

Sistema de Visión Artificial para la Detección y el Reconocimiento de Señales de Tráfico basado en Redes Neuronales

Casimiro. A. Rocha

UNIVERSIDAD AUTONOMA DEL CARIBE, Barranquilla, Colombia, crocha@uautonoma.edu.co

José R. Escorcía Gutierrez

UNIVERSIDAD AUTONOMA DEL CARIBE, Barranquilla, Colombia, jescorcia@uautonoma.edu.co

RESUMEN

Hoy en día, las actividades que pasan desapercibidas por el hombre o la falta de información necesaria para comprender situaciones son causas de muchos accidentes a nivel general. En este proyecto se pretende orientar al hombre en su actividad de conducción por medio de visión artificial. La visión artificial ha venido trabajando en muchas disciplinas como Biomedicina, Reconocimiento de rostros, Reconocimiento de huellas dactilares, etc. Para este proyecto se ha diseñado un sistema de detección basado en redes neuronales y momentos invariantes capaz de adaptarse y entrenarse con ciertas señales de tránsito con la meta de asistir al conductor de no cometer una infracción o en el peor de los casos un accidente la red neuronal será capaz de reconocer una señal de tránsito a cierta distancia para que así el conductor a priori tenga el conocimiento de esta. El sistema fue implementado en MATLAB y este presenta mejorías frente a sistemas basados en lógica difusa o en procesamiento de imágenes básico.

Palabras claves: Momentos invariantes, Redes Neuronales, segmentación

ABSTRACT

Nowadays, the activities that go unnoticed by the man or the lack of information needed to understand situations are the causes of accidents at large. This project aims to guide men in their activity of driving through artificial vision. The vision has been sold to work in many disciplines such as biomedicine, face recognition, fingerprint recognition, and so on.

For this project, it has been created a detection system based on neural networks and moment invariants capable of learning and train with some traffic signals with the goal of assisting the driver not to commit a breach or in the worst case an accident, the neural network will be able to recognize a traffic signal some distance so that the driver has a priori knowledge of this.

The system was implemented on MATLAB and this shows better performance than others based in fuzzy logic and basic image processing.

Keywords: Invariant moments, neural networks, segmentation.

1. INTRODUCCIÓN

En recientes años, el uso de los sistemas inteligentes para mejorar situaciones de la vida diaria, debido a que estos sistemas se adaptan a la situación o al entorno donde se desea realizar la aplicación. Es por esto que técnicas de soft-computing, no solo son utilizadas en el procesamiento de señales e imágenes, sino que puede ser usadas en una gran variedad como la bioingeniería (Watrous and Shastri, 1987), (Center and Verna, 1998), la gestión y predicción de sistemas de inventarios, procesamiento paralelo y no lineal, telecomunicaciones, control automático, entre otros.

Entre estas aplicaciones, una de las que más ha recibido aportes de estos sistemas inteligentes es la visión artificial. Como se muestra en muchas aplicaciones es posible utilizar los sistemas inteligentes para detectar el movimiento de personas utilizando video y además identificarlas. También se pueden tener combinaciones de aéreas como la robótica y la visión artificial para detectar movimientos en robots móviles como el propuesto en Thorpe, Carlson et al, 2003.

Entre los sistemas inteligentes usados actualmente podemos encontrar, la lógica difusa propuesto por L. Zadeh (Zadeh, 1965), los algoritmos genéticos (Barricelli, 1954), los filtros Kalman (Kalman,1962), los sistemas multiagentes (Wooldridge, 2002) y cada uno de ellos presenta una forma de inteligencia que es capaz de aprender y adaptarse al entorno y situación que se encuentra.

Otra técnica de inteligencia artificial que ha tomado bastante auge en especial en aquellos que requieren la clasificación de datos o de información han sido las redes neuronales. La mayoría de los sistemas actuales de cómputo se han diseñado para llevar a cabo funciones matemáticas y lógicas a una velocidad que resulta asombrosamente alta para el ser humano. Sin embargo la destreza matemática no es lo que se necesita para solucionar problemas de reconocimiento de patrones en entornos ruidosos, característica que incluso dentro de un espacio de entrada relativamente pequeño, puede llegar a consumir mucho tiempo.

Uno de los usos que han tenido las redes neuronales es en la detección y reconocimiento de imágenes; que para este proyecto resulta importante debido a que de una imagen se desea tomar solo aquella parte de la imagen que contenga la señal de tránsito.

El principal objetivo de este proyecto es el de proponer un sistema de detección y reconocimiento de señales de tránsito para la prevención de accidentes, ya que ante las cifras de mortalidad por accidentes de tránsito recopiladas por el Fondo de Prevención Vial (Accidentalidad Vial, 2007), las cuales a partir de las estadísticas para el año 2007, se presentaron en Colombia 181.076 accidentes de los cuales 5.409 personas terminaron muertas, resulta importante garantizar a los conductores y pasajeros las herramientas necesarias para poder prevenir cualquier consecuencia por la no detección de una señal de tránsito.

Se desea llevar a cabo el diseño de una red neuronal capaz de reconocer las señales de tránsito en rango de distancia con el objetivo de poder asistir al conductor. Con esto se pretende saber que señales de tránsito existen en una imagen tomada para su análisis.

Para este sistema se pretende además estudiar las señales de tránsito en Colombia, cuáles son estas, qué características tienen, físicamente como se encuentran compuestas. Todo esto teniendo en cuenta que las señales de tránsito usadas en Colombia y América Latina son distintas a las que se manejan en Europa y Asia principalmente. La red neuronal a diseñar debe ser adaptable a todo tipo de señales para que el problema no quede limitado a las instancias presentes. Aunque, la red no opere en tiempo real se busca que esta sea capaz de llevar a cabo el proceso de reconocimiento en el menor tiempo posible, tiempo que estará supervisado por el número de operaciones que se lleve a cabo en la herramienta computacional de MATLAB. La etapas que se proponen en esta artículo son las de un marco teórico donde se estudian las señales de tránsito en Colombia, que métodos de descripción de color y de forma se usan para estas; luego se propone la estructura del sistema en diagramas de bloques y se estudia cada diagrama hasta obtener el mejor resultado. Finalmente se realizan unas conclusiones en base a los resultados obtenidos.

2. MARCO TEÓRICO

2.1 SEÑALES DE TRÁNSITO

Las señales de tráfico, son los signos usados en postes o pintadas en la calle ubicadas en el lado de caminos usadas para impartir la información necesaria a los usuarios que transitan por un camino o carretera, en especial los conductores de vehículos. Puesto que las diferencias de idioma pueden crear barreras, las muestras internacionales usan símbolos en lugar de palabras. Se han desarrollado principalmente en Europa y se han

adoptado en la mayoría de los países. La convención de Viena sobre muestras del camino y las señales del 8 de noviembre de 1968 define ocho categorías de muestras:

- A. Señales de peligro
- B. Muestras de la prioridad
- C. Muestras prohibitorias o restrictivas
- D. Muestras obligatorias
- E. Muestras de regla especiales
- F. Información, instalaciones, o muestras del servicio
- G. Dirección, posición, o muestras de la indicación
- H. Paneles adicionales

Sin embargo, los países y las áreas categorizan muestras del camino de diversas maneras. En Estados Unidos, el tipo, la colocación y los estándares gráficos de las señales de tráfico y de las marcas del pavimento se regulan legalmente. Otros países, como el resto de países del continente americano y Colombia, así como Australia suelen seguir el sistema estándar estadounidense, que es diferente al europeo (que por lo general es seguido por la mayoría de países africanos y de Oriente Medio).

Para el caso de Colombia, según el código del tránsito (Código de tránsito, 2002) las señales de tránsito se encuentran divididas en: reglamentarias (rojas), preventivas (amarillas), informativas (azules) y transitorias (anaranjada).

2.2 SEGMENTACIÓN

La segmentación es un paso imprescindible en diversos procesos de tratamiento de imagen. Entre otros, es necesaria para tomar medidas sobre una región, para realizar reconstrucciones tridimensionales de una zona de la imagen, para la clasificación o diagnóstico automático o para reducir la información de las imágenes. Si de una serie de imágenes para un determinado estudio sólo nos interesa una región concreta podemos segmentarlas y almacenar sólo las regiones para el análisis posterior.

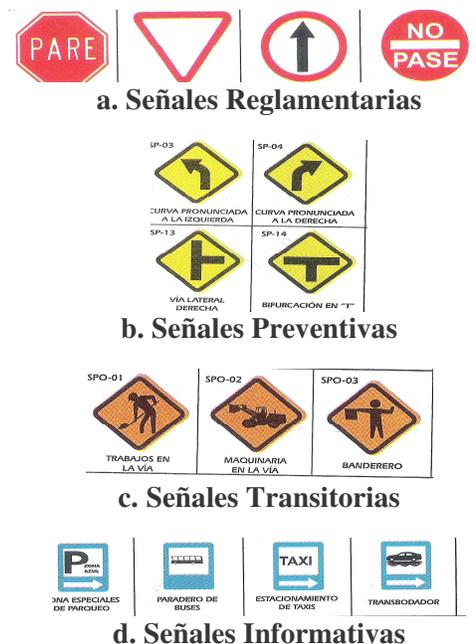


Figura 1. Ejemplos de señales de tránsito según el código de tránsito colombiano.

Aunque a la vista, la detección de regiones pueda parecer una tarea sencilla, nos encontramos con una serie de dificultades a la hora de realizar la segmentación de una imagen.

- *Artefactos*: Por problemas en la adquisición podemos encontrar ruido en la imagen, que distorsiona las características de las distintas regiones.
- *Volumen parcial*: Al ser las imágenes representaciones 2D de objetos 3D podemos encontrar en la imagen información que no es propiamente de ese plano y que limita la definición de las regiones.
- *Inhomogeneidad*: Los objetos no poseen siempre una iluminación uniforme, sino que muchas veces es sólo una textura característica lo que los diferencia de otros.
- *Forma*: Podríamos clasificar las formas de los objetos *grosso modo* en objetos patata, con una forma más o menos regular y convexa, objetos con forma de coliflor, con bordes no definidos y ciertas ramificaciones, y objetos spaghetti, objetos finos y muy ramificados. Estos últimos pueden ser difíciles de segmentar ya que pueden estar entrelazados con otros objetos.
- *Ruido*: La imagen puede ser de baja calidad debido a la presencia de ruido o a una mala adquisición de la misma.
- *Conocimiento previo*: Muchas veces empleamos en la segmentación visual el conocimiento previo que tenemos sobre una determinada región. Este conocimiento previo no es sencillo de obtener ni de ponerlo en estructuras aplicables a algoritmos.

En cuanto al grado de interacción del usuario en el proceso de segmentación, estos se pueden clasificar en:

- Manual: El usuario realiza la segmentación él mismo con la ayuda de una herramienta informática.
- Automática: El ordenador realiza todo el proceso de forma automática.
- Semi-automática o interactiva: El ordenador realiza el proceso, pero el usuario interviene en determinados momentos sobre el mismo para definir parámetros o corregir resultados. Es el método empleado generalmente.

2.3 REDES NEURONALES

Uno de los grandes avances logrados con el algoritmo Backpropagation (Redes Neuronales) es que esta red aprovecha la naturaleza paralela para reducir el tiempo requerido por un procesador secuencial para determinar la correspondencia entre unos patrones dados. Además el tiempo de desarrollo de cualquier sistema que se esté tratando de analizar se puede reducir como consecuencia de que la red puede aprender el algoritmo correcto sin que alguien tenga que deducir por anticipado el algoritmo en cuestión.

El algoritmo Backpropagation (Werbos, 1974) para redes multicapa es una generalización del algoritmo LMS, ambos algoritmos realizan su labor de actualización de pesos y ganancias con base en el error medio cuadrático. La red Backpropagation trabaja bajo aprendizaje supervisado y por tanto necesita un set de entrenamiento que le describa cada salida y su valor de salida esperado de la siguiente forma:

$$1) \{p_1, t_1\}, \{p_2, t_2\}, \dots, \{p_Q, t_Q\}$$

Donde p_Q es una entrada a la red y t_Q es la correspondiente salida deseada para el patrón q -ésimo. El algoritmo debe ajustar los parámetros de la red para minimizar el error medio cuadrático. Gráficamente, en la Figura 2 se observa la estructura de una red neuronal.

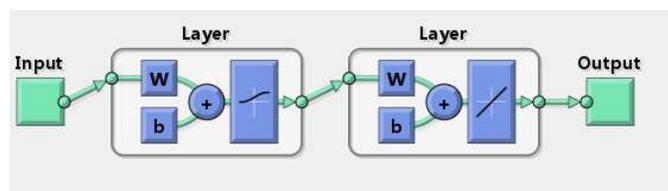


Figura 2. Ejemplo de la estructura de una red neuronal del tipo Backpropagation.

El entrenamiento de una red neuronal multicapa se realiza mediante un proceso de aprendizaje, de tal manera que los pesos W y los umbrales b satisfagan una ecuación óptima del sistema que deseamos obtener. Para realizar este proceso se debe inicialmente tener definida la topología de la red esto es: número de neuronas en la capa de entrada el cual depende del número de componentes del vector de entrada, cantidad de capas ocultas y número de neuronas de cada una de ellas, número de neuronas en la capa de la salida el cual depende del número de componentes del vector de salida o patrones objetivo y funciones de transferencia requeridas en cada capa, con base en la topología escogida se asignan valores iniciales a cada uno de los parámetros que conforma la red.

Es importante recalcar que no existe una técnica para determinar el número de capas ocultas, ni el número de neuronas que debe contener cada una de ellas para un problema específico, esta elección es determinada por la experiencia del diseñador, el cual debe cumplir con las limitaciones de tipo computacional.

2.4 MOMENTOS INVARIANTES

Los momentos invariantes son descriptores de forma que permiten identificar características de objetos en imágenes que no dependen del tamaño, la posición, la rotación, las distorsiones, entre otros. Estos fueron inicialmente implementados por Hu (Hu, 1962), en el que a partir de los momentos centrales y los momentos simples es posible adquirir características únicas para distintas imágenes. A partir de estos momentos es posible reconocer distintos objetos de una imagen teniendo en cuenta el conocimiento a priori de los momentos que se deben usar como se propone en Rizon (Rizon et al, 2006). Para este proyecto se usan los momentos invariantes como alternativa de reconocimiento de cada una de las señales de tránsito a clasificar.

Para las señales de tránsito de prueba se utilizan se han escogido las señales PARE, GLORIETA y CURVA PRONUNCIADA A LA IZQUIERDA, debido a que estas representan las características más importantes tanto en forma como en color de las señales de tránsito. La señal de PARE es del tipo reglamentaria y a diferencia de la mayoría de señales reglamentarias esta es hexagonal; tanto la GLORIETA como la CURVA PRONUNCIADA A LA IZQUIERDA son preventivas y su forma externa es un rombo.

3. PRUEBAS DEL SISTEMA

Se ha seguido una metodología basada en iteraciones controladas. Básicamente, consiste en realizar cuatro fases: inicio, elaboración, construcción y transición. Cada una de estas fases se subdivide en iteraciones que incluyen los pasos de diseño cumpliendo con el diagrama de bloques de la Figura 3: Transformación, subsegmentación, segmentación y normalización y reconocimiento de las señales de tránsito. Es decir, las necesidades y requerimientos se van refinando y corrigiendo en cada iteración, y no están completamente definidos desde el principio del proyecto. No obstante, cada paso en la iteración tendrá un peso distinto dependiendo de la fase en que se encuentre el proyecto. A continuación se describe cada una de las etapas para llevar a cabo el presente trabajo.

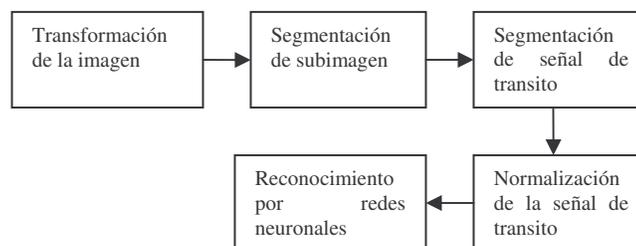


Figura 3. Diagrama de bloques del sistema de detección y reconocimiento de señales de tránsito.

1. Transformación de la imagen: Para esta etapa lo que se pretende es convertir las imágenes digitales en datos que se pueden procesar y comprender. Para esto se utiliza la herramienta de simulación de MATLAB que tienen en su Toolbox de procesamiento de imágenes, que permite convertir una imagen en una matriz de datos de RGB (Rojo, verde y azul), aún así es posible convertir en otros formatos como por ejemplo el HSI

o el HSV que son formatos los cuales permiten separar las características de color de las imágenes separando la intensidad que puede existir. Este formato es ventajoso para la detección de señales de tránsito ya que estas pueden variar su intensidad fácilmente, esto debido a que las señales de tránsito en Colombia se encuentran revestidas de material reflectivo, esto con el objetivo de poder observar las imágenes en la oscuridad. El inconveniente de estos formatos de conversión es que ninguno es lineal y los colores de la señal quedan en el mismo rango.

2. Segmentación de la subimagen: Después de transformar la imagen, se obtiene la subimagen que contenga la señal de tránsito para realizar la segmentación a partir de los bordes característicos de la imagen. Esto se realiza por dos razones: la primera porque al reducir el tamaño de la imagen las fuentes de ruido también se pueden reducir como se observa en la Figura 4. Este ruido proviene de varias fuentes, para entre ellos se pueden considerar el color de las placas vehiculares de los carros particulares que se maneja en Colombia, estas son exactamente del mismo color que las señales preventivas. Lo cual reducirá la efectividad del sistema propuesto. También taxis y obreros de mantenimiento público pueden ser fuentes de ruido.



Figura 4. Subsegmentación de imagen. Existen distintas fuentes de ruido como los STOPS de los carros, las placas vehiculares y los obreros de mantenimiento público.

Otra razón es que se espera que el sistema detecte las señales a grandes distancias, de tal forma que la probabilidad de detectar la imagen cuando el vehículo este cerca sea baja y se garantice que se prevenga a tiempo la señal de tránsito. La imagen se subdividirá a $1/16$ de la imagen original y a $1/4$ de la imagen. Para que no ocurra una doble detección, se referencia cada subimagen y si se encuentran traslapadas solo tomará la que presente un mejor rendimiento.

3. Segmentación de la señal de tránsito: Después de la segmentación primaria, se realiza una segunda segmentación de la cual se obtendrán las características internas de la imagen como lo es la señal de pare, cruce, etc. Esto se realiza de dos formas: la primera determinando los rangos RGB o HSV donde se deben encontrar las señales de tránsito. La segunda es utilizando redes neuronales que procesen las entradas RGB o HSV de tal forma que clasifique solo los colores que se asemejen bastante a los de las señales de tránsito.
4. Normalización de la información de la señal de tránsito: Debido a que la señal de tránsito puede ser tomada a distancias diferentes se debe obtener una dimensión de referencia que permita la comparación y el análisis con respecto a la imagen de referencia.
5. Reconocimiento de las señales de tránsito: Por ultimo con la red neuronal entrenada, se obtiene que señal de tránsito es la que se está observando. Para entrenar la red neuronal se toman los momentos invariantes básicos, los momentos invariantes difusos y los momentos invariantes afines de cada una de las imágenes

de prueba. El tipo de red neuronal diseñada será de tipo Backpropagation, ya que reduce el error correlacionándolo con los campos a clasificar. El número de entradas y salidas serán los mismos (tres) ya que lo que se busca es segmentar la imagen dejando solo los colores deseados.

Para las señales de prueba se han tomado 73 imágenes de muestra las cuales se encuentran en distintas posiciones, distancias, momentos del día y obstáculos; de tal forma que representen en gran medida los fenómenos que pueden ocurrir al momento de tomar una imagen en ambientes externos. De estas 73 imágenes, solo 30 serán usadas para entrenar las redes neuronales tanto para el proceso de segmentación como para el de reconocimiento a partir de sus momentos invariantes. El resto serán usadas para verificar la eficiencia de las redes neuronales. Para los momentos invariantes se escogieron los momentos I1, I2 e I7 de Hu, los momentos A11 y A22 de los afines y los momentos B11 y B22 de los difusos.

4. RESULTADOS

En la Tabla 1 se presentan las características de las redes neuronales que se propusieron para el proceso de segmentación. La idea de esto es el de obtener a partir de prueba y error que red escogida tiene el mejor rendimiento y permita obtener la mejor segmentación. Para medir el rendimiento de cada red neuronal se analiza el porcentaje de señal de tránsito segmentada. En la Tabla 2 se presentan el número de imágenes que alcanzan un porcentaje de segmentación alcanzado.

Tabla 1. Características de las redes neuronales propuestas.

Neuronas	Capa 1	Capa 2	Capa 3	Capa 4	Capa 5
Red 1	3	4	7	4	3
Red 2	3	7	4	3	
Red 3	3	4	8	3	
Red 4	3	4	3		

Tabla 2. Rendimiento de las redes neuronales frente al porcentaje de segmentación

Imágenes Segmentadas	>20%	>40%	>50%	>70%	>90%
Red 1	41	40	38	31	28
Red 2	41	32	31	30	25
Red 3	41	41	38	34	32
Red 4	38	36	29	28	24

A pesar de que la red neuronal número 1 es la que mayor número de capas posee, no es la que presenta el mejor rendimiento, en especial para mantener una segmentación por encima del 70%. Para este caso la que mejor representa la red que representa las mejores características es la red número 3; ya que el número de imágenes que han sido segmentadas por encima del 90% equivalen al 74.4% de imágenes simuladas. Aunque este rendimiento no es concluyente, la red neuronal propuesta (red 3) es la mejor candidata a nuestro sistema. Para determinar que

el sistema presenta un mejor rendimiento, se ha comparado frente al método de umbralización el cual se basa en el rango de colores que se debe encontrar la imagen a segmentar. Para esto se toma la información de las señales de tránsito en el espacio de color RGB. En la Tabla 3 se presenta las características en color de las señales de tránsito.

Tabla 3. Rango de colores de señales de tránsito.

Rango de Colores	R	G	B
Señal reglamentaria (Rojo)	183-230	49-105	38-62
Señal preventiva (Amarillo)	238-255	220-242	0-15

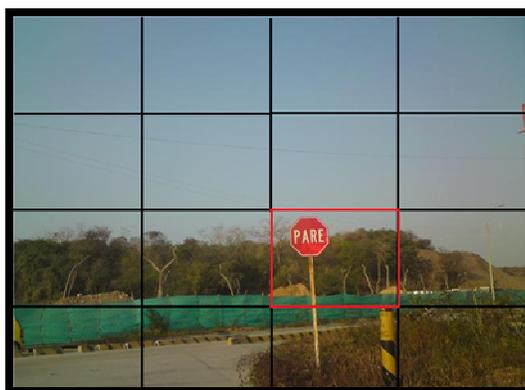


Figura 5. Imagen correctamente segmentada y detectada por la red neuronal numero 3.

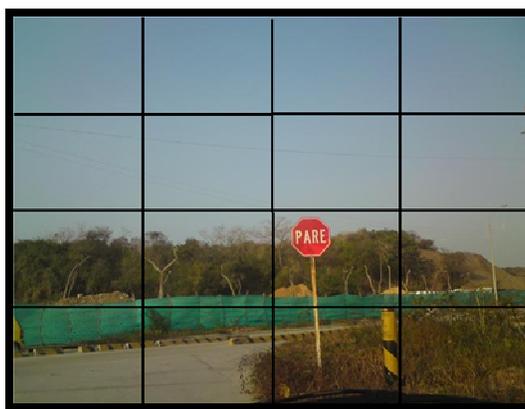


Figura 6. Imagen no detectada por el método de umbralización.

En la Figura 5 se muestra una señal de PARE correctamente detectada mientras que la Figura 6 presenta la misma señal de tránsito que no ha sido detectada por el método de umbralización. Esto se debe a que el método aunque posea la información sobre el rango de colores de la señal, no correlaciona el espacio de colores donde se encuentra la imagen, es decir no la clasifica. Esto si ocurre en el caso de de la red neuronal.

Finalmente, para realizar el reconocimiento como anteriormente se había comentado, se utilizan siete momentos invariantes con los cuales se representan las traslación, rotación, reflexión, difusión y afinidad. Para esto se

diseñan dos redes neuronales las cuales tendrán como salida la red neuronal que deben representar. En la Tabla 4 se presentan las redes neuronales Backpropagation escogidas. En la Tabla 5 se presenta el porcentaje de imágenes correctamente detectadas de cada uno de las redes neuronales. Se observa que la segunda red presenta un mejor rendimiento que el primero, de hecho la segunda red neuronal realiza una mejoría de la segmentación, teniendo una mayor cantidad de imágenes que las segmentadas en por lo menos un 90%.

5. CONCLUSIONES

El sistema de detección y reconocimiento propuesto tiene la ventaja que se adapta a fenómenos que son ajenos a las señales de tránsito, como puede ser la reflectividad o la opacidad de una imagen, o distorsiones por obstáculos pequeños o daño en esta. Frente a soluciones basadas en procesamiento de imágenes básico, el porcentaje de acierto es considerablemente mayor el sistema propuesto por redes neuronales (88.6% el basado por redes neuronales versus 48% el basado en segmentación por umbralización). Aparte, el tiempo de respuesta del sistema basado en redes neuronales resulta ser considerablemente superior que el propuesto por umbralización en un factor de 2,797.

Tabla 4. Características de las redes neuronales propuestas.

Neuronas	Capa 1	Capa 2	Capa 3	Capa 4
Red 1	7	4	2	
Red 2	7	7	4	2

Frente a otros sistemas propuestos como el de Fleyed, 2005 tiene la ventaja que al subdividir la imagen, el ruido producido por objetos con los mismos colores tiende a ser mucho menor y a su vez los valores obtenidos de los momentos invariantes tendrán menos ruido.

Tabla 5. Señales reconocidas correctamente.

Señales Acertadas	Red 1	Red 2
Señal reglamentaria (Rojo)	9	12
Señales preventivas (Amarillo)	21	26

Para próximas investigaciones se pretende realizar modificaciones respecto a la subdivisión de la imagen y también se pretende buscar otro método de invarianza que tenga mejores resultados como el propuesto por Reza, Erol et al, 2006, que presenta los momentos de Zernike.

REFERENCIAS

- Amayeh, G, Erol, A, Bebis, G and Nicolescu, M (2005). "Accurate and efficient computation of high order Zernike moments". First ISVC: 462--469.
- Barricelli, N (1957). "Symbiogenetic evolution processes realized by artificial methods". *Methodos*: 143–182.
- Código Nacional de Tránsito Terrestre (2002). Ley 769 de 2002.
- Kalman, R.E. (1960). "[A new approach to linear filtering and prediction problems](#)". *Journal of Basic Engineering* 82 (1): 35–45.
- Mohamed Rizon, Haniza Yazid, 2Puteh Saad, Ali Yeon Md Shakaff, Abdul Rahman Saad, Mohd Rozailan Mamat, Sazali Yaacob, 1Hazri Desa and M. Karthigayan. "Object Detection using Geometric Invariant Moment". *American Journal of Applied Sciences* 2 (6): 1876-1878, 2006.
- Watrous, R and Shastri, L (1987). "Learning phonetic features using connectionist networks: an experiment in speech recognition". *Proceeding 1st IEEE International Conference in Neural Networks*.

Werbos, P (1974). "Beyond regression: New tools for prediction and analysis in the behavioral sciences". Ph.D dissertation, Harvard University.
Michael Wooldridge, *An Introduction to MultiAgent Systems*, John Wiley & Sons Ltd, 2002.
L. A. Zadeh (1965) "Fuzzy sets". *Information and Control* 8 (3) 338–353.

Autorización y Renuncia

Los autores autorizan a LACCEI para publicar el escrito en los procedimientos de la conferencia. LACCEI o los editors no son responsables ni por el contenido ni por las implicaciones de lo que esta expresado en el escrito