5) y hacer más fácil su extracción. Para la realización de esta tarea se utilizan las ecuaciones anteriormente presentadas.

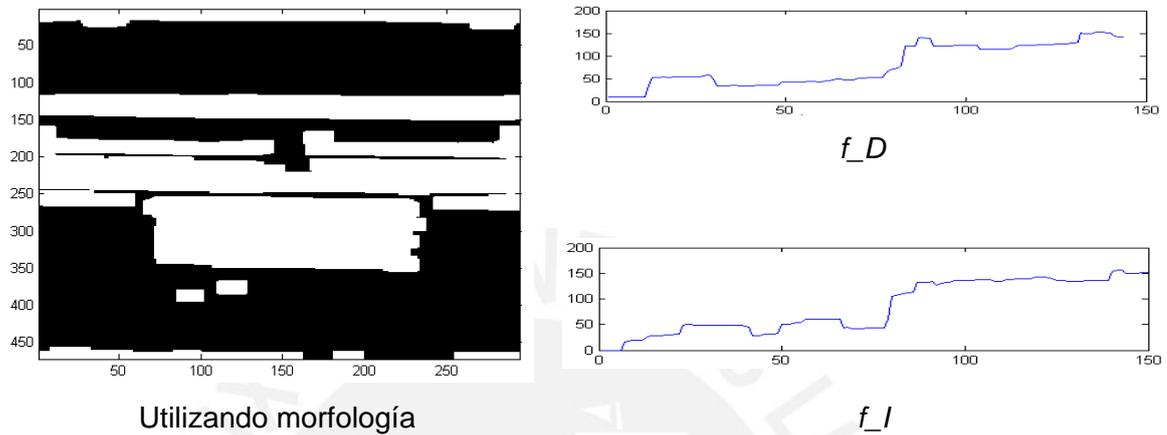


Figura 2.5 Empleo de morfología matemática- dilatación

Dada la gráfica, se hace bastante claro observar donde se debe realizar el corte para tener la estimación horizontal de la placa. El corte debe ser tomado en el cambio más brusco de amplitud de las funciones: f_D y f_I .

Para la realización de este corte se empleó una red neuronal multicapa con una sola capa oculta. La arquitectura empleada es 150-180-1, la cual incluye una repartición de 150 neuronas en la capa de entrada, 180 neuronas en la capa oculta y 1 neurona en la salida. Más adelante, en la sección 2.8, se explica con mayor detalle las redes neuronales. A continuación en la figura 2.6 y 2.7 se muestran el resultado y las etapas de la estimación horizontal, respectivamente.

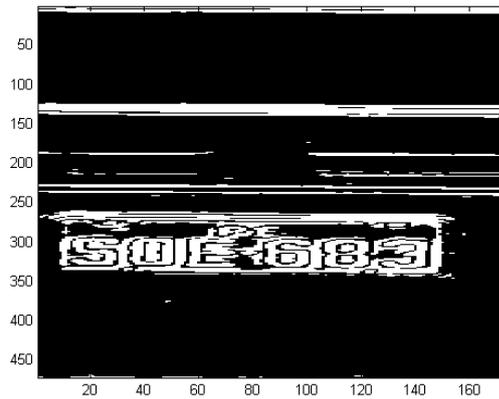


Figura 2.6 Placa estimada horizontalmente

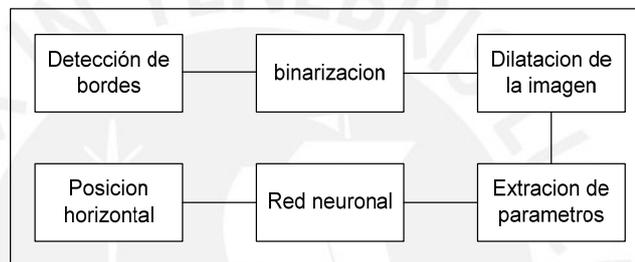


Figura 2.7 Etapas para la estimación horizontal

2.3.3 Estimación vertical

Así como en el caso de la estimación horizontal, en la estimación vertical también se toma en cuenta las propiedades geométricas del vehículo. Como ya se tiene una idea del ancho de la placa es fácil determinar cual sería su altura aproximada gracias a la proporción que se tiene entre la altura y el ancho de la placa (1:2). Para su ubicación se toma en cuenta que las placas están localizadas en la parte inferior del vehículo y debajo de ellas sólo está el piso.

Se comienza a realizar el barrido desde abajo hacia arriba hasta encontrar una región significativa que permita obtener una aproximación de las distribuciones entre el ancho y la altura (ver figura 2.8). El inicio y fin de la región dan las coordenadas.

Como se observa, el uso de Laplaciano ayuda notablemente en la localización de la placa y éste es un hecho que justifica su utilización.

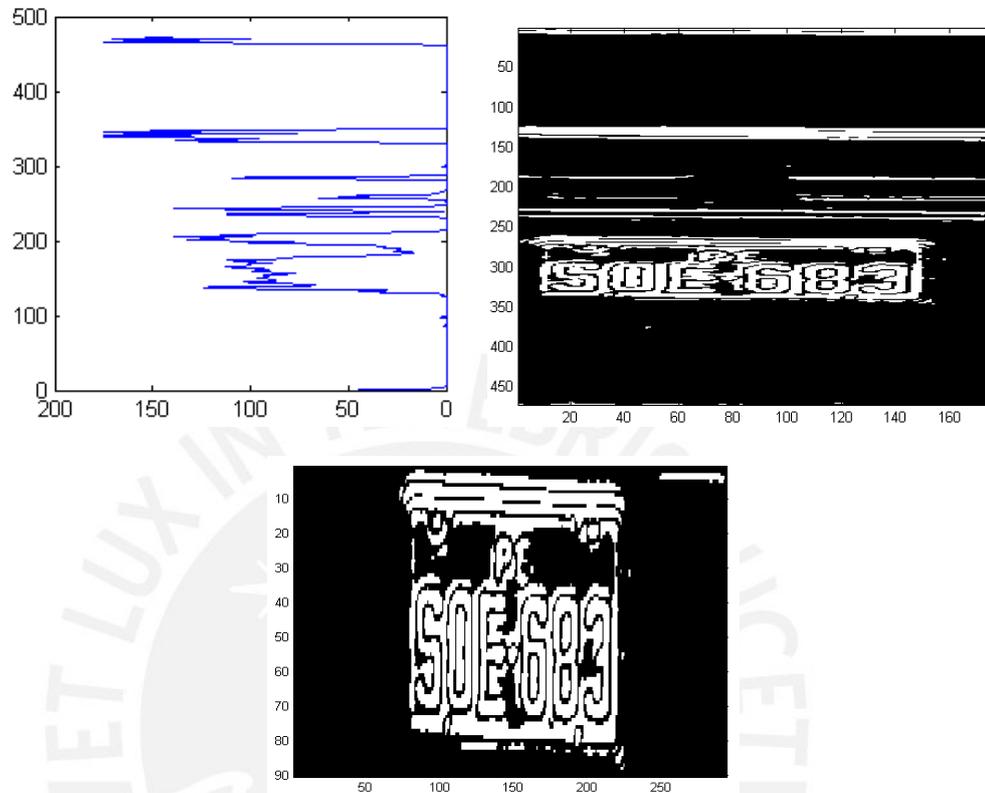


Figura 2.8 Estimación vertical

2.3.4 Segunda estimación horizontal

Se tiene localizada la placa, pero se debe realizar mejoras para poder obtener una más exacta ubicación de la placa. En la primera estimación se empleó dilatación de morfología matemática, originando que los valores obtenidos para la estimación no fuesen los más apropiados. Por ello la necesidad de realizar otra aproximación.

Tomando la imagen producto de la estimación vertical, como se muestra en la fig. 2.8, se procede a determinar las posiciones horizontales de la placa y a continuación se construye una gráfica que representa la cantidad de información que tiene cada columna (ver figura 2.9).

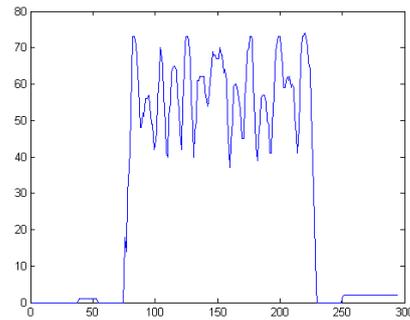


Figura 2.9 Información de las columnas de la imagen anterior

De la gráfica anterior se extraen fácilmente las posiciones de inicio y final de la placa. A continuación se muestran la placa ya extraída de la imagen y sus coordenadas respectivas.

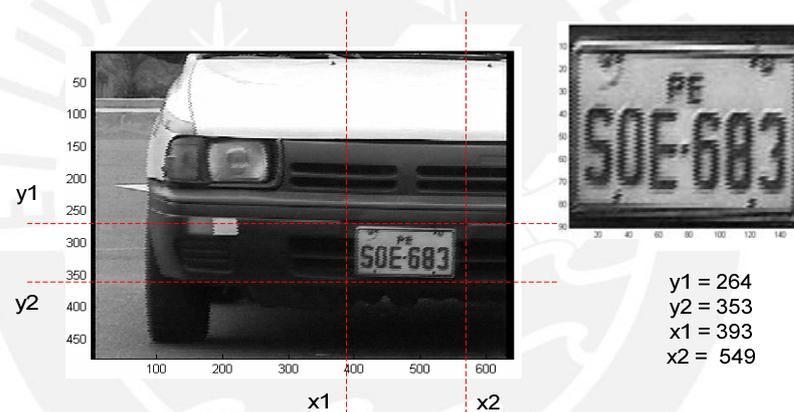


Figura 2.10 Placa extraída de la imagen de un automóvil y sus coordenadas correspondientes.

Para el procedimiento anteriormente descrito se tiene un éxito de 70%. En este análisis el error se debe al factor climático, principalmente a los días soleados en que se grabaron las imágenes. En esa línea, al aplicar la detección de bordes (Laplaciano) la imagen arroja bordes en el suelo (sombras provocadas por el sol) produciéndose regiones significativas con las mismas distribuciones de la placa. Por ese motivo se agregó una etapa más para discernir entre piso y placa logrando reducir el nivel porcentual del error hasta en un 14% (éxito de 86%).

Para finalizar, es necesario señalar que para esta etapa de la segunda estimación horizontal sólo se toma en cuenta la mitad de la imagen, es decir, la mitad de la altura de la imagen. Por lo tanto, si la amplitud del manchón localizado es mayor al 92% del ancho de la imagen, ésta corresponderá al suelo; de no ser así la imagen obtenida será la de la placa vehicular.

2.4 Umbralización

Se requiere que la imagen original en tonos de grises (255 niveles) sea modificada a sólo dos valores (1 y 0). De allí la necesidad de implementar una etapa de umbralización. Esta etapa busca encontrar un valor tal que persiga la siguiente fórmula:

$$f(x, y) = \begin{cases} 1 & f(x, y) \geq u \\ 0 & f(x, y) < u \end{cases}$$

Las imágenes obtenidas fueron tomadas bajo la luz natural por lo que no se tiene una distribución uniforme de la luz.

Para la determinación del valor de umbralización se precede a dividir la imagen en varias regiones (ver figura 2.11). A cada región se le extrae un valor, este es el promedio existente entre el promedio de los valores que están por encima de él y el promedio de los que están por debajo; una vez que se tienen los valores de todas las regiones se procede a sacar la media, obteniéndose como resultado el valor de umbralización (u).

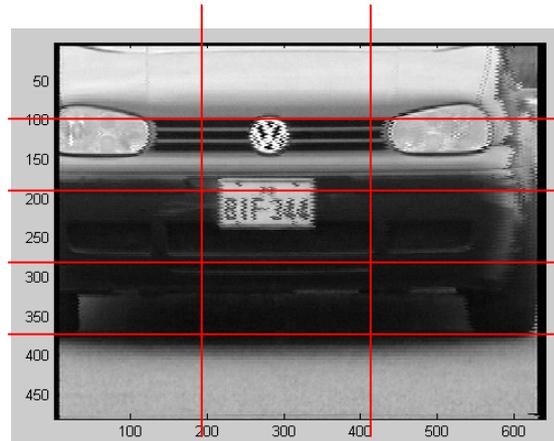


Figura 2.11 División de la imagen en regiones

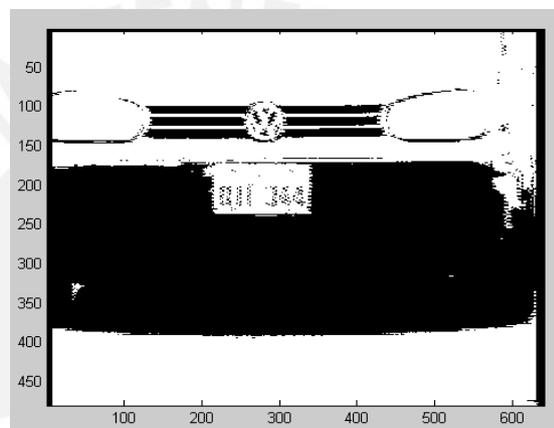


Figura 2.12 Imagen umbralizada aplicando el método descrito

2.5 Ubicación de los caracteres

Una vez que se tiene localizada la ubicación de la placa del vehículo, el siguiente paso es localizar los caracteres dentro de esta estimación, ya que estos constituyen el área de interés del proceso de análisis posterior: el reconocimiento de la placa vehicular.

Así, para lograr la ubicación o localización de los caracteres nos valemos nuevamente de las estimaciones obtenidas en función a las características del vehículo. En el caso de la placa de un vehículo, se toma en cuenta que los caracteres siempre se encuentran debajo de las letras “PE”.

A continuación se describirán los pasos planteados en la localización de los caracteres de la placa:

- Primero: se tiene la imagen estimada en binario
- Segundo: como las imágenes fueron capturadas utilizando una cámara de video con los vehículos en movimiento, se hace necesario mejorar estas imágenes. Para lograr esta mejora, se crea una nueva imagen empleando una máscara [11]. La realización de este paso consiste en recorrer la imagen tomando dos píxeles consecutivos verticalmente para tomar el mayor valor y colocarlo en la nueva imagen.

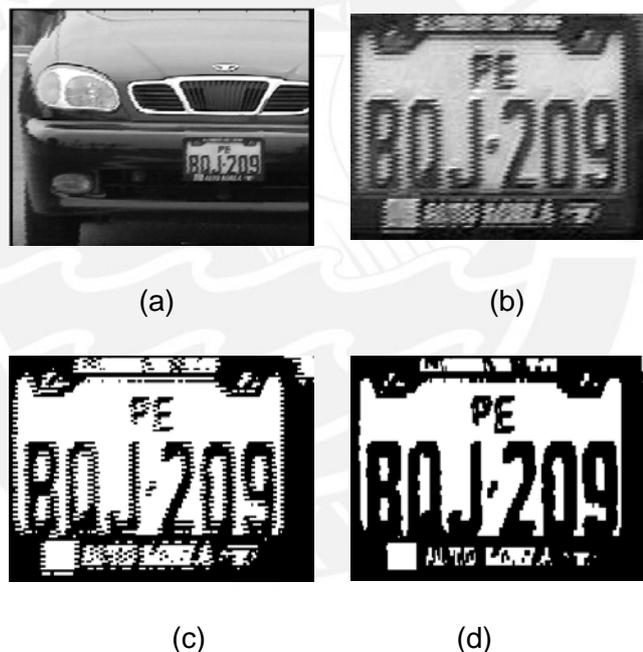


Figura 2.13 (a) Imagen capturada (b) Placa estimada (c) Placa umbralizada
(d) Aplicación de la máscara desarrollada a (c)

- Tercero: se realiza la sumatoria a las columnas, obteniéndose en los extremos la mayor cantidad de información producto del margen negro ocasionado por la máscara, la cual señala el lugar en donde se debe realizar los cortes horizontales. En el caso de que no se cuente con el margen en la placa, se

realiza la búsqueda de menor información en los extremos, pero antes de realizar este paso es necesario aplicar dilatación a la imagen con el fin de eliminar la PE y así evitar errores en el corte como el que se encuentra en la figura 2.14 (a). En dicha imagen se muestra el posible error.

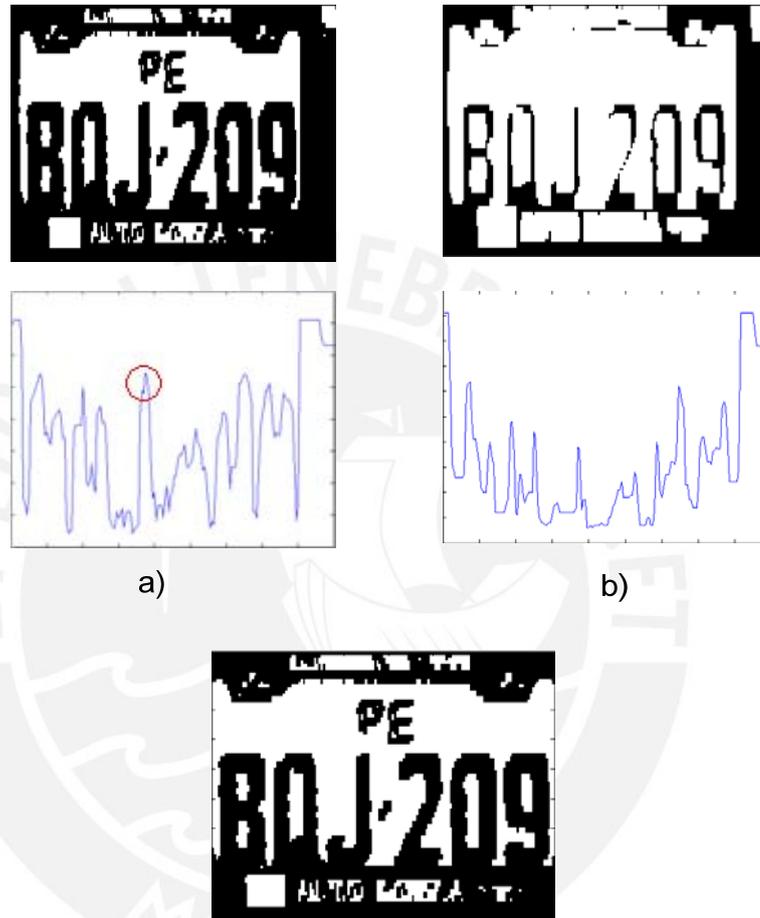


Figura 2.14 Ubicación horizontal de los caracteres

- Por último se realiza la sumatoria de las filas como se muestra en la figura 2.15 y se busca la PE. Se sabe que una de las características de las placas en el Perú es ubicar los caracteres debajo de la PE. En la última parte de este procedimiento se toma la mayor cantidad de información en caso de tener un margen negro, y la menor cantidad en caso de no tenerlo.

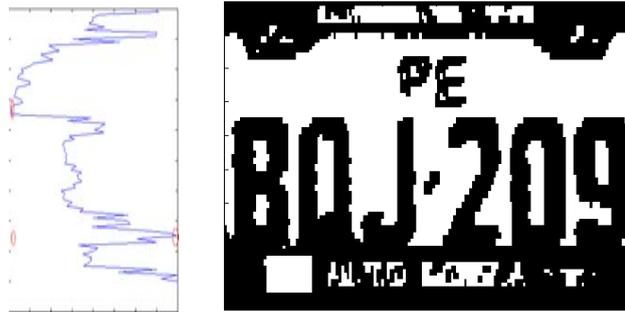


Figura 2.15 Sumatoria de filas



Figura 2.16 Ubicación de los caracteres

2.6 Segmentación

Después que la posición de la placa ha sido estimada dentro de la imagen capturada, la parte de la imagen del vehículo correspondiente a la sección de la placa es representada usando sólo dos niveles de intensidad (cero y uno). Con este fin se hace uso de la umbralización y luego se extraen los caracteres de la placa.

El proceso de segmentación consiste en subdividir una imagen en sus partes constituyentes u objetos. Esta sólo se realiza en la placa ya que es el área de interés para el análisis. Los objetos de interés a tomar en cuenta son las letras, los números y el guión. En esa medida, los objetos no tomados en cuenta serán los manchones producto de la suciedad o deterioro de las placas.

La segmentación se realiza utilizando la técnica de etiquetado con vecindad de 8 [10]. Para ello se utilizan dos punteros (puntero de inicio y puntero de parada) que sirven

para etiquetar los objetos. La técnica del puntero es muy útil para caracteres con algún tipo de inclinación y de diferentes tamaños, debido a que en esta situación no es posible segmentar los caracteres usando sólo líneas horizontales y verticales. Es por ello que el empleo de esta técnica en la segmentación de los caracteres de la placa se realiza independientemente de su inclinación, posición, tamaño y distancia entre la cámara y la posición de la placa.

Para explicar el procedimiento de la segmentación se realizará un ejemplo en el cual se describirán los pasos a seguir. Por comodidad la imagen que se muestra en la figura 2.17 no está representada en ceros y unos como si lo están las imágenes reales luego de aplicar la umbralización.

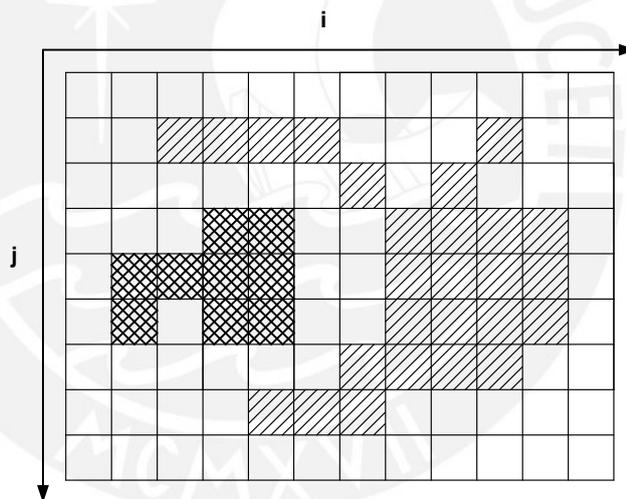


Figura 2.17 Dos patrones diferentes

La imagen a segmentar o extraer es la más pequeña. Se aclara que sólo por comodidad se han representado de la anterior forma las dos imágenes.

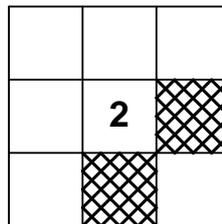
Primer paso. Una vez realizada la umbralización se procede a realizar un barrido total de la imagen con la finalidad de encontrar algún tipo de objeto. El barrido se hace de

arriba hacia abajo y de izquierda a derecha según como se ve la imagen. En la realidad, según como se vea la imagen se seguirán estos ejes de coordenadas.

Segundo paso. Al encontrarse un nuevo objeto, se crea una nueva imagen donde se ubica la imagen segmentada y a continuación se etiqueta el punto de coordenadas donde se encontró la imagen. Por un lado habría que advertir que la etiqueta de la imagen depende del número de objeto segmentado, comenzando con el número dos; y del otro, que se inicializan los punteros (puntero de inicio y puntero de parada) con el número uno y dos respectivamente.

Tercer paso. Se almacena en memoria la coordenada donde se localizó el primer píxel del objeto a segmentar, en el caso del ejemplo sería (2,5). Para el desarrollo del programa se realizó una pequeña variación por comodidad. Así, en lugar de tener un valor de (2,5) se tiene (5,2), donde la primera coordenada corresponde al eje “j” y la segunda al eje “i”.

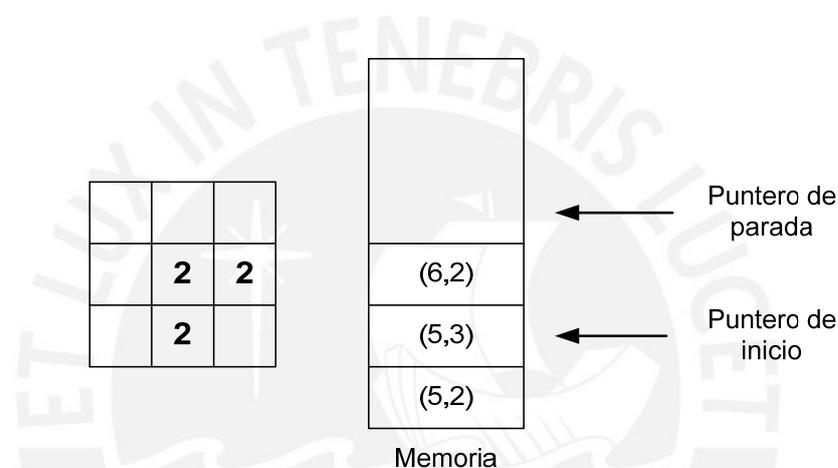
Cuarto paso. A continuación se analiza la vecindad de la coordenada encontrada, ésta es de 3 x 3. Si tomamos dichas dimensiones como ejemplo, obtendríamos el siguiente gráfico:



En el gráfico anterior, el número dos representa la etiqueta que le corresponde al objeto encontrado.

Quinto paso. En el siguiente paso se analiza el contenido de la vecindad comenzando por el extremo izquierdo superior y terminado en el extremo derecho inferior, cuando

se encuentra con el valor de uno. Para el ejemplo el valor son las líneas inclinadas. Enseguida se almacena esta posición de memoria, siendo esta (5,3), se incrementa en uno el puntero de parada y se continúa con el análisis. Como resultado se vuelve a encontrar con información que indica posición, la cual se almacena en la siguiente dirección de memoria (6,2) incrementándose de nuevo el puntero de parada. Cabe señalar que cuando se encuentra información se procede a etiquetarla. Se tiene la gráfica siguiente como resultado al final de esta vecindad:



Se incrementa el puntero de inicio en uno cuando se termina de analizar la vecindad

Sexto paso. El puntero de inicio indicará donde se realizará el siguiente análisis de vecindad.

Séptimo paso. Se repiten el cuarto, quinto y sexto paso sucesivamente hasta que el puntero de inicio sea igual al puntero de parada.

Octavo paso. Se reconstruye la imagen segmentada teniendo como base la memoria ya que ésta nos indica la posición de cada coordenada, obteniendo como resultado toda la forma del objeto segmentado.

Noveno paso. Una vez terminado de segmentar el objeto se continúa con el barrido de la imagen principal, solo que esta vez la nueva etiqueta será incrementada en uno.

Décimo paso. En el caso de encontrarse con más información en la imagen se repiten del segundo al noveno paso hasta terminar con el barrido de la imagen.

Producto de este proceso no es solo la segmentación, también es posible ubicar las coordenadas donde se inicia la imagen o mejor dicho la coordenada desde la que se extrae todo el objeto. Además, también puede extraerse información como el tamaño del objeto, tanto en su altura como en su anchura. Todos estos datos son útiles ya que ayudarán a mejorar la imagen de los objetos encontrados.

Como se observa el procedimiento de segmentación utilizado en el desarrollo de esta tesis no toma en cuenta la forma del objeto, tampoco su orientación ni tamaño, lo cual es bastante útil ya que no delimita la distancia entre la cámara y la placa. Otra ventaja de esta técnica es que se pueden eliminar pequeñas manchas o mejor dicho no tomarlas en cuenta dada la facilidad para determinar el área del objeto segmentado con solo observar el valor del puntero de parada.

2.7 Reconocimiento de los caracteres

Para el proceso de reconocimiento de caracteres se emplean tres redes neuronales, una para los números y dos para las letras. Las tres redes son del tipo red backpropagation [8] y los patrones de entrada utilizados para entrenarla pertenecen a las peculiaridades de los caracteres.

2.7.1 Creación de patrones

De cada carácter se extraen dos patrones. Cada patrón está constituido por la sumatoria de cada fila y de cada columna, según el esquema matemático siguiente:

$$f(j) = \sum_{i=1}^n ima(i, j)$$

$$g(j) = \sum_{j=1}^m ima(i, j)$$

$$ima = \underbrace{\begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}}_m \Bigg\}^n$$

Como las placas no son capturadas a una misma distancia, dado que la distancia entre la cámara y el vehículo no son siempre las mismas, se obtienen diferentes tamaños para cada carácter. Por eso se necesita un proceso de normalización, tanto en el dominio como en el rango de las funciones “f” y “g”.

2.7.1.1 Normalización

El rango establecido para todos los patrones es de 0 a 1, lo cual se logra dividiendo cada valor de la función entre el valor más alto de la misma.

$$f_n(j) = \frac{f(j)}{\max(f(j))}$$

Para que todas las funciones tengan el mismo dominio se establece un valor fijo de elementos por cada función y para lograr este paso se emplea un sub-muestreo y/o sobre-muestreo a cada una (ver figura 2.18). Se emplea un sobre-muestreo cuando la cantidad de elementos que se tiene es menor a la deseada y un sub-muestreo cuando es mayor. Para el sobre-muestreo se agregan valores al dominio. Dichos valores están representados por el valor intermedio existente entre diferentes muestras adyacentes,

lo cual da como resultado el doble de elementos menos el primero que se tenía originalmente. Si al realizarse este proceso la cantidad de elementos sigue siendo menor al valor fijado, se vuelve a realizar el sobre-muestreo hasta tener un valor mayor o igual al deseado. Para el sub-muestreo se toma la diferencia existente entre el valor deseado y el valor de la función actual, siendo esta última la cantidad de elementos a no tomar en cuenta dada su separación equidistante. A continuación se tiene la figura 2.18 en la cual se muestran las diferentes funciones realizadas.

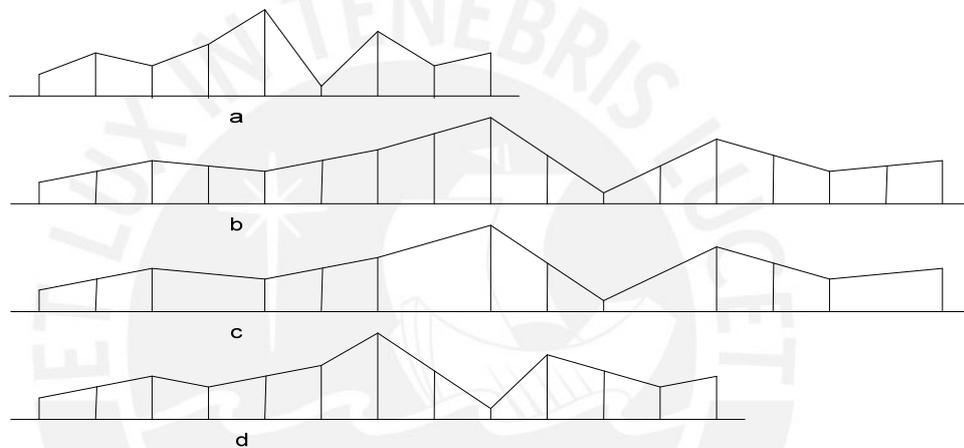


Figura 2.18 Se muestra: a) La función real, b) Función sobre-muestreada
c) Sub-muestreo de la función y d) La función normalizada

Las funciones normalizadas no son suficientes para entrenar a la red neuronal, por lo que es necesario introducir una etapa previa dedicada al procesamiento de los patrones.

2.7.2 Procesamiento de los patrones

Se observa que los caracteres pueden ser dibujados en una cuadrilla de 7x4 píxeles, donde se muestra que cada uno de los patrones está dividido en tres secciones. Para el patrón de las columnas se tienen dos secciones iguales y una de doble tamaño; y

para el patrón de las filas dos secciones iguales y otra de un tamaño cinco veces mayor (ver figura 2.19).

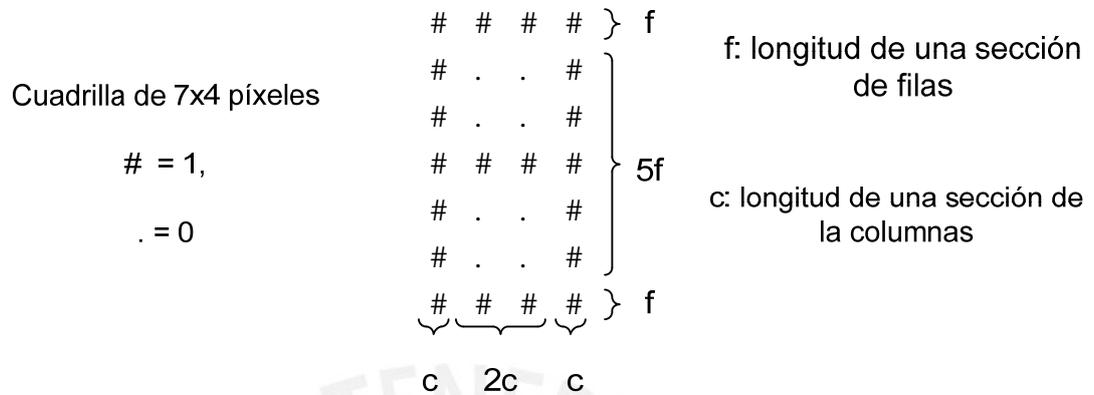


Figura 2.19 Procesamiento de los patrones

Se modifican los patrones a fin de que se tengan las proporciones anteriormente dichas, esto con el fin de reducir el error e igualar los patrones utilizados para diferentes caracteres del mismo símbolo (letra o número). Las modificaciones se realizan utilizando sobre-muestreo y/o sub-muestreo hasta obtener las dimensiones deseadas.

2.8 Redes neuronales

En la actualidad las Redes Neuronales Artificiales son las que están causando un mayor impacto en la ciencia computacional, debido a su extraordinaria aplicabilidad práctica. Recientemente esta tecnología ha captado la atención de los profesionales dedicados a la estadística y al análisis de datos [12], quienes comienzan a incorporar las redes neuronales al conjunto de herramientas estadísticas orientadas a la clasificación de patrones y la estimación de variables continuas.

En consonancia con la incorporación de esta nueva herramienta en diferentes campos científicos, en esta tesis se emplea la red multicapa (Backpropagation) de una sola

capa oculta, dada su probada eficacia. La red multicapa es una de las redes neuronales más utilizadas en la actualidad debido a los buenos resultados derivados de su uso. Esta red permite obtener un margen de éxito en el proceso de clasificación de patrones y la estimación de variables continuas [8], [12], [13].

Las dimensiones de la arquitectura de las redes neuronales empleadas son: 110-200-26 para las letras y 110-200-10 para los números. Cabe señalar que el número de entradas está determinado por los patrones, cada uno de ellos comprende 110 elementos: 70 y 40 para las funciones f y g , respectivamente. El número de neuronas de salidas es determinado por el sistema, el cual reparte 10 elementos para los números (del 0 al 9) y 26 para las letras. Para determinar el número de neuronas de la capa oculta no existe una regla específica, se determina según la necesidad y la experiencia del diseñador de la red. Únicamente cabría añadir que el número no debe ser muy alto puesto que traería como problema el sobre-entrenamiento de la red, lo cual no es ideal; ni muy bajo, puesto que el error sería muy alto.

Para la realización de esta tesis se emplearon 200 neuronas en la capa oculta. Se llegó a esta cifra después de realizar varias pruebas con diferentes cantidades de neuronas, siendo este el valor que arrojó menor error. Tanto para la capa de entrada como para la oculta se empleó un bias⁵ con el fin de mejorar el desempeño de la red neuronal.

⁵ El termino bias actúa como los pesos en las redes neuronales cuya salida a las neuronas es siempre uno

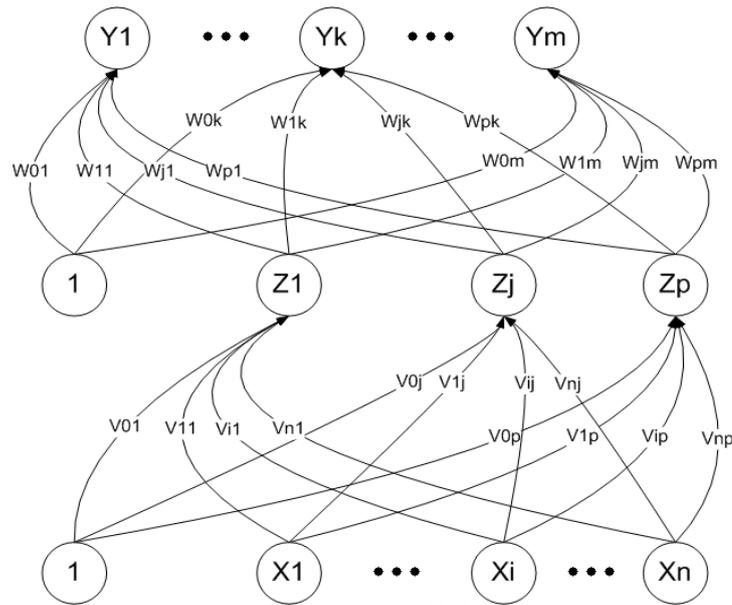


Figura 2.20 Arquitectura de la red neuronal multicapa con una sola capa

Las funciones de activación empleadas son: lineal para la entrada, y sigmoide bipolar para la capa oculta y para la capa de salida. Se utilizó sigmoide bipolar en la capa de salida ya que las variables de salidas asignadas fueron de 1 y -1 y lo que se quiere lograr es que se saturen las salidas que no corresponden a la entrada asignada.

Función sigmoide:

$$\begin{array}{c}
 \text{Diagram: } \sum \xrightarrow{m} z_j \xrightarrow{n} \\
 n = \frac{2}{1 + e^{-\left(\frac{m-c}{a}\right)}} - 1
 \end{array}$$

$$\begin{array}{c}
 \text{Diagram: } \sum \xrightarrow{\theta} y_k \xrightarrow{y} \\
 Y = \frac{2}{1 + e^{-\left(\frac{\theta-b}{q}\right)}} - 1
 \end{array}$$

Donde:

- Xi: neurona de entrada i.
- Zj: neurona j de la capa oculta
- Yk: neurona de salida k.

w_{jk} : peso de conexión entre la neurona de la capa oculta j y la neurona de salida k .

v_{ij} : peso de conexión entre la neurona de la capa de entrada i y la neurona de la capa oculta j .

Ecuaciones para la red:

$$w_{jk} = w_{jk} - \eta(Y_k - \bar{Y}_k) \frac{\partial Y_k}{\partial \theta_k} n_j$$

$$v_{ij} = v_{ij} - \eta \sum_k (Y_k - \bar{Y}_k) \frac{\partial Y_k}{\partial \theta_k} w_{jk} \frac{\partial n_j}{\partial m_j} X_i$$

$$b_k = b_k - \eta(Y_k - \bar{Y}_k) \frac{\partial Y_k}{\partial b_k}$$

$$q_k = q_k - \eta(Y_k - \bar{Y}_k) \frac{\partial Y_k}{\partial q_k}$$

$$c_j = c_j - \eta \sum_k (Y_k - \bar{Y}_k) \frac{\partial Y_k}{\partial \theta_k} w_{jk} \frac{\partial n_j}{\partial c_j}$$

$$a_j = a_j - \eta \sum_k (Y_k - \bar{Y}_k) \frac{\partial Y_k}{\partial \theta_k} w_{jk} \frac{\partial n_j}{\partial a_j}$$

Además se tiene:

$$\frac{\partial Y}{\partial \theta} = -2 \left(1 + e^{-\left(\frac{\theta-b}{q}\right)} \right)^{-2} e^{-\left(\frac{\theta-b}{q}\right)} \left(-\frac{1}{q} \right)$$

$$e^{-\left(\frac{\theta-b}{q}\right)} = \frac{1-Y}{1+Y}$$

Reemplazando se obtiene:

$$\frac{\partial Y}{\partial \theta} = -2 \left(\frac{1+Y}{2} \right)^2 \left(\frac{1-Y}{1+Y} \right) \left(-\frac{1}{q} \right) \Rightarrow \frac{\partial Y_k}{\partial \theta_k} = \frac{(1+Y_k)(1-Y_k)}{2q_k}$$

$$\frac{\partial n}{\partial m} = -2 \left(1 + e^{-\left(\frac{m-c}{a}\right)} \right)^{-2} e^{-\left(\frac{m-c}{a}\right)} \left(-\frac{1}{a} \right) \Rightarrow \frac{\partial n_j}{\partial m_j} = \frac{(1+n_j)(1-n_j)}{2a_j}$$

$$\frac{\partial Y}{\partial b} = -2 \left(1 + e^{-\left(\frac{\theta-b}{q}\right)} \right)^{-2} e^{-\left(\frac{\theta-b}{q}\right)} \left(-\frac{1}{q} \right) \Rightarrow \frac{\partial Y_k}{\partial b_k} = -\frac{(1+Y_k)(1-Y_k)}{2q_k}$$

$$\frac{\partial Y}{\partial q} = -2 \left(1 + e^{-\left(\frac{\theta-b}{q}\right)} \right)^{-2} e^{-\left(\frac{\theta-b}{q}\right)} [-(\theta-b)](-1)q^{-2} \Rightarrow \frac{\partial Y_k}{\partial q_k} = -\frac{(1+Y_k)(1-Y_k)}{2q_k^2} (\theta_k - b_k)$$

$$\frac{\partial n}{\partial c} = -2 \left(1 + e^{-\left(\frac{m-c}{a}\right)} \right)^{-2} e^{-\left(\frac{m-c}{a}\right)} \left(-\frac{1}{a} \right) \Rightarrow \frac{\partial n_j}{\partial m_j} = -\frac{(1+n_j)(1-n_j)}{2a_j}$$

$$\frac{\partial n}{\partial a} = -2 \left(1 + e^{-\left(\frac{m-c}{a}\right)} \right)^{-2} e^{-\left(\frac{m-c}{a}\right)} [-(m-c)](-1)a^{-2} \Rightarrow \frac{\partial n_j}{\partial a_j} = -\frac{(1+n_j)(1-n_j)}{2a_j^2} (m_j - c_j)$$

Se utilizaron dos redes neuronales para las letras debido a su mayor cantidad de valores de salida. Cada red fue adiestrada mediante diferentes patrones de entrenamiento. El funcionamiento por separado de cada red no alcanza un éxito considerable siendo este menor al 50%. La unión de estas redes se realiza comparando el valor de cada neurona de salida y tomando el valor mayor como se muestra en la figura 2.21, ya que esto significa que para ciertos patrones de entrada la red tiene una mejor respuesta.

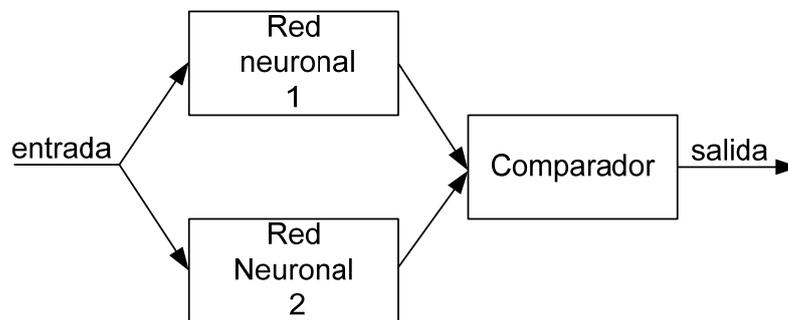


Figura 2.21 Diagrama de bloques del reconocimiento de letras

2.9 Problemas en el desarrollo del sistema

Las imágenes capturadas y empleadas en el desarrollo de esta tesis, fueron tomadas bajo la luz natural en diferentes días y ante diferentes condiciones climatológicas (se tomaron durante días nublados y soleados). En ese sentido, de acuerdo a la iluminación natural de cada día en que se capturaron las imágenes se presentaron una serie de problemas. Por ello se hizo necesario proponer un sistema que fuera lo suficientemente robusto como para que pudiera funcionar durante las diferentes etapas del sistema: el proceso de ubicación de la placa y el del reconocimiento de los caracteres.

En relación a lo anterior se realizará una descripción de los problemas presentados así como sus soluciones planteadas. El orden propuesto se realiza en concordancia con el modo en el que fueron apareciendo dichos problemas durante el proceso de localización y reconocimiento de los caracteres de la placa de un vehículo.

2.9.1 Localización de la placa

A continuación se señalan los errores producidos durante el proceso de localización de las placas vehiculares, así como las mejoras realizadas para contrarrestar sus efectos al interior del sistema propuesto. Los errores se presentaron, principalmente, en las etapas de estimación vertical y en la segmentación. Las mejoras realizadas son las producidas en la parte inferior y en el tamaño de los caracteres, y las planteadas en los cortes extremos de la placa y frente a los caracteres que rodean al guión.

2.9.1.1 Error en la etapa de estimación vertical

En la etapa de estimación vertical se utilizó la función laplaciano bidimensional con el fin de detectar los bordes de la placa frente a los de la pista, dado que este rasgo no es homogéneo y uniforme en las tonalidades de intensidad. El problema radica en que las tonalidades se intensifican en diferentes grados durante los días soleados originando que durante el empleo del laplaciano se encuentren bordes en la pista, lo cual provoca la confusión entre ésta y la placa de rodaje. Lo anterior dado que, en algunos casos, la proporción del ancho de esta región es parecida a la proporción de la placa.

Otro problema encontrado se debe a que algunos vehículos son más altos que otros, lo cual provoca que al fondo de la imagen de un vehículo se vea la pista, originando la detección de falsos bordes debajo de la placa. Para evitar esta confusión entre la pista y la placa se analiza el valor mayor de la región significativa. Si ésta es aproximada al ancho de la imagen se está haciendo referencia a la pista, por tanto se ignora la información y se sigue con la búsqueda de la placa; de no ser así, se ha logrado la localización exitosa de la placa.

2.9.1.2 Error en la segmentación

En la segmentación se presentan varios tipos de problemas. Los principales son:

1. Producto del deterioro de la placa.
2. Sombra producto del mismo vehículo en días soleados.
3. La inclinación de la placa.
4. Adhesivos pegados en la placa.

5. El marco negro que se coloca como protección a la placa originando la unión de los caracteres en la parte inferior.
6. Los pernos que se utilizan para fijar la placa en el vehículo.
7. La suciedad en la placa, lo cual produce manchas en la imagen.

En el caso del deterioro de la placa podemos encontrar errores como:

- Falta del guión, carácter que separa las letras y los números (ver figura 2.21). La ausencia del guión trae como principal problema la dificultad para saber si se trata de una placa de tres o dos letras, lo cual origina que un carácter numérico sea procesado por la red neuronal entrenada para el reconocimiento de letras en lugar del destinado al reconocimiento de los números.

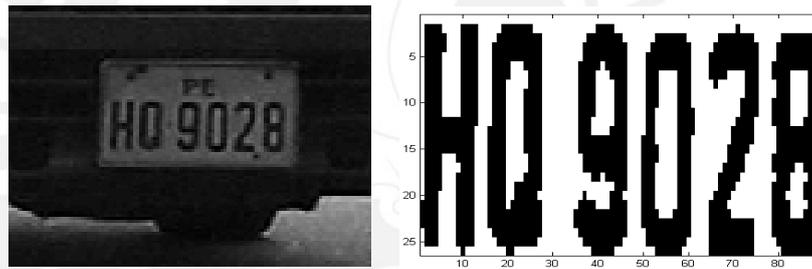


Figura 2.21 Error producido por falta de guión

- Falta de pintura en los caracteres. Este defecto en la placa produce una falta de información en los caracteres durante la umbralización (ver figura 2.22). En el peor de los casos puede suceder que un carácter sea analizado como dos caracteres, esto produciría error no solo en los caracteres analizados sino en los otros caracteres a los que desplaza en una posición.

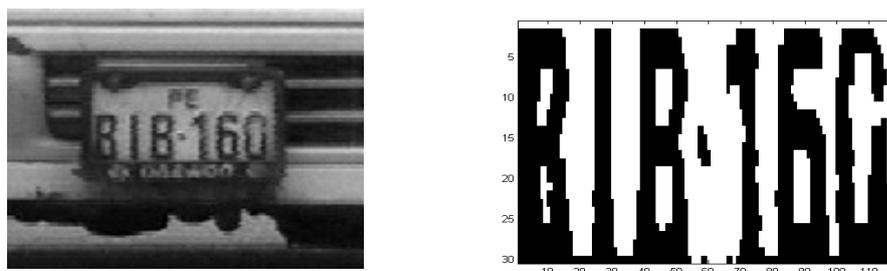


Figura 2.22 Error producido por falta de pintura

Las sombras producidas por el vehículo, al igual que el borde negro que se encuentra en algunas placas de rodaje, trae como error la unión de uno o más caracteres. El defecto se debería a la forma de segmentación empleada, puesto que esta unión permite considerar dos caracteres como uno solo. Se tienen varios casos de unión, estos se agrupan según el número de objetos encontrados que puedan ser considerados como caracteres. Como resultado se tienen ocho casos, en los que cada caso corresponde al número de objetos encontrados.

Algunas placas que no cuentan con un marco poseen en su lugar pernos lo suficientemente grandes como para unir los caracteres o para considerarlos como objetos válidos en el momento de hacer la segmentación. Por otro lado, la inclinación de la placa trae consigo pérdida de información en la parte inferior del último o primer carácter (según sea la inclinación), y falta de información en la parte superior del carácter del otro extremo; además, produce la aparición de un nuevo objeto ubicado en la parte que está entre la placa inclinada y la posición vertical inferior de la estimación vertical.

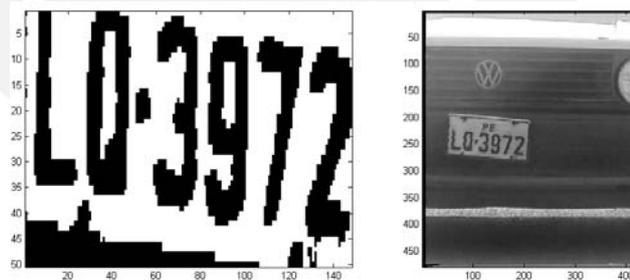


Figura 2.23 Placa inclinada

A continuación se mostrarán las soluciones planteadas según los casos de error presentados. Como el sistema no sabe si es afectado por algún error en particular, se debe considerar el peor de los casos en el ensayo, es decir, hacer como si la placa estuviera afectada por todos los errores anteriormente presentados.

El primer paso es no tomar en cuenta los objetos que son producidos por la suciedad de la placa, valiéndonos de la geometría de los caracteres. Dichos caracteres ocupan dentro de la placa una determinada área, por lo que en base al tamaño estimado podría calcularse el área que podría corresponder a cada objeto. Por ello es necesario obtener el área de cada objeto encontrado durante el proceso de segmentación. Únicamente de este modo podría establecerse el área de estos objetos, así ellos tengan un área menor no correspondiente a los caracteres analizados, en cuyo caso no se tomarían en cuenta. La pregunta que salta ante esta discriminación de objetos es: ¿qué pasa con el guión que no tiene la misma área que los demás objetos? Para solucionar este problema se establece como área de comparación el área estimada que podría tener el guión dentro de la placa analizada.

En caso de encontrarse con pernos que posean un área lo suficientemente grande como para ser tomada en cuenta, es decir, que su área se confunda con el área de un guión, se tomará en consideración las características de la placa. Los pernos siempre se localizan entre el segundo objeto y el penúltimo objeto, es decir, la ubicación dos y la ubicación siete; mientras que el guión se localiza en la posición tres o cuatro según se tenga dos o tres letras, respectivamente. De este modo, al encontrarse con un perno éste no será tomado en cuenta gracias a su ubicación dentro de la placa.

A continuación se elimina el borde inferior, producto de la inclinación de las placas y del borde negro. Para saber si es necesaria esta etapa se toma en cuenta el tamaño de cada objeto, si uno de ellos es casi igual al ancho de la sección de la placa localizada se procede a reemplazar los valores de 1 por el de 0 de ese objeto, evitando afectar al resto de la imagen. Después se procede a contar los objetos encontrados en la segmentación. En caso de que el conteo sea diferente a siete (el guión, dos letras y cuatro números o tres letras y tres números), se tendría que

analizar si el conteo es mayor o menor que siete. En el caso de ser menor se señalaría que existe unión en los objetos o falta del guión (seis objetos). Por el contrario, de ser mayor implicaría que la estimación no fue tan precisa en la obtención de los extremos, ya sea el lado derecho o izquierdo o en ambas partes del vehículo. Al final, durante la umbralización se obtienen nuevos objetos (Ver figura 2.24).

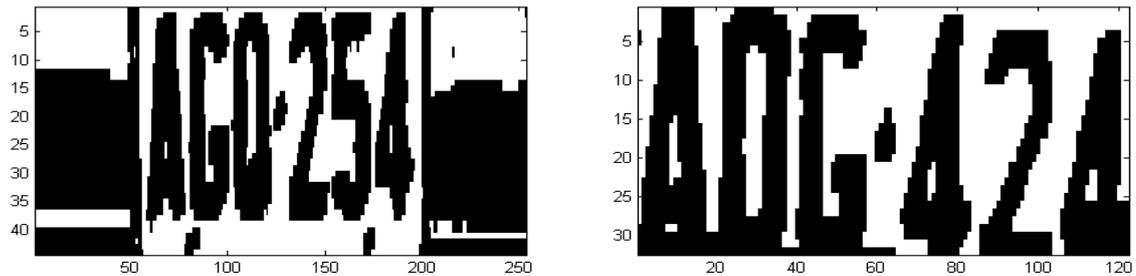


Figura 2.24 Muestra de casos en que se tienen más y menos de siete caracteres

En caso de que los objetos encontrados sean menores a siete existen tres formas de separar a los caracteres:

- Eliminando la franja horizontal inferior que une a estos elementos
- Eliminando la franja vertical que une a estos elementos.
- Eliminación de una línea que une a los caracteres.

Primero se busca el objeto más ancho, el cual estaría compuesto por dos o más caracteres unidos. Si la anchura de este objeto comparado con el resto de objetos es mayor a la mitad de estos, se pasa a analizar la posibilidad de eliminar la parte inferior o una franja del objeto; si no se cumple este criterio, nos encontramos ante el caso de una placa de vehículo que no cuenta con el carácter del guión. Para saber donde representar el guión se toma en cuenta las diferencias posiciones entre el segundo y tercer objeto, y entre el tercer y cuarto objeto. En este proceso la mayor diferencia representa la falta de un objeto lo que, según el caso del guión, permite saber si la placa es de tres o dos letras.

Regresamos al problema de si se debe eliminar la franja inferior o una franja vertical de la imagen. Para decidir que acción tomar se introduce un nuevo bloque, lo cual es necesario debido a la problemática que presenta el carácter F (ver figura 2.25). En la gráfica se tiene la unión del carácter F con otro (letra O) por la parte inferior. En ese caso, de tomar como separación las franjas verticales, se podría leer en lugar de F el carácter E y reconocer como tal el carácter E, lo cual sería una decisión errada para el operador del sistema y para el propio sistema.

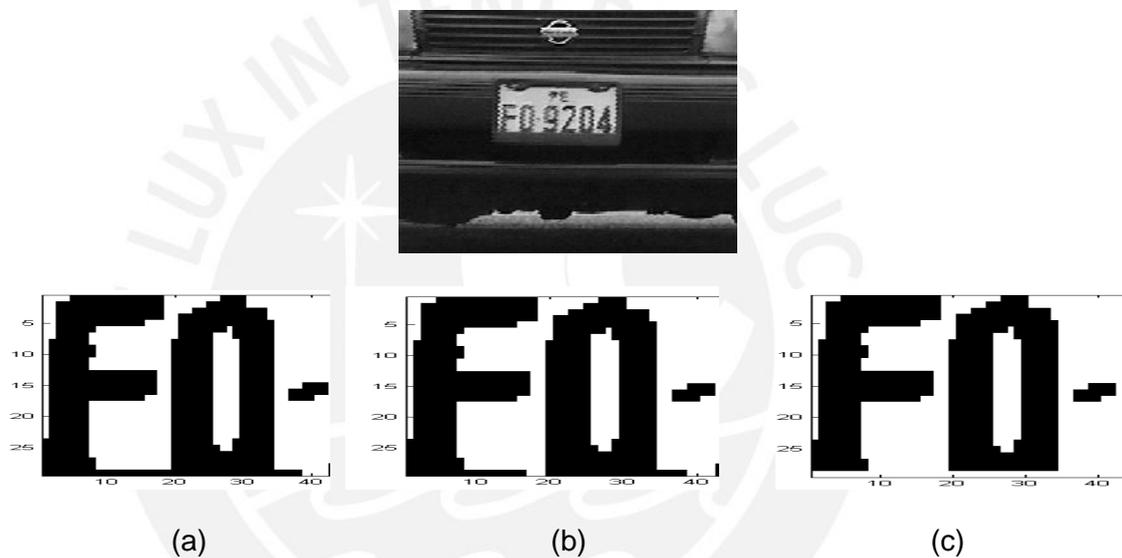


Figura 2.25 Unión del carácter F

En los gráficos de la figura anterior se muestran:

- (a) La unión entre los caracteres producto del margen negro.
- (b) La utilización de franjas verticales en la separación de objetos.
- (c) La eliminación de la franja horizontal inferior.

En el caso de que sea mayor a siete caracteres, se procede a analizar los objetos que se encuentran en los extremos de la placa estimada. Como paso final los objetos son eliminados si no corresponden al tamaño de los demás objetos. Tanto para objetos mayores o menores a siete el análisis es realizado hasta que cumplan con las

características impuestas para una placa de siete caracteres, ya que una placa puede presentar varios tipos de errores a la hora de segmentarla.

Cuando se tienen siete objetos se analiza el primer objeto o el último, ya que puede darse el caso de que a la hora de estimar la ubicación de la placa se haya producido el corte del primer carácter (ver figura 2.26). Para evitar el corte del primer carácter se analiza un tramo más a la izquierda y a la derecha de la placa, si se encuentra continuidad en el objeto se procede a ajustar el corte de la placa.

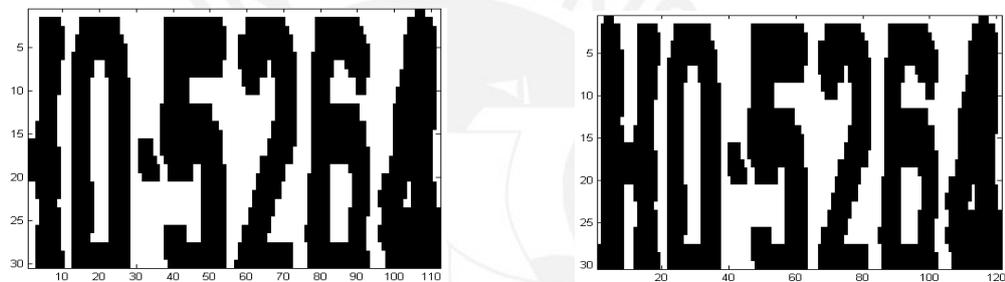


Figura 2.26 Error en el corte y su respectivo ajuste

Finalmente se procede a realizar retoques a los caracteres con la finalidad de que al momento de extraer sus patrones estos mantengan cierta homogeneidad y obtengan un mayor grado de acierto en el proceso de reconocimiento.

2.9.1.2.1 Mejoramiento de la parte inferior de los caracteres

Generalmente, durante la etapa de umbralización de la imagen de las placas de rodaje, sobre todo en las placas que cuentan con el margen negro, algunos caracteres se unen en la parte inferior por lo ya expuesto anteriormente. En algunos casos no se llega a tener la unión por unos cuantos píxeles (ver figura 2.27), lo cual origina que la parte inferior de los caracteres presente cierta deformidad y provoque que los patrones extraídos del carácter no constituyan una buena representación de los mismos.

La solución propuesta es reemplazar la última fila por la antepenúltima, siempre y cuando la última fila del carácter cuente con una mayor cantidad de píxeles con respecto a la antepenúltima fila.

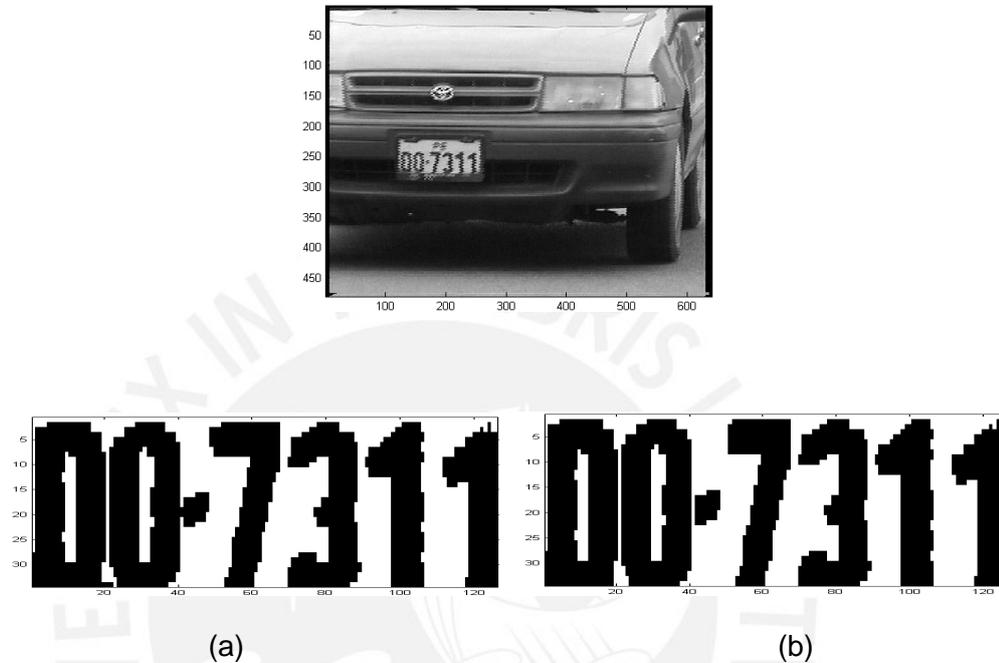


Figura 2.27 (a) La parte inferior del primer carácter muestra una deformidad producida por la umbralización; b) Se muestra el mejoramiento de la parte inferior de la placa.

2.9.1.2.2 Mejoramiento del tamaño de los caracteres

Un problema característico de las placas es el del deterioro de sus esquinas, ya sea por la falta de pintura o por las manchas de suciedad. Dicho desperfecto sólo es importante cuando se produce en las esquinas inferiores de la placa, dado que puede darse el caso que afecte a los caracteres que se encuentran en los extremos (ver figura 2.28). En casos como éstos, durante la umbralización de la placa se manifiesta una deformidad en la imagen procesada (justo en el lugar en el que se localiza el deterioro), lo que ocasiona que el carácter contenga información innecesaria.



Figura 2.28 Imagen de una placa con deterioro en sus esquinas. Se muestra como la umbralización de la placa ocasiona una deformidad del carácter B y 7.

Frente al problema señalado, la solución propuesta es el recorte del tamaño de los caracteres extremos (B y 7) con la finalidad de obtener una altura homogénea en todos los caracteres de la placa. Para lograr la homogeneidad de los caracteres, primero se busca el carácter más alto de todos los presentes en la placa y se pasa a eliminar la información que se encuentra por debajo de la altura asignada. Durante este procedimiento se evita tocar el resto de la imagen, trabajando únicamente con el carácter que corresponde al problema en cuestión (ver figura 2.29).

Para la determinación de la altura se emplea el resto de los caracteres. Como medida eje se tiene el tamaño del carácter que posea la misma o más cercana dimensión a la coordenada-y con respecto al carácter examinado. Lo anterior se debe tomar en cuenta para evitar problemas con placas inclinadas, principalmente.



Figura 2.29 Recorte en el tamaño de los caracteres

2.9.1.2.3 Mejoramiento en los cortes extremos de la placa

Las placas que poseen un margen negro también presentan otros problemas. Uno de ellos es la sombra que se produce en los extremos de la placa, debido a diferentes razones como el que la placa no se encuentre unida completamente al vehículo, la mala colocación de la misma o por el deterioro de la placa. Lo anterior origina que los caracteres de los extremos se encuentren unidos con el margen, provocando deformidad en estos caracteres y produciendo que el sistema arroje un carácter errado durante el proceso de reconocimiento.



Figura 2.30 Error en el corte de los extremos

Como se observa en la figura 2.30 los caracteres de los extremos cuentan con partes que no deben ser analizadas, ya que esto provocaría un mal reconocimiento del carácter en cuestión. La solución propuesta es tomar el tamaño del carácter más ancho entre el segundo y sexto objeto, y comparar este valor con el ancho de los caracteres de los extremos. Para ello la diferencia de valores es el indicativo de si nos encontramos o no ante un carácter que posee parte innecesaria en sus extremos izquierdo o derecho, así como también en el primer y último carácter. De ser imperiosa la eliminación de parte del carácter, se analiza sólo una parte del extremo derecho o izquierdo del carácter, según sea el caso. Enseguida formamos un patrón con la suma

de las columnas de la región a analizar, en donde la menor indica el lugar en el que se deberá realizar el nuevo corte a la placa.



Figura 2.31 Mejora en el corte de los extremos

2.9.1.2.4 Mejoramiento en los caracteres que rodean al guión

Al realizarse la umbralización algunos caracteres se encuentran unidos, veamos el caso de la unión entre un carácter y el guión. Como se explicó anteriormente, se tiene una etapa de separación para la cual se describieron tres diferentes técnicas. En el caso de la unión de un carácter con el guión, se emplea la separación utilizando la eliminación de una línea vertical. Se obtienen buenos resultados bajo el punto de vista de una simple separación, pero el inconveniente que se presenta es la deformidad del carácter que estaba unido. Ello dado que puede dejarse parte del guión en el carácter.

A continuación se presenta la figura 2.32 donde se muestra la unión del carácter G y el guión de una placa y el proceso de separación respectivo. Como resultado, se observa que el carácter G se queda con parte del guión de la placa.



Figura 2.32 Muestra de unión entre los caracteres G y – (guión),

La solución frente al problema consiste en analizar la vecindad del guión. Para ello se toma una distancia hacia la derecha y la misma para la izquierda y se analiza cada vecindad por separado tomando en cuenta la suma de elementos de cada columna. Si la cantidad de elementos encontrados es menor o igual al ancho del guión se reemplaza la columna vertical.

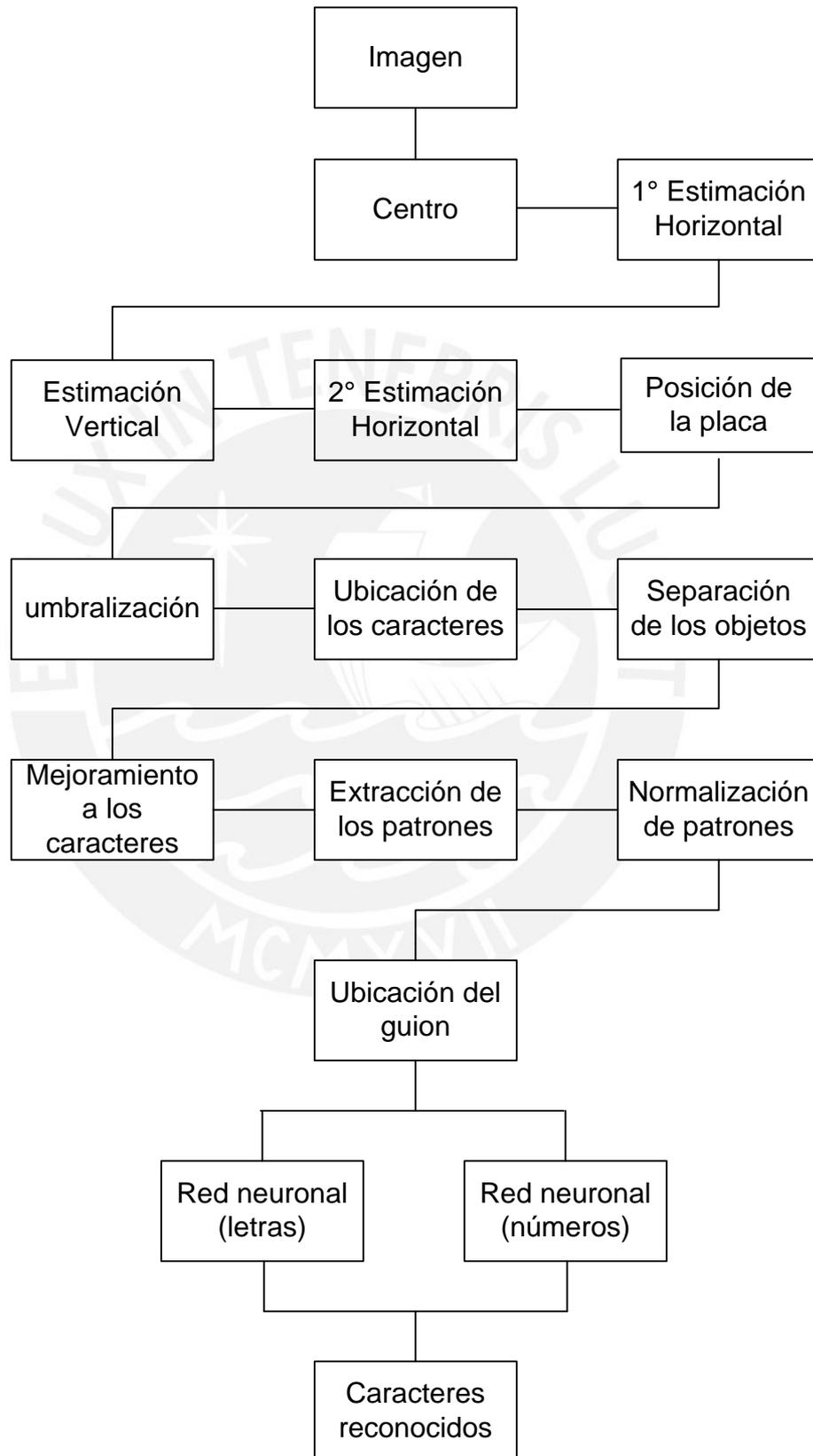


Figura 2.33 Mejoramiento del carácter que rodea al guión

2.9.2 Error en el reconocimiento

Los errores que se producen en la etapa de reconocimiento se deben, mayormente, a que durante la etapa de entrenamiento de las redes neuronales se trabaja con un margen de error. Otros errores también se producen por causa de los diferentes problemas que se presentan en la etapa de segmentación, debiéndose el error, en muchos casos, a las etapas previas y no a la red. Frente a ello no se puede hacer mucho dentro de esta etapa para mejorar el acierto en el reconocimiento.

2.10 Diagrama de bloques del sistema propuesto



CAPÍTULO 3

ANÁLISIS DE RESULTADOS

Para realizar las pruebas de la presente tesis se tomaron 200 imágenes de vehículos mediante una cámara de video analógica instalada al interior de la Pontificia Universidad Católica del Perú. La ubicación de la cámara se estableció en la entrada de la avenida Riva Agüero (ver figura 3.1) con un horario para tomas de entre 10am a 12am y 2pm a 4pm. La grabación de todas las imágenes en video se realizaron bajo iluminación natural en días nublados y soleados. Una vez obtenidas las grabaciones requeridas se pasó a realizar la digitalización de las imágenes vehiculares y, posteriormente al análisis cuyos resultados serán expuestos en este capítulo.

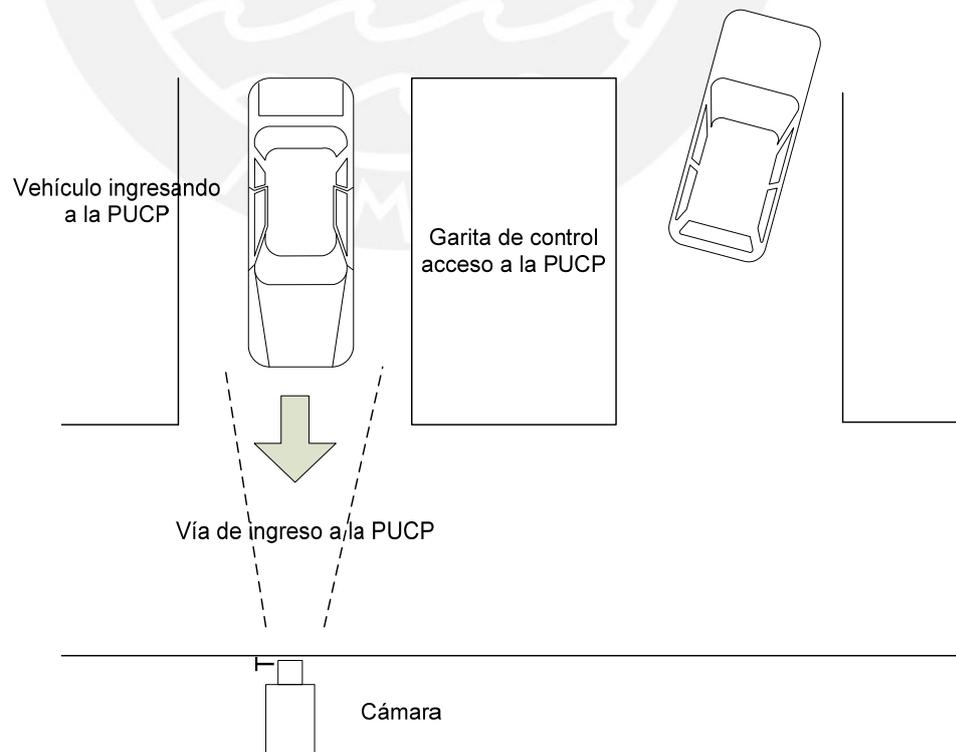


Figura 3.1 Esquema del sistema

Como se ha detallado a lo largo de esta investigación, el sistema propuesto se divide en dos procesos: proceso de localización de la placa dentro de la imagen capturada y el proceso de reconocimiento de los caracteres. Según la descripción ya realizada en el segundo capítulo, el proceso de reconocimiento tiene como pieza fundamental la localización del centro en la imagen, ya que esta permite determinar si se puede o no llegar a reconocer la placa vehicular.

En el primer proceso se obtuvo un éxito de 80.4%, es decir, se logró ubicar satisfactoriamente un alto porcentaje del total de placas incluidas en la muestra de la investigación. Este resultado se debe a que la localización del centro permite un éxito de 88.9% y, a su vez, condiciona los resultados positivos posteriores. De allí la importancia de ubicarlo como el primer paso básico en el sistema propuesto, dado que permite un mejor nivel de efectividad en la aplicación del primer proceso.

En el segundo proceso sólo se tomó en cuenta las placas que se pudieron localizar satisfactoriamente, es decir, el 80.4% del total de placas analizadas. En la figura 3.2 se muestra el porcentaje gráfico de aciertos tanto para los números como para las letras.

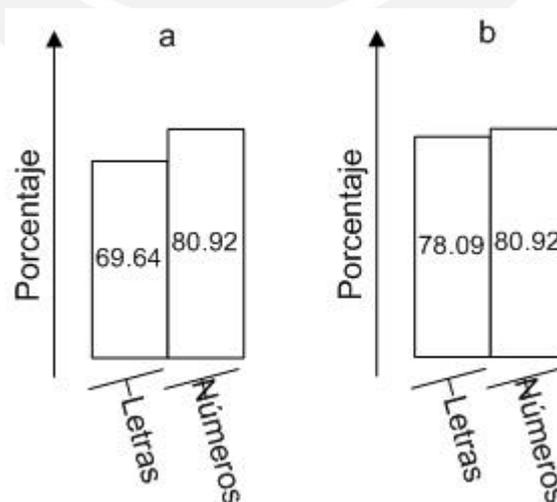


Figura 3.2 (a) Porcentaje de aciertos b) Porcentaje tomando en cuenta el guión

En la figura 3.2 la diferencia entre los gráficos radica en que el gráfico b se toma en cuenta el guión como parte del sistema neuronal dedicado al reconocimiento de letras, es decir, se considera como carácter alfabético al guión. Con respecto a ello se tiene que el reconocimiento del guión tiene un éxito del 100%.

A continuación se muestra el análisis de los resultados de éxito en cada carácter. Se comienza por el análisis de los números y después se realiza el de las letras.

3.1 Análisis de los números

Se realiza el análisis de cada número por separado para obtener el porcentaje de éxito en la identificación. Ello con la finalidad de encontrar qué número es el menos eficiente o qué números son los que presentan mayores dificultades de identificación durante el proceso de reconocimiento del sistema propuesto.

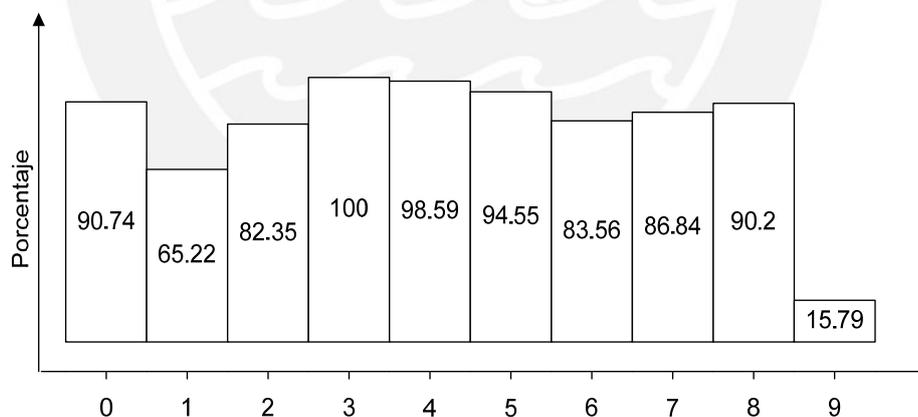


Figura 3.3 Porcentaje de éxito en la identificación

La figura 3.3 muestra los niveles de éxito en la identificación por cada carácter pero no es de mucha ayuda ya que no permite visualizar mayores detalles sobre el real funcionamiento del sistema. Por ejemplo, para el carácter tres la gráfica muestra un 100% de aciertos, es decir, que el sistema siempre reconoce todos los 3 que se

presentan en la muestra de análisis. Sin embargo, en la práctica esto no implica que siempre que se tiene 3 como valor de respuesta, esta corresponda necesariamente a dicho número impar. Por tanto, se puede deducir que el margen de error o acierto en el reconocimiento de un número no es exacto. Ante ello se hace necesario establecer una gráfica porcentual que permita establecer una mejor descripción del sistema y encontrar, a su vez, una mayor certeza sobre la eficiencia de la red neuronal. Con ese fin se muestra la siguiente gráfica (figura 3.4).

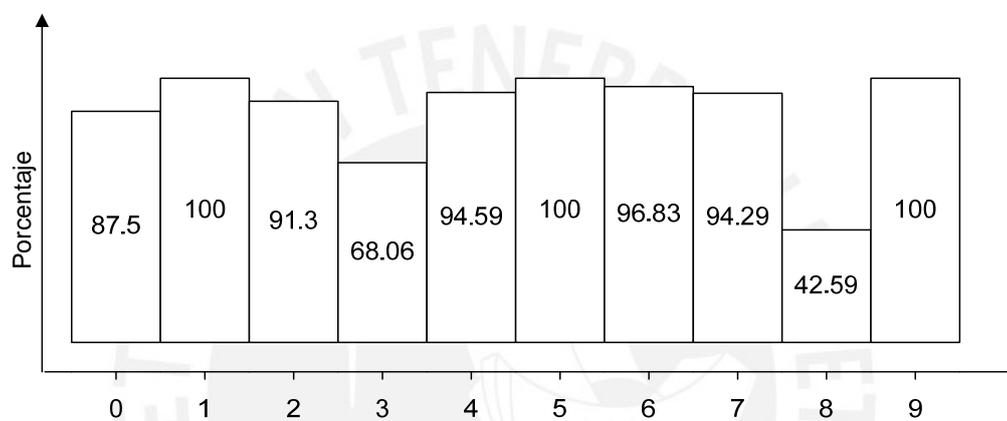


Figura 3.4 Eficiencia de la red

Como se ve en la figura 3.4, los valores para el carácter numérico tres han cambiado. Los porcentajes señalan que el reconocimiento de dicho carácter tiene una eficiencia de 68.06%, lo que equivale a decir que el sistema arrojó como resultado el carácter 3 más veces de las que debía, lo cual implica un nivel de error en la identificación del carácter. Otra diferencia que se observa es que el carácter 1 en la figura 3.3 muestra un éxito de 65.22% y, en cambio una eficiencia de 100% en la figura 3.4. Esto quiere decir que no se reconocen todos los unos introducidos en el sistema, pero que sí la salida del sistema corresponde al carácter uno puede tenerse la completa seguridad de que el número ha sido correctamente identificado. Como se observa, la figura 3.4 nos ayuda a reconocer el nivel de eficiencia de la red neuronal empleada durante el

reconocimiento de números, lo cual permite establecer los lineamientos para mejorar el sistema.

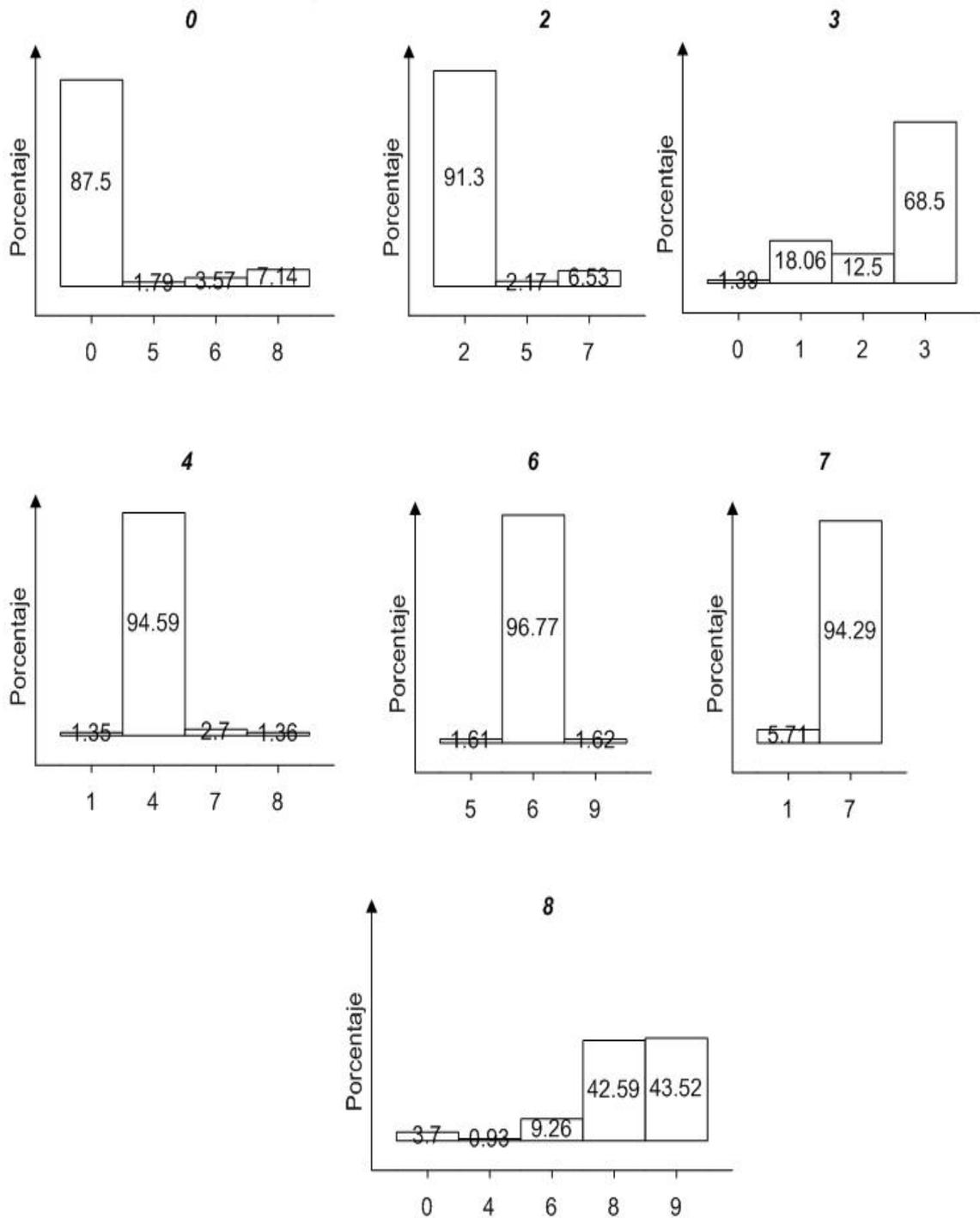


Figura 3.5 Distribución del acierto a la salida de la red

Los gráficos de la figura 3.5, anteriormente presentadas, representan la eficiencia en el reconocimiento durante el empleo de redes neuronales, es decir, el nivel de

funcionamiento y activación de una neurona correspondiente a un patrón de entrada específico. Para una mejor y más detallada descripción de las gráficas se analizará la correspondiente al número tres.

Como se observa, la gráfica número tres tiene diferentes valores de porcentaje y para cada valor un carácter correspondiente. Cada carácter corresponde al patrón de entrada y el nivel porcentual muestra la eficiencia o equívoco de la red. En otras palabras, los porcentajes muestran que la red se activó una cierta cantidad de veces de las cuales sólo un porcentaje correspondía al patrón correcto. Así, las cifras porcentuales mostradas en la gráfica son:

1.39% al patrón de entrada 0

18.06% al patrón de entrada 1

12.5% al patrón de entrada 2

68.5% al patrón de entrada 3

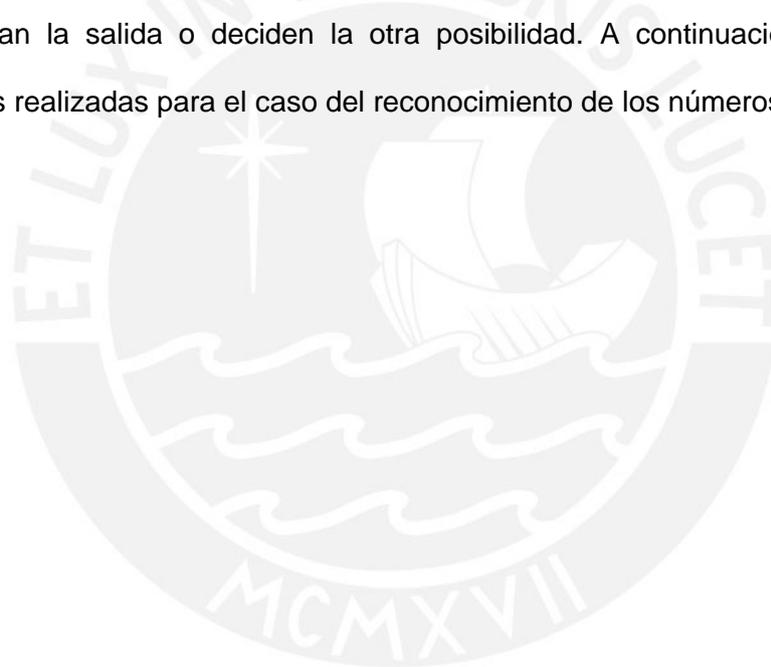
Según lo anterior, el sistema tiende a confundir solamente los patrones de entrada 1, 2 y asignarles como salida 3. Los casos más críticos se muestran a continuación:

Salida de la red	Patrón de entrada
0	8
2	7
3	1 y 2
8	9 y 6

A la luz de estos resultados se hace necesario agregar un sistema de confirmación de salida, es decir, un sistema que permita saber de manera efectiva si la salida asignada

por la red neuronal corresponde al patrón de entrada o si se trata de otro carácter. Las gráficas mostradas anteriormente dan la solución a cuales podrían ser los otros caracteres logrando una más fácil tarea de confirmación de salida, ya que en algunos casos sólo se trata de dos posibilidades. Un ejemplo de lo anterior es el caso de la salida correspondiente al carácter 2, en ese caso lo que se debe analizar es si se trata efectivamente del dos o del siete.

Para mejorar el sistema se introducen bloques a las salidas previamente seleccionadas. Estos bloques permiten analizar el patrón de entrada y según estos confirman la salida o deciden la otra posibilidad. A continuación se muestran las mejoras realizadas para el caso del reconocimiento de los números.



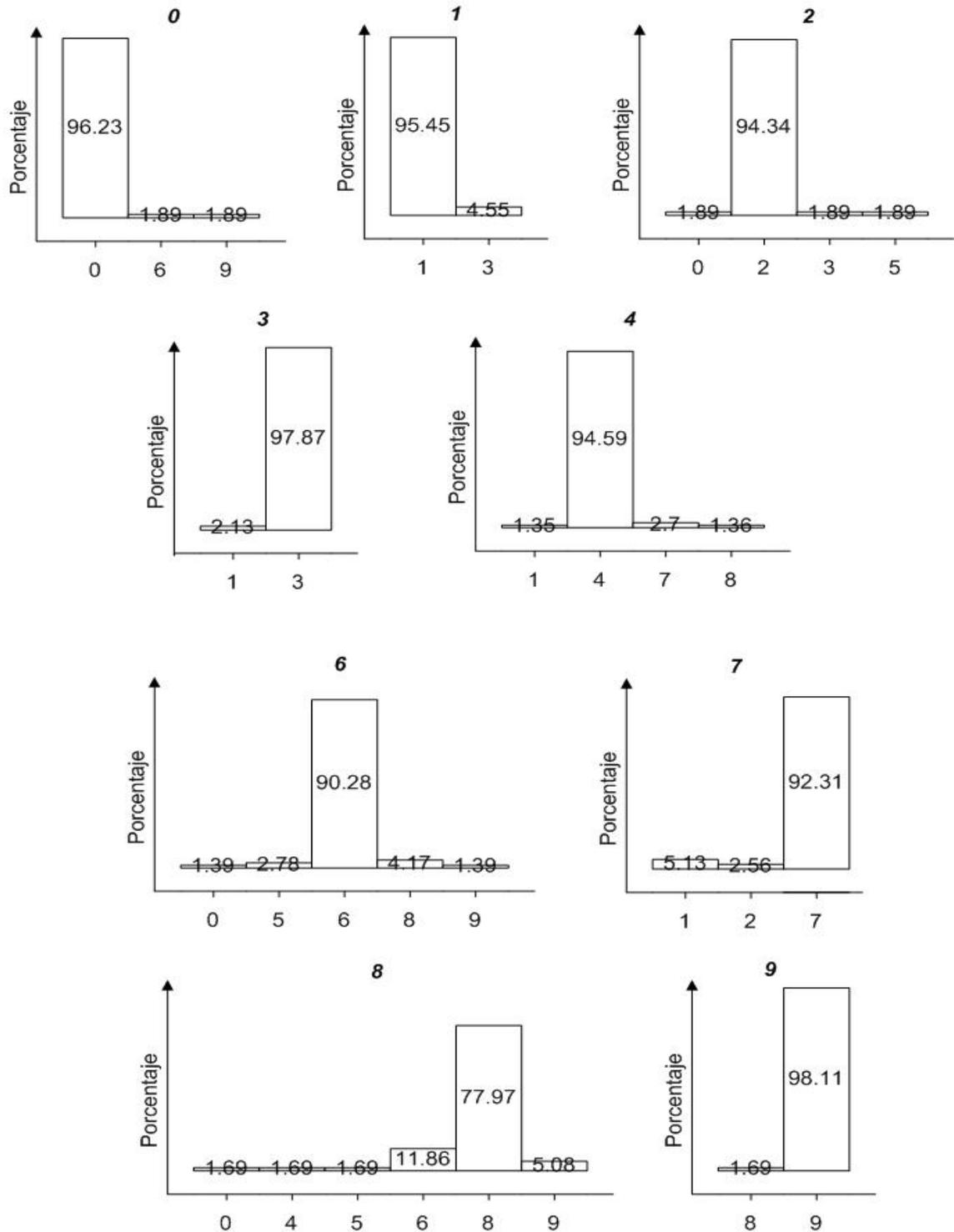


Figura 3.6 Porcentaje de eficiencia a la salida del sistema utilizando bloques de confirmación

El carácter correspondiente al número ocho es el más difícil de determinar ya que su patrón muestra mayor concordancia con otros caracteres. A pesar de ese problema se logró mejorar la eficiencia del sistema.

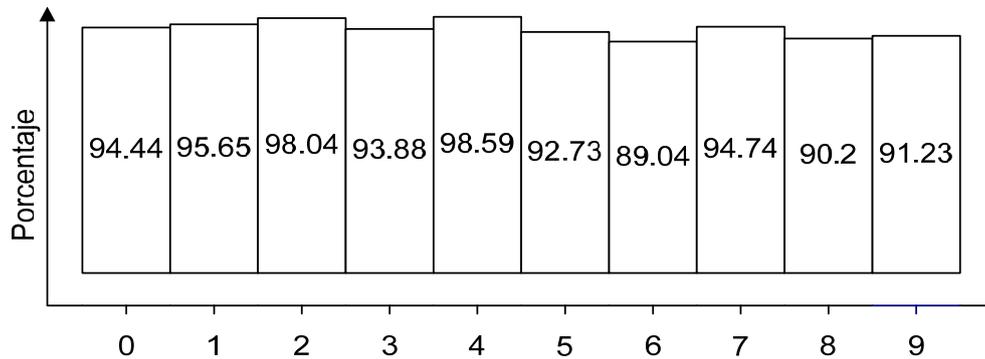


Figura 3.7 Acierto a la salida del sistema

La figura 3.7 muestra los porcentajes de aciertos a la salida del sistema. Como se ve, estos valores porcentuales son mucho más óptimos comparados con el que se tenía inicialmente en la figura 3.3, lo cual significa una mejora en la eficiencia del proceso de reconocimiento de números.

3.2 Análisis de las letras

El análisis de los caracteres correspondientes a las letras se hace de forma parecida a la de los números. En primera instancia se analiza el éxito de identificación en los caracteres (ver figura 3.8) y la eficiencia del sistema (ver figura 3.9), sobre la base de estos análisis se agregan bloques de confirmación de salida. La diferencia que se presenta con respecto a los números es que se tiene que tomar en cuenta no sólo el porcentaje de eficiencia, sino también el número de muestras que se tiene. Habría que recordar que se cuenta con 26 caracteres diferentes correspondientes a cada letra y algunos caracteres no son tan comunes o no se tienen demasiadas muestras para hacer un análisis exacto.

Se cuentan con 200 imágenes capturadas de forma aleatoria, para las cuales no se realizó ningún tipo de discriminación, siempre y cuando cumplieren con las consideraciones previamente establecidas. Como resultado de la toma circunstancial de las placas incluidas en el análisis se obtuvo una mayor cantidad de muestras para algunos caracteres y menor para otros.



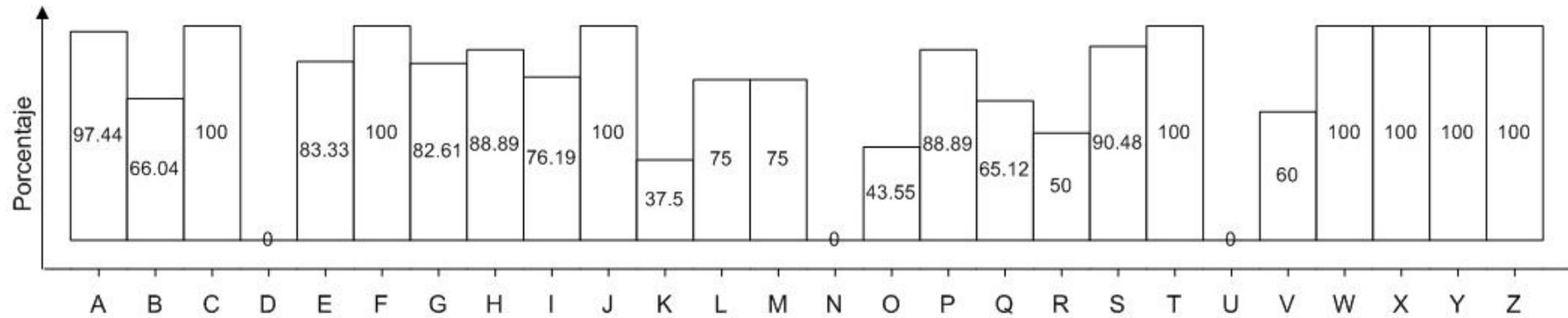


Figura 3.8 Porcentaje de éxito en la de identificación a la salida de la red

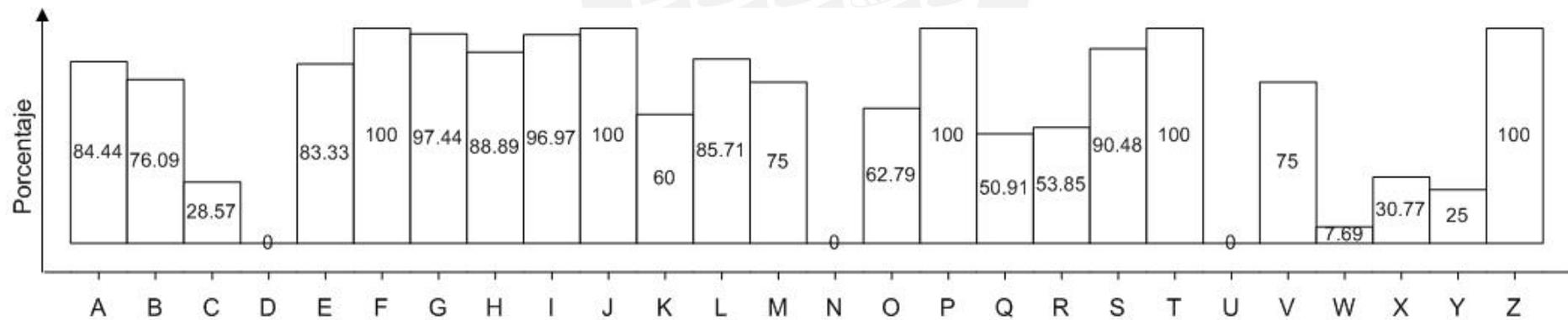


Figura 3.9 Porcentaje de eficiencia a la salida de la red

Según las gráficas, las neuronas de salida correspondientes a los caracteres N, U no se activaron. Esto se debe a la poca cantidad de muestras que se tienen con ese carácter no a la falla en el reconocimiento. También se encuentra un nivel de activación nulo en el caso del porcentaje correspondiente al carácter D. Lo anterior se debe al parecido que dicha letra mantiene con la O, Q, B, y también a la poca cantidad de muestras. A continuación se señala el cuadro con el número de muestras trabajadas en referencia a cada carácter alfabético.

Letra	cant	Letra	cant	Letra	cant
A	39	J	3	S	21
B	53	K	8	T	5
C	4	L	8	U	2
D	10	M	4	V	5
E	6	N	2	W	2
F	7	O	62	X	4
G	46	P	9	Y	1
H	9	Q	43	Z	6
I	42	R	14		

Tabla 3.1 Cantidad de muestras

El carácter W es el menos eficiente por lo que es el primero en analizarse. Con el fin de dicho análisis se tiene la siguiente distribución:

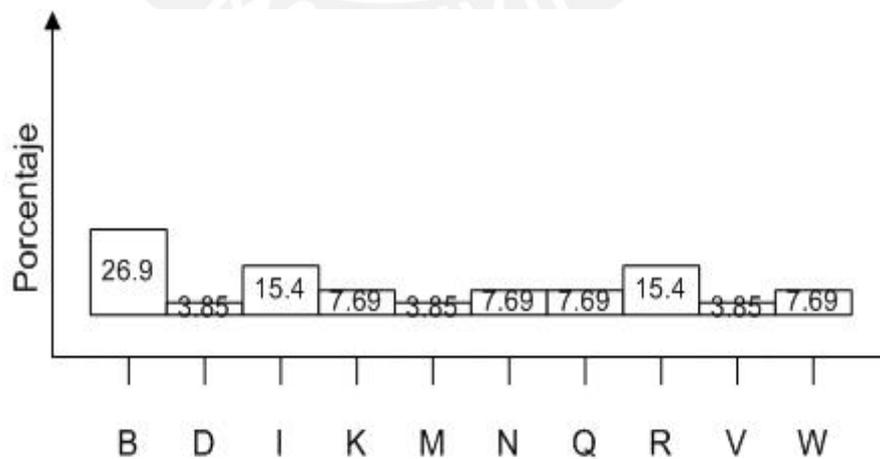


Figura 3.10 Distribución de eficiencia del caracter W

Como se observa en la figura 3.10 se hace complicada la integración de un solo bloque de confirmación. Esto se debe a que cada bloque cuenta con varios caracteres representativos como B, I, K, N, Q, R y W; y además, a la poca cantidad de muestras (dos). Por ese motivo, en la presente tesis el reconocimiento del carácter W no será tomado en cuenta. Una vez hecha esta modificación se logró tener una mejor distribución inicial (ver figura 3.11), así como un mayor nivel de reconocimiento (ver figura 3.12).



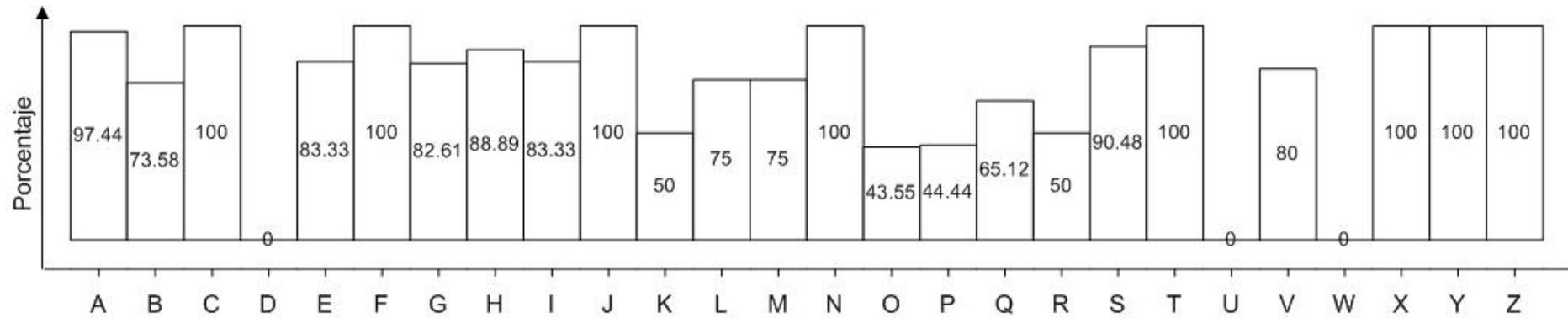


Figura 3.11 Porcentaje a la salida de la red

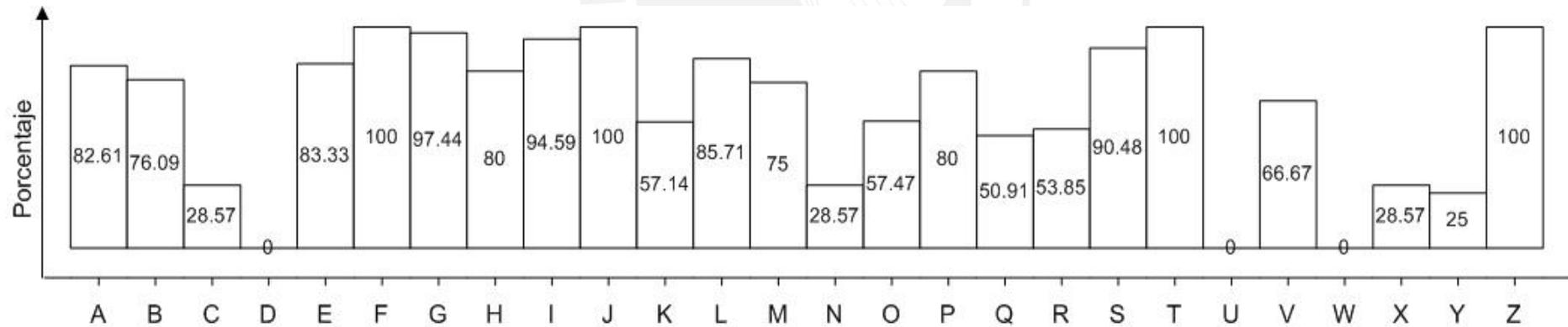


Figura 3.12 Porcentaje de eficiencia a la salida de la red

Las distribuciones porcentuales de eficiencia en el reconocimiento de los caracteres alfabéticos son:

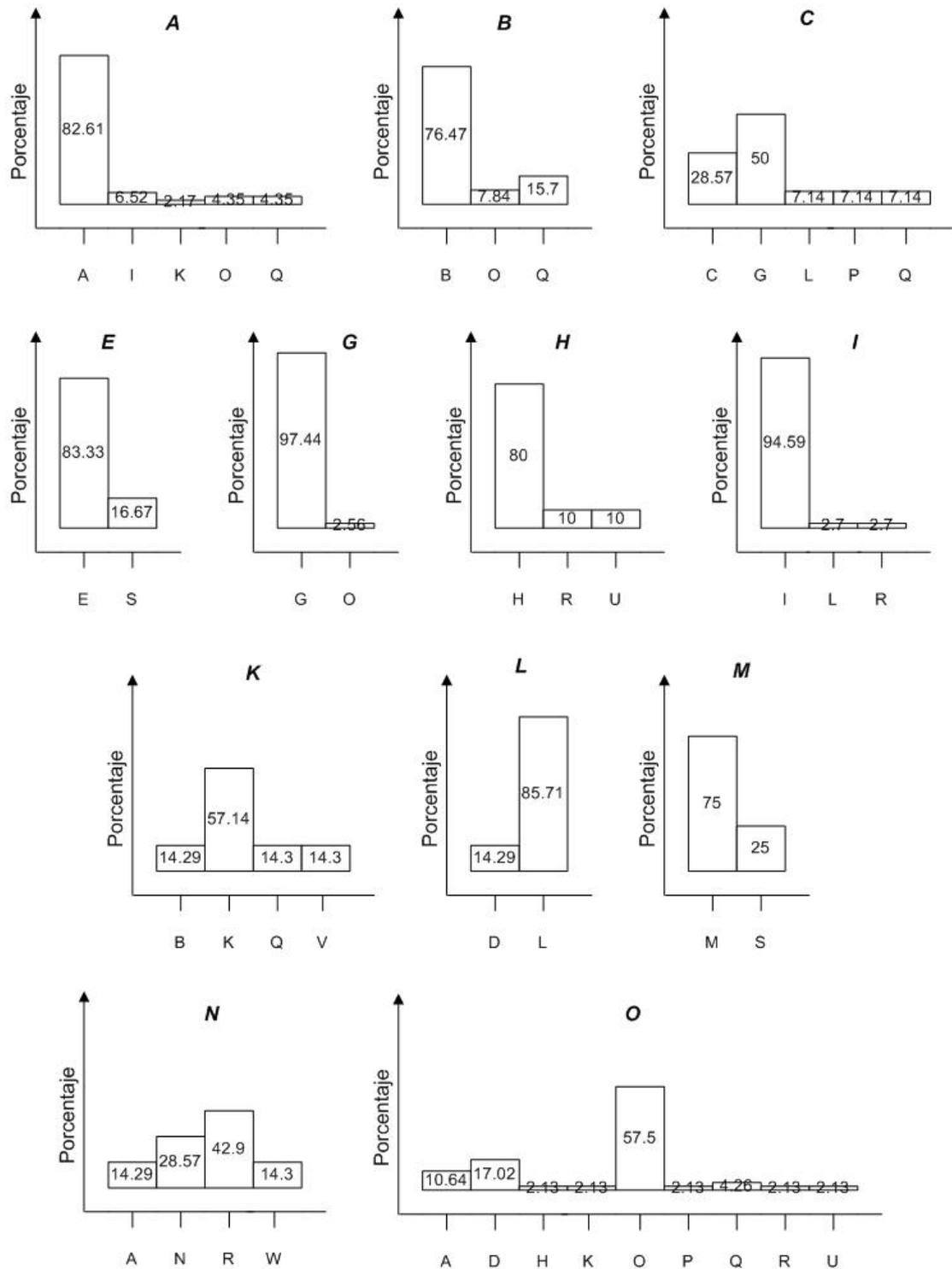


Figura 3.13 Distribución de eficiencia de la A a la O

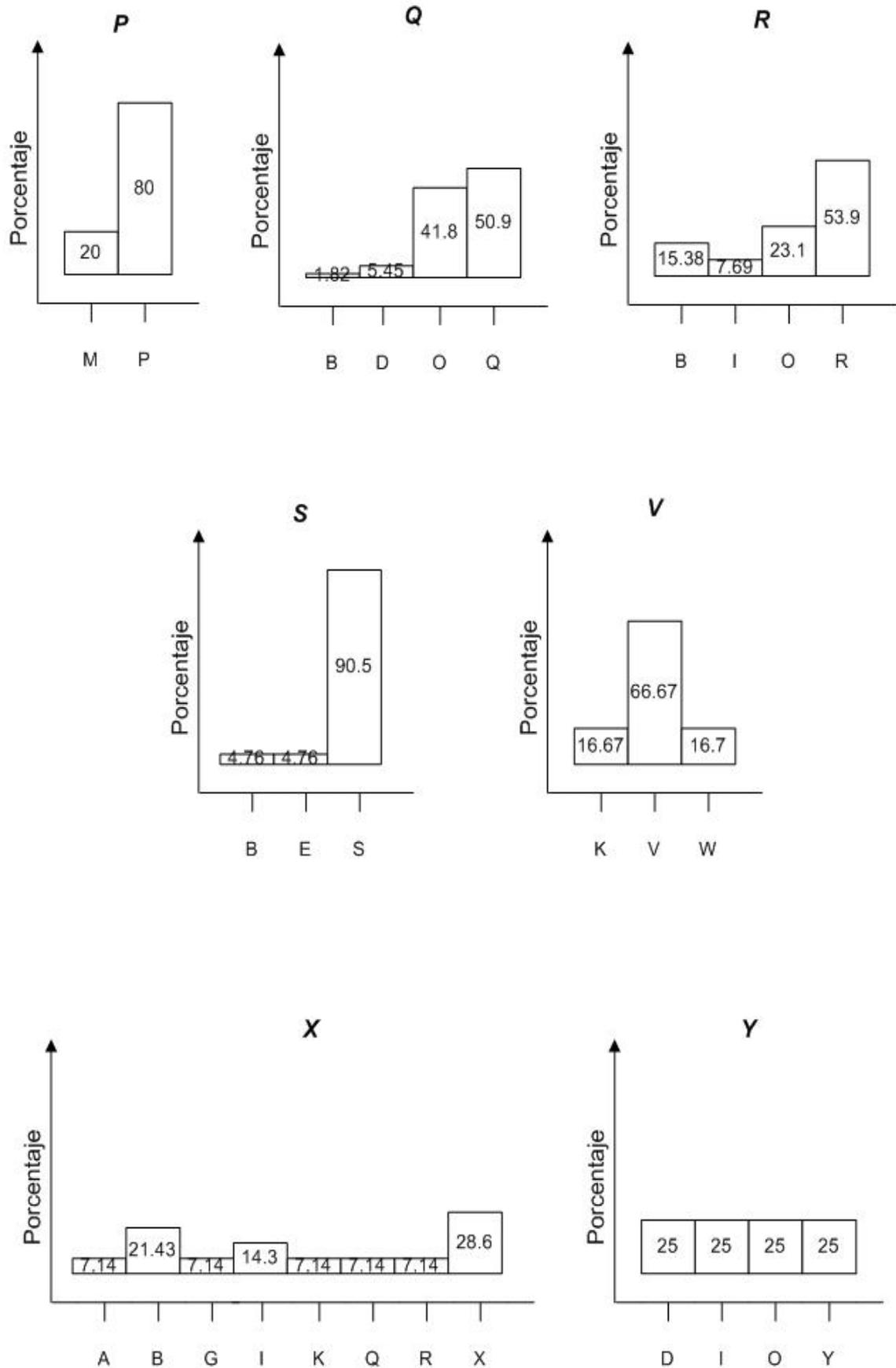


Figura 3.14 Distribución de eficiencia de la P a la Y

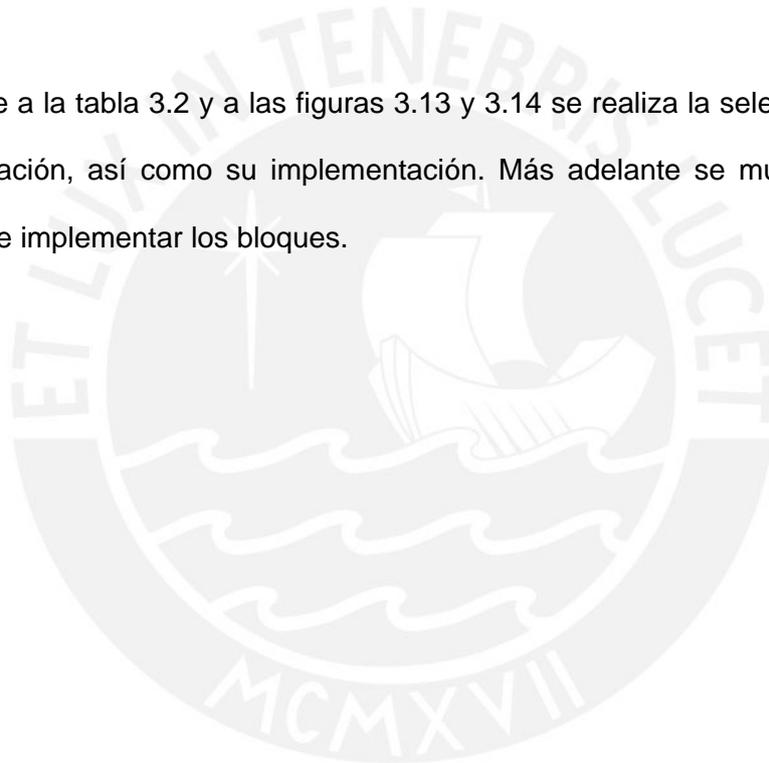
La gráfica correspondiente al carácter X presenta el mismo inconveniente que el carácter W cuando se tiene un porcentaje significativo para los caracteres A, B, G, I, K, Q, R y X. Sin embargo, para este caso no se adopta la misma solución ya que la letra X presenta una mayor cantidad de muestras (cuatro) para el análisis. La solución que se planteó fue introducir un bloque que tuviera como única misión determinar si se trataba del carácter X o no durante el proceso de identificación de los caracteres. Con ese fin, el nuevo bloque se activa cuando la red neuronal arroja como salida dicho carácter. En caso de que el bloque determine que se trata de otro carácter que no es la X, se toma como salida a la neurona que tiene el siguiente mayor peso.

Para determinar que el reconocimiento de las letras se debe mejorar, se toma en cuenta el número de apariciones que tienen cada una, es decir, el número de veces en que las salidas de la red neuronal son activadas. Esto ayudará a realizar un mejor criterio en la selección de bloques de confirmación, ya que se podría tener el caso de que una salida se active cuatro veces y que una de ellas no corresponda a la letra correcta. En ese sentido se tendría una eficiencia de sólo 75% y 25% para el otro carácter. Dada esta situación, si nos fiamos únicamente de los porcentajes obtenidos estaríamos creando un bloque para que corrija un solo error lo cual carece de sentido. Sobre todo en la medida que el error producido podría deberse no a un mal funcionamiento de la red neuronal, si no a que el carácter presenta en sí mismo anomalías diversas que harían difícil la identificación, por ejemplo la existencia de manchas o el deterioro de las placas que dificultan la tarea de reconocimiento.

Letra	cant	Letra	cant	Letra	cant
A	46	J	3	S	21
B	51	K	7	T	5
C	14	L	7	U	0
D	0	M	4	V	6
E	6	N	7	W	0
F	7	O	47	X	14
G	39	P	5	Y	4
H	10	Q	55	Z	6
I	37	R	13		

Tabla 3.2 Aparición en la salida de la red

En base a la tabla 3.2 y a las figuras 3.13 y 3.14 se realiza la selección de bloques de confirmación, así como su implementación. Más adelante se muestran sus mejoras luego de implementar los bloques.



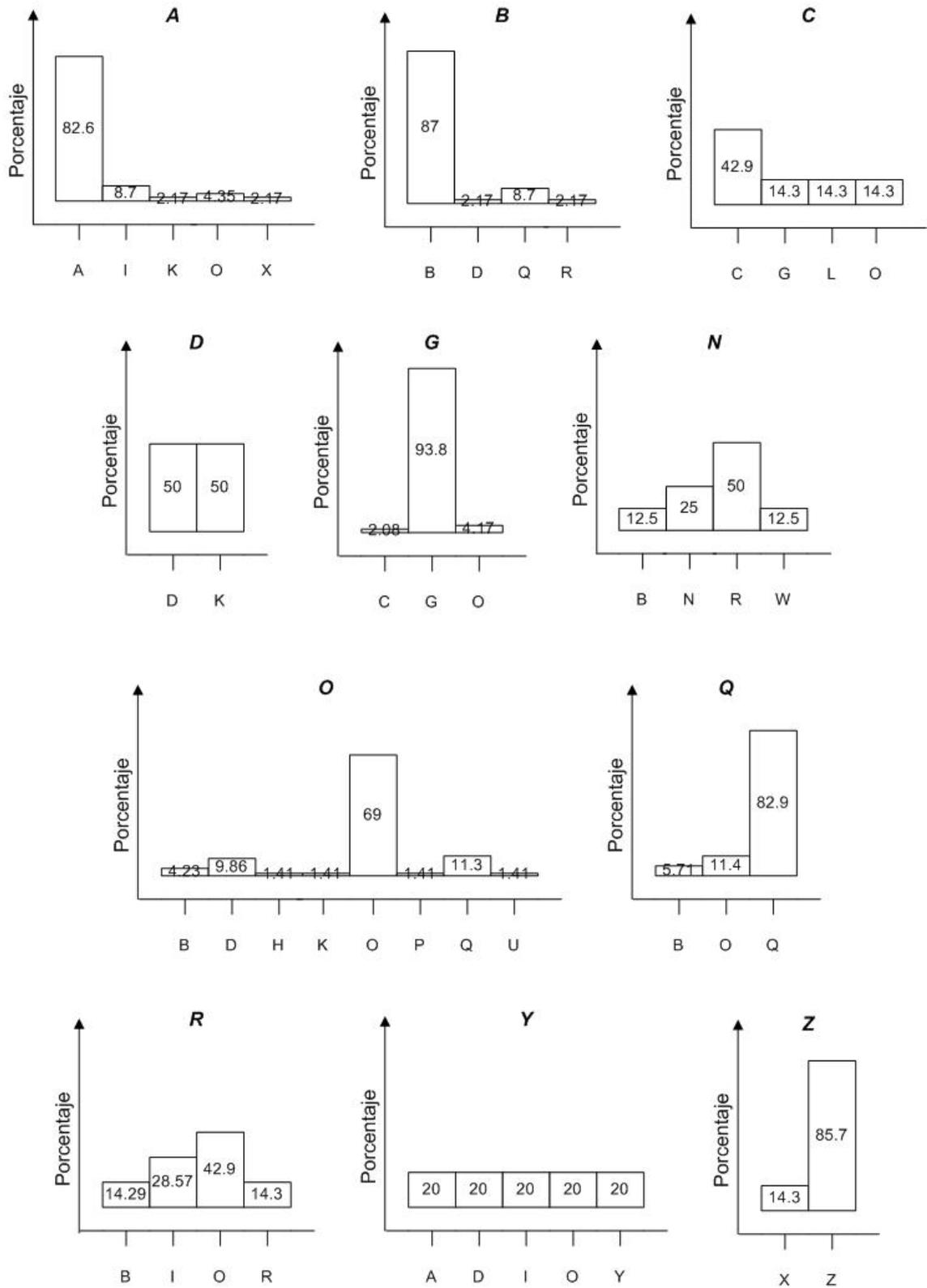


Figura 3.15 Porcentaje de eficiencia a la salida del sistema utilizando bloques de confirmación

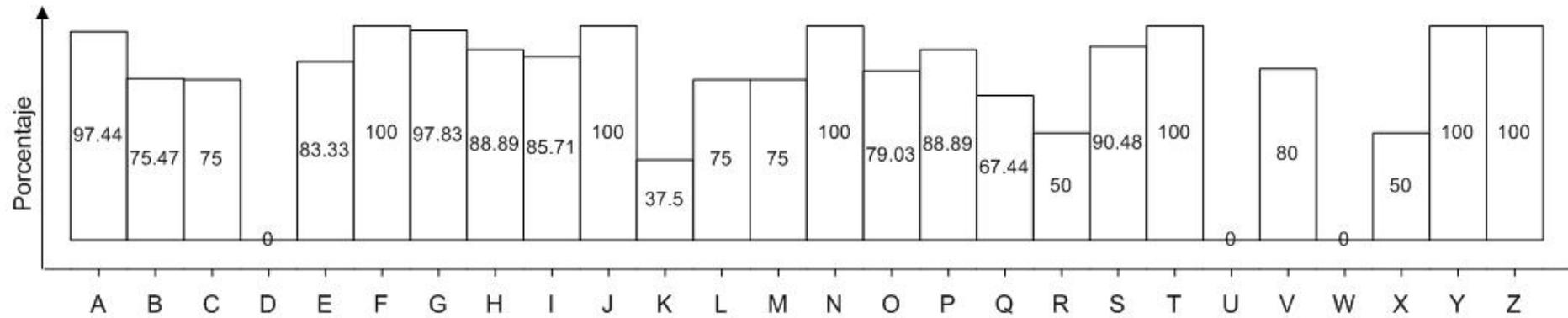


Figura 3.16 Porcentaje de acierto a la salida del sistema

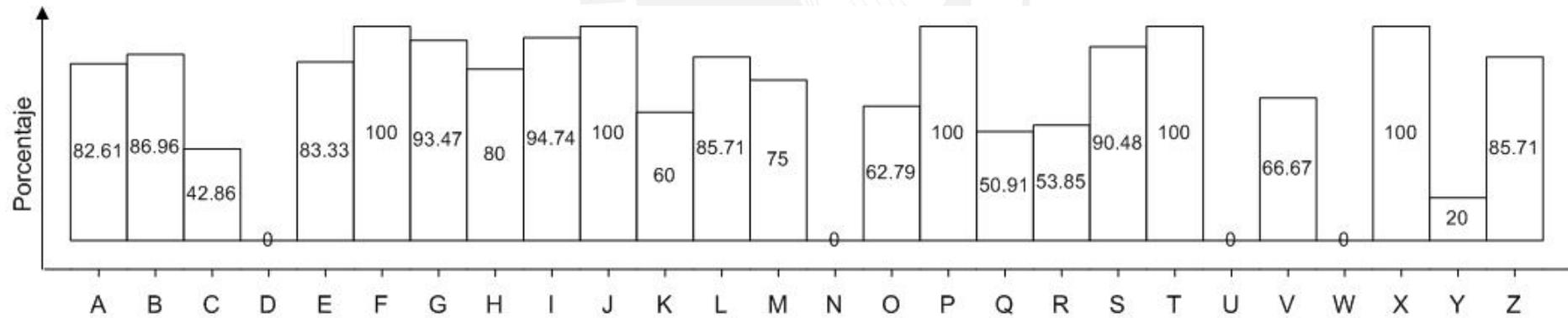


Figura 3.17 Porcentaje de eficiencia a la salida del sistema

El porcentaje de acierto final que se muestra a continuación ha sido establecido en función a su comparación con el que se presentó al comienzo de este capítulo. Al respecto, la siguiente gráfica muestra una mejoría notable.

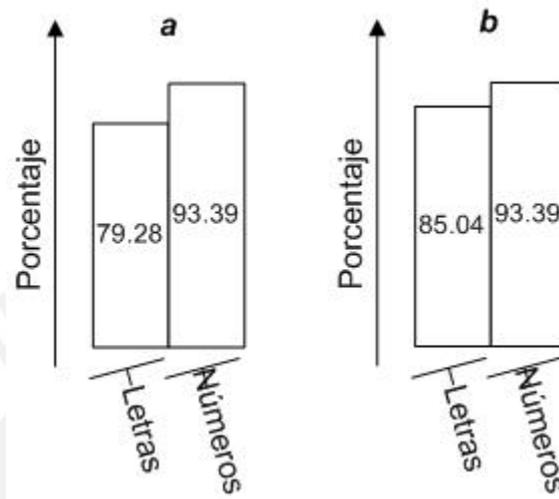


Figura 3.18 Porcentajes de acierto del sistema total

En la figura anterior, la diferencia entre el gráfico que corresponde a la letra **b** con respecto al que corresponde a la letra **a** se basa en que **b** toma en cuenta el guión de la placa en la datación de sus porcentajes totales. En ese sentido, se muestra el reconocimiento del guión mediante un porcentaje que señala una mayor eficacia o nivel de acierto. Otro dato curioso se da cuando se intenta analizar qué placas permiten una mayor eficiencia del proceso de reconocimiento: las placas de fondo blanco o de fondo de color, las placas que tienen dos letras y cuatro números o las placas que tienen tres letras y tres números.

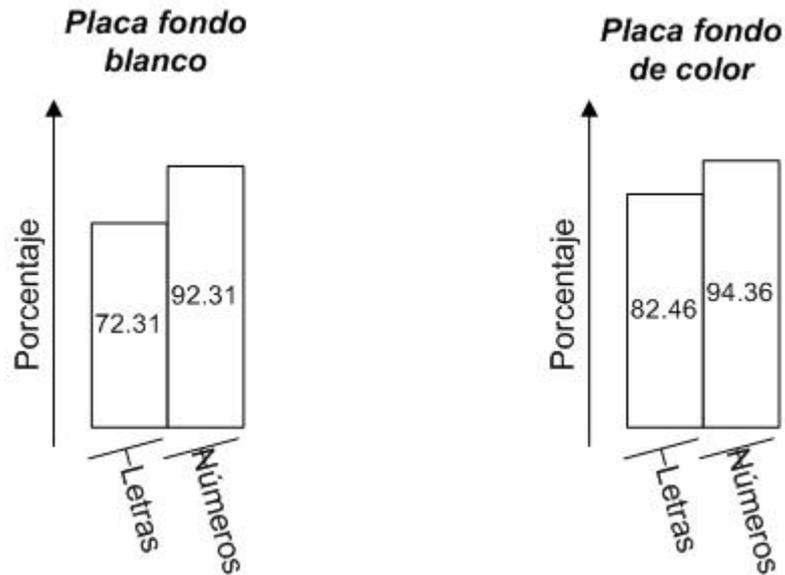
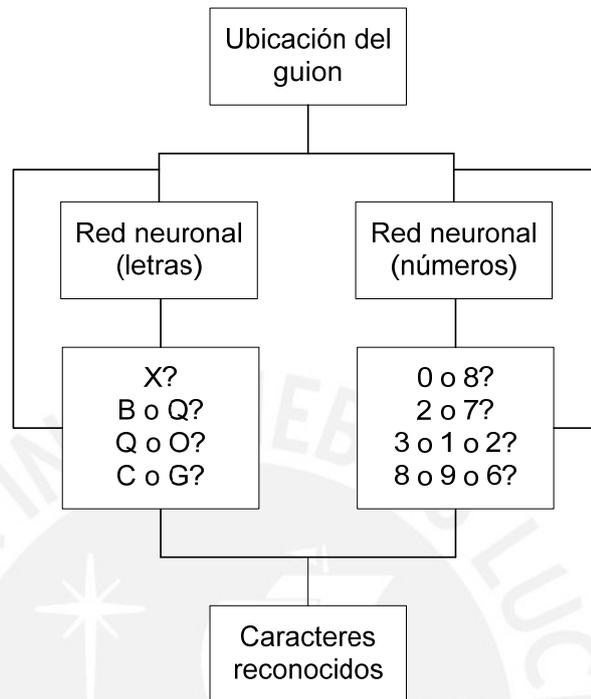


Figura 3.19 Porcentaje de eficiencia según el fondo de la placa

Estos resultados demuestran que el sistema es más eficiente para placas con fondo de color. Lo anterior se debe, principalmente, a que estas placas están constituidas por tres letras y tres números mientras que las de color blanco por dos letras y cuatro números. Al contar con dos letras una placa cualquiera, la posibilidad de que el sistema arroje un valor incorrecto es mayor, lo cual disminuye considerablemente la eficiencia del sistema. Para el caso de los números se podría pensar que sería mejor la eficiencia al contar con cuatro números, pero éste no es el caso ya que al presentarse demasiados caracteres la probabilidad de que al menos uno falle aumenta. En esa medida se tiene una mejor distribución en las placas de color, ya que como se muestra en la figura 3.19 tienen mejor probabilidades de acierto al contar con tres caracteres en cada región. Ello la señala como la distribución más óptima.

Los porcentajes mostrados se han establecido sin considerar al guión en el reconocimiento de los caracteres correspondientes a las letras.

3.3 Diagrama de bloques final del sistema propuesto



En el diagrama anterior se muestra la agregación de los bloques de confirmación. Según el sistema propuesto se puede visualizar que se cuenta con dos entradas: la salida a confirmar y el patrón de reconocimiento.

CAPÍTULO 4

RECOMENDACIONES

La etapa que se debe mejorar, principalmente, es la del reconocimiento de letras, ya que ésta presenta la menor tasa de efectividad dentro de los porcentajes globales de la investigación. En ese sentido, se necesita encontrar otro método que permita el reconocimiento de caracteres alfabéticos con un mayor grado de acierto. También, habría que tomar en cuenta que el problema de identificación de caracteres alfabéticos se debe no tanto al uso de redes neuronales, ya que ellas presentan muy buen rendimiento para el reconocimiento de números, si no más bien a la extracción de los patrones de reconocimiento. En esa medida se necesitan elaborar patrones que muestren mejor las características que se encuentren en un solo carácter.

A continuación se enumera una serie de alternativas para posibles estudios que permitan reducir la tasa de error.

- Esqueletización es una palabra técnica procedente del vocablo inglés “skeletonization”. La palabra da nombre al proceso mediante el cual se transforma una determinada forma u objeto de una imagen digital compuesta de una determinada cantidad de píxeles, en un objeto basado en líneas. Este proceso es muy útil a la hora de reconocer caracteres, dado que se constituye como una herramienta muy potente dentro de este campo digital. Una muestra del proceso de esqueletización son los mapas autoorganizados de Kohonen

(SOFM), capaces de detectar la topología y la distribución de probabilidad de diferentes patrones de entrada. No obstante, la red citada se basa en el método de descenso del gradiente alcanzando soluciones que son mínimos locales (pudiendo no ser globales). Además, la solución que encuentra la red depende también de los valores iniciales de los pesos neuronales respectivos. Ello conlleva a que un segmento no represente a un único grupo, sino también a dos, mientras otros pueden estar representados por más de un segmento. Este inconveniente hace que la red no sea adecuada para la detección. Una solución sería establecer o incorporar una fase más a la red utilizando diferentes mecanismos con la finalidad de controlar el tamaño de los segmentos.

- Varios métodos de selección consisten en tener varios patrones de reconocimiento y en elegir el que tenga mayor nivel de repeticiones. En esa línea, cada nivel por separado debe contar con un buen nivel de reconocimiento y en caso que se obtenga diferentes respuestas escoger el sistema que tenga menor error en la etapa de pruebas.
- Uso de super-resolución. Este proceso consiste en tener cuatro imágenes consecutivas correspondientes a una secuencia (en el caso de esta investigación al auto en movimiento) y en formar una nueva imagen en base a las anteriores. Para esto se toma un píxel de cada imagen que corresponda a las mismas coordenadas y se forma una nueva imagen cuadrada de cuatro píxeles. Este último proceso de formación de imágenes se repite sucesivamente hasta culminar con todos los píxeles. La apariencia que origina este proceso es mejor en cuanto a resolución, lo cual ayudaría como etapa previa al análisis del sistema.

- Otra alternativa es agregar una etapa final al proceso de reconocimiento y que esta última etapa sea la encargada de reconocer o dar la decisión final.
- Para la localización, se podría emplear la transformada de Hough realizando algunas variaciones o agregando una etapa que evite el problema de localización de la placa y las confusiones que suelen darse entre ésta, y un faro del vehículo.
- Análisis de varias imágenes del mismo vehículo. Se necesita tener tres o más secuencias de imágenes de un mismo vehículo y analizar cada imagen por separado, ya sea con el sistema propuesto en esta tesis o con algún nuevo sistema. Para ello se requiere colocar como resultado final a los caracteres que se repitan más veces. Las imágenes deben tener una pequeña variación en el espacio sin presentar problemas. Ello dado que las imágenes capturadas pertenecen a vehículos en movimiento.
- Selección previa. Se necesita extraer ciertas características que tiene algunos caracteres como, por ejemplo, el carácter B que se diferencia de la D por la barra en el medio que une sus extremos. Una vez identificados los rasgos de las letras tipo se realizaría una previa selección con la finalidad de formar grupos y de que cada grupo cuente con su propia red neuronal. Lo anterior haría más fácil la tarea de reconocimiento, ya que las redes tienen mejor exactitud en el reconocimiento a menor cantidad de salidas.
- Para la realización del proceso de reconocimiento en tiempo real, se hace necesaria la creación de dos módulos adicionales: un módulo de captura de video y otro de detección de un vehículo al interior de una imagen. Lo último

con el fin de no procesar imágenes que no contengan vehículos y no confundir vehículos con otros elementos, como peatones o aves, por ejemplo.

- También sería necesario crear un módulo de selección del grupo de imágenes que contienen a un mismo vehículo, lo cual permitiría la selección automática de las mejores que se pasaran a procesar. Lo anterior significaría un ahorro en esfuerzo y la posibilidad de una mayor rapidez en el proceso de localización y reconocimiento automático de placas vehiculares.



CONCLUSIONES

El sistema fue evaluado mediante la utilización de 200 imágenes capturadas, las que permitieron ensayar la estimación de la placa localizada y el reconocimiento de los caracteres. En el primer caso se obtuvo una localización correcta de 80.4%, en el segundo caso, un éxito de 79.28% para las letras y 93.39% para los dígitos. Estos valores fueron posibles gracias a que no dejó toda la tarea de reconocimiento a las redes neuronales, ya que si bien éstas obtuvieron un buen nivel porcentual (69.64% y 80.92% para las letras y dígitos, respectivamente), este se pudo mejorar con la ayuda de bloques adicionales de confirmación de salida que ayudaron en el reconocimiento.

La diferencia porcentual entre las letras y los dígitos se debe a que la red neuronal backpropagation funciona mejor o alcanza un menor error cuadrático medio cuando se tiene menos salidas, dado que sólo se tiene 10 neuronas de salida para los números mientras que 26 para las letras. Se tiene un éxito de 100% en la etapa de segmentación.

Se logra un éxito de 100% para diferenciar entre placas de tres letras y tres dígitos frente a las de dos letras y cuatro dígitos. Este sistema fue desarrollado principalmente para identificar placas en el Perú, país en el que se presentan dos tipos de placas. Sin embargo, con muy pocas modificaciones el sistema puede ser usado para identificar placas de diferentes países.

El sistema responde considerablemente bien ante caracteres con deterioro, ya sea por manchas de suciedad o despintado de los mismos. Los resultados obtenidos toman en cuenta a todo este tipo de placas. No se ha hecho ningún tipo de pre-selección a las imágenes capturadas, tan solo se han tomado en cuenta las condiciones impuestas inicialmente. Habría que agregar que en el Perú existen placas que presentan una máscara que une los caracteres en la parte inferior y provocan una sombra en los mismos. Este problema fue tomado en cuenta, dado que el sistema propuesto arroja resultados extraídos de placas con estas características. La muestra presentó más de la mitad de imágenes de placas que contenían dichas máscaras.

A la luz de estos resultados y a las dificultades que se presentan se concluye que el sistema propuesto para la localización y reconocimiento automático de la placa de un automóvil tiene un funcionamiento bastante óptimo.

BIBLIOGRAFÍA

- [1] CHUTAPE L Li, y XIAODONG,
1999 Automatic License Number Extraction and Its Parallel Implementation.
En: *Proceeding of The 25th Annual Conference of the IEEE, Industrial Electronic Society IECON'99*. Vol. (2): pp.704-709. California.
- [2] ZIMIC, FICZKO, MRAZ y VIRANT,
1997 The Fuzzy Logic Approach to the Car Number Plate Locating Problem.
En: *Proceedings of Intelligent Information System, IIS'97*. pp.227-230
Bahamas.
- [3] COETZEE, BOTHA. y WEBER,
1998 PC Based Number Plate Recognition System. En: *Proceedings of IEEE International Symposium on Industrial Electronics, ISIE'98*. Vol. (2): pp. 605-610. Pertoria.
- [4] VENA.
1993 Number Plate Recognition by Hierarchical Neural Network. En:
Proceedings of 1993 International Joint Conference on Neural Network.
Vol. (3): pp. 2106-2108. San Francisco.
- [5] GENDY, SMITH y LACHOWICZ,
1997 Automatic Car Registration Plate Recognition Using Fast Hough Transform. En: *Proceedings of International Carnahan Conference on Security Technology*. pp. 209-218. Canberra.
- [6] HARVEY, He M. G y VINAY, A. T,
1996 Hough Transform in Car Number Plate Skew Detection, *Proceedings of Fourth International Symposium on Signal Processing and Its Application, ISSPA96*. Vol. (2): pp. 593-596. Queensland: Gold Coast.

- [7] ZININO y ROVETTA,
1999 Visual Location Of License Plates by Vector Quantization. En: *Proceedings of the 1999 IEEE International Symposium on Circuits and Systems, ISCAS'99*. Vol. (4): 135-138. Orlando
- [8] FAUSETT Laurene
1994 Fundamentals of Neural Networks. Architectures, Algorithms, and Applications. New Jersey: Prentice-Hall and Englewood Cliffs.
- [9] VÁZQUEZ, M. Nakano y H. PÉREZ-Meana
2003 Automatic system for localization and recognition of vehicle plate numbers. En: *Journal of Applied Research and Technology*. Vol. (1): 63-77. D. F., México
- [10] CUI Yuntao y HUANG Qian
Automatic License Extraction from Moving Vehicles. En: *Siemens Corporate Research, Inc*. New Jersey.
- [11] GONZÁLEZ Rafael C. y WOODS Richard E.
1996 Tratamiento digital de imágenes. Massachusetts: Addison-Wesley Iberoamericana, S. A.
- [12] MORENO MONTAÑO Juan José
2002 Redes neuronales artificiales aplicadas al análisis de datos. Tesis (Dr.) Universitat De Les Illes Balears. Facultad de Psicología Palma de Mallorca, 220p.
- [13] VILLANUEVA ESPINOZA María del Rosario
2002 Las Redes Neuronales Artificiales y su Importancia Como Herramienta en la Toma de Decisiones. Tesis (Lic.). Universidad Mayor de San Marcos. Facultad de Ciencias Matemáticas. Lima, 68p.
- [14] ATENCIA Javier y NESTAR Raúl
2001 Aprenda MatLab 6.0 Como si Estuviera en Primero. San Sebastián

- [15] INTERNATIONAL PARKING INSTITUTE
Frequently Asked Questions. International Parking Institute. 2002. [Consulta: 18 de enero del 2004]. <http://www.parking.org/Value/FAQ/Default.aspx>
- [16] MARINA STASENKO – 2001
Number of Cars. The Physics Factbook. 29 agosto del 2003. [Consulta: 25 de febrero del 2004]. <http://hypertextbook.com/facts/2001/MarinaStasenکو.shtml>
- [17] SOLARI De LA FUENTE, Martín.
“Realidad Actual y Estrategias para la Prevención de los Accidentes de Tránsito”. Seguridad y educación vial. Touring Perú. el 2004. [Consulta: 23 de enero del 2005]. <http://www.touringperu.com.pe/gxpfiles/sitio/ztema10.htm>

