

# RECONHECIMENTO AUTOMÁTICO DE PLACAS DE VEÍCULOS AUTOMOTORES ATRAVÉS DE REDES NEURAIS ARTIFICIAIS

**Bruno Clemente Guingo**

bcinguigo@posgrad.nce.ufrj.br

**Antonio Carlos Gay Thomé**

thome@nce.ufrj.br

**Roberto José Rodrigues**

rjr@nce.ufrj.br

Área de Ensino e Pesquisa, NCE/IM

Universidade Federal do Rio de Janeiro

Cidade Universitária - Ilha do Fundão, Caixa Postal 2324 - CEP: 20001-970 - Rio de Janeiro-RJ

BRASIL

## RESUMO

Neste artigo descreve-se a estratégia adotada para a concepção e desenvolvimento de um sistema para reconhecimento automático de placas de veículos automotores. A abordagem do problema é modular, dividida em seis etapas, que cobrem desde a tomada da imagem até o reconhecimento de cada um dos caracteres que compõe a placa. Técnicas de processamento de imagens e de inteligência computacional - redes neurais, são intensivamente empregadas.

**Palavras-chave:** monitoração do tráfego, localização de carros roubados, controle de acesso, aplicação de multas, reconhecimento de placas, segmentação de caracteres, extração de características, redes neurais.

## ABSTRACT

This article describes the adopted strategy for the conception and development of a system for automatic recognition of automotives vehicles plates. We used a modular solution to approach this problem that is divided in six phases starting on the capture of the image and ending on the recognition of each one of the characters in the plate. Image Process and computational intelligence techniques more specifically neural networks system, are strongly used.

**Key-words:** traffic controllers, stolen localization cars, access controls, fines application, plates recognition, characters segmentation, extraction of characteristics, neural networks.

**Categoria do Artigo:** Sessões Técnicas - Inteligência Artificial

# RECONHECIMENTO AUTOMÁTICO DE PLACAS DE VEÍCULOS AUTOMOTORES ATRAVÉS DE REDES NEURAIIS ARTIFICIAIS

## RESUMO

Neste artigo descreve-se a estratégia adotada para a concepção e desenvolvimento de um sistema para reconhecimento automático de placas de veículos automotores. A abordagem do problema é modular, dividida em seis etapas, que cobrem desde a tomada da imagem até o reconhecimento de cada um dos caracteres que compõe a placa. Técnicas de processamento de imagens e de inteligência computacional - redes neurais, são intensivamente empregadas.

**Palavras-chave:** monitoração do tráfego, localização de carros roubados, controle de acesso, aplicação de multas, reconhecimento de placas, segmentação de caracteres, extração de características, redes neurais.

## ABSTRACT

This article describes the adopted strategy for the conception and development of a system for automatic recognition of automotives vehicles plates. We used a modular solution to approach this problem that is divided in six phases starting on the capture of the image and ending on the recognition of each one of the characters in the plate. Image Process and computational intelligence techniques more specifically neural networks system, are strongly used.

**Key-words:** traffic controllers, stolen localization cars, access controls, fines application, plates recognition, characters segmentation, extraction of characteristics, neural networks.

## 1 Introdução

Com o constante crescimento da frota mundial de veículos surgem problemas cada vez mais complicados de serem rapidamente resolvidos e que por isso, carecem de soluções criativas e automatizadas. Existe hoje uma grande necessidade da engenharia de tráfego em conseguir informações rápidas e precisas no que diz respeito ao reconhecimento de placas dos automóveis, pois através dele é possível aumentar a eficiência do controle e da monitoração do tráfego, bem como baixar os custos de operação.

O reconhecimento de placas não tem utilização somente no controle e na monitoração do tráfego. Existem outras aplicações que podem ser beneficiadas, tais como: localizar carros roubados, aplicar multas aos condutores infratores, controlar o acesso a estacionamentos, coletar estatísticas de fluxo de entrada e saída de certos locais ou mesmo gerar subsídios para a área de turismo. Mesmo nas grandes metrópoles brasileiras, a identificação do veículo, a partir das fotografias tiradas pelos inúmeros equipamentos espalhados nas vias públicas, ainda é feita manualmente e não em tempo real. Com um sistema automatizado o ganho não se restringe tão

somente ao tempo de resposta, mas principalmente na imediata disponibilidade da informação para uso em inúmeras outras aplicações, como nas áreas de segurança, planejamento e turismo.

Tendo em vista a grande concentração de veículos nos grandes centros e, que a maioria das grandes cidades brasileiras já utilizam câmeras para o controle de tráfego ou mesmo para a aplicação de multas, a utilização de um sistema de reconhecimento automático de placas seria um excelente aliado aos órgãos fiscalizadores de trânsito, aos estacionamentos de grande rotatividade, às concessionárias das rodovias privatizadas e a outros setores que possam fazer uso da informação colocada rapidamente à disposição.

## **2 Obtenção dos Dados**

Para a realização desta pesquisa, foi imprescindível a obtenção de uma grande quantidade de fotografias digitais, pois é através delas que o estudo e o desenvolvimento de um modelo computacional se torna possível. O Departamento de Estrada e Rodagem do Estado do Rio de Janeiro (DER-RJ), através da empresa Perkons Equipamentos Eletrônicos Ltda, firma que presta serviços de informática ao DER-RJ, forneceu um conjunto de 9.079 (nove mil e setenta e nove) imagens digitais de veículos. As imagens fornecidas foram obtidas através das diversas câmeras espalhadas pelas rodovias do Estado do Rio de Janeiro, utilizadas na fiscalização e aplicação de multas.

O uso de fotos obtidas em condições reais enriquece o estudo, uma vez que apresenta problemas concretos que precisam ser avaliados e abordados, tais como: neblina, chuva, luminosidade, ângulos, distâncias e velocidades diferentes de obtenção das imagens.

## **3 Análise das Imagens**

Foi realizada uma primeira análise no conjunto de imagens visando verificar a quantidade de fotografias capazes de serem identificadas pelo ser humano. Das 9.079 imagens, levantou-se que 42,4% não puderam ser identificadas, por diversos problemas, tais como: imagens de veículo sem placa, imagens sem veículo, completamente escuras, com excessiva luminosidade na placa, com caracteres mal definidos na placa e com baixo grau de resolução.

Em resumo, o estudo foi realizado com base nas 5.233 (cinco mil e duzentos e trinta e três) fotos restantes.

## **4 Metodologia Utilizada**

O sistema é constituído de forma modular, cabendo a cada módulo uma funcionalidade específica.

- **Módulo 1: Obtenção da fotografia do veículo**

Responsável pela captura da imagem. Nesta fase, as imagens são obtidas pelas câmeras espalhadas pelos diversos pontos das vias públicas. Este é o procedimento atualmente empregado pelos órgãos fiscalizadores de trânsito brasileiro. Na Fig.1 pode-se verificar como isto é feito.



Fig.1: Câmera de detecção

Existem três formas de obtenção das imagens: através de vídeo, de fotografias digitais ou de fotografias analógicas (convencional). Uma boa qualidade de imagem é fundamental para o desempenho do sistema de reconhecimento. Na Fig.2 pode-se observar o padrão de qualidade de uma imagem capturada pelo sistema utilizado pelo DER-RJ. Apesar da baixa definição, a imagem deste exemplo é considerada uma das melhores em relação a todo o conjunto de dados.



Fig.2: Imagem obtida pela câmera de detecção

- **Módulo 2: Localização da região da placa**

Responsável por localizar a placa ou regiões candidatas à placa. Tem como entrada a foto inteira e seu objetivo é fornecer como saída uma região onde esteja provavelmente localizada a placa.

- **Módulo 3: Seleção e extração da região da placa**

Responsável por selecionar, dentre as regiões candidatas, aquela mais provável, extrair a região da imagem e guardá-la em um arquivo para a execução da fase seguinte.

- **Módulo 4: Segmentação dos caracteres da placa**

Responsável por separar os caracteres um do outro e constituir sete novos arquivos menores, cada um contendo a imagem de apenas um dos caracteres que compõem a placa.

- **Módulo 5: Extração das características dos caracteres segmentados**

Responsável por extrair de cada caractere segmentado, informações tais que permitam que o mesmo seja mais facilmente classificado pelo módulo de reconhecimento.

- **Módulo 6: Reconhecimento dos caracteres**

Responsável por reconhecer cada caractere que compõe a placa, através das informações disponibilizadas pelo módulo de extração das características. O reconhecimento é realizado através das técnicas de redes neurais.

## **5 Módulos Correntemente em Desenvolvimento**

Os módulos do sistema de reconhecimento automático de placas estão sendo desenvolvidos conforme o seguinte:

### **5.1 Localização da região da placa**

Conforme o artigo 1º da resolução nº 45/98 do Conselho Nacional de Trânsito - CONTRAN [9], as dimensões (em milímetros) aceitas para as placas são as seguintes:

a) ideal: altura=130 - largura=400; b) máxima: altura=143 - largura=440; c) mínima: altura=117 - largura=360.

A dimensão dos caracteres da placa é de 63 mm de altura para todos os caracteres. A largura é definida para cada letra e algarismo, sendo o maior valor para a letra “W” (49 mm) e o menor para a letra “I” (10 mm). O processo de localização em estudo envolve a aplicação de algoritmos de processamento de imagens para a localização de regiões candidatas à placa. Após esta localização, as regiões da placa são extraídas e guardadas cada uma em um arquivo diferente, os quais são utilizados na fase seguinte.

### **5.2 Seleção e extração da região da placa**

Os algoritmos utilizados no módulo de localização da região da placa nem sempre são capazes de identificar uma única região para a placa, devido à diversidade de condições das diferentes fotografias. Há fotos em que os veículos não possuem placa e há fotos em que mais de uma região candidata é encontrada.

Um conjunto neural é então utilizado para, dentre as regiões de imagem encontradas, selecionar aquela com a maior probabilidade de ser uma placa. A decisão final sobre a aceitação ou não das regiões candidatas à placa também é feita por rede neural. Para suportar o estudo e a concepção de ambos os módulos, foi desenvolvido um programa que possibilita a extração manual de regiões com e sem a placa incluída.

### 5.3 Segmentação dos caracteres da placa

Após a extração da região que contém a placa, a próxima etapa é a segmentação dos caracteres, onde é realizada a separação de cada caractere um do outro, criando sete arquivos, cada um com a imagem de um caractere da placa [1], conforme mostra a Fig.3. Nem sempre a segmentação consegue separar todos os caracteres em sete arquivos distintos, pois devido aos problemas de baixa resolução da imagem e de distorção, alguns caracteres ficam com uma proximidade muito grande entre seus vizinhos, tendo como consequência à união de dois caracteres. O percentual de segmentação da placa em sete caracteres distintos, até o presente momento é de 76.61%, isto devido ao conjunto de dados do DER-RJ, ter em sua grande maioria fotos de baixíssima qualidade. Em fotos de boa qualidade a taxa de acerto da segmentação tem sido de 100%.



Fig.3: Segmentação dos caracteres da placa

Pode ser visto nas Fig.4 e 5 abaixo, um exemplo com a imagem inicial (entrada) e a imagem segmentada (saída) de uma placa.



Fig.4: Imagem inicial



Fig.5: Imagem segmentada

### 5.4 Extração das características dos caracteres segmentados

Nesta fase são extraídas características de cada imagem segmentada, de forma a compor uma assinatura para cada caractere. A extração é feita por uma técnica desenvolvida pelo grupo, que é derivada dos algoritmos de detecção de contorno [4] e [5]. A técnica é baseada na projeção do contorno da imagem sobre os lados de um polígono regular, posicionado em torno de cada caractere. Os vetores de características são baseados na computação de um conjunto de distâncias tiradas a partir do contorno da imagem até um polígono de referência.

O polígono de referência deve ser regular porém pode ter qualquer número de lados. Este deverá ser posicionado em torno da imagem-alvo como mostrado na Figura 6. Qualquer polígono pode ser construído com base nas equações Equações (1,2, 3 e 4).

O processo básico toma a distância de cada lado do polígono até o contorno da imagem e armazena esta distância em um vetor que também contém o número de lados e o número de pontos (tamanho do lado) para cada lado. O número de características extraídas pode não ser o mesmo para diferentes tipos de polígonos, não obstante o número de características desejadas seja o mesmo. O número de características computado é sempre muito próximo do desejado e isto acontece devido a diferença na geometria e inclinação dos lados para os polígonos escolhidos. Esta diferença é sempre pequena e portanto, não é relevante para o problema em si.

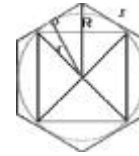
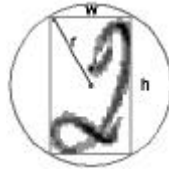


Figura 6a: círculo circunscrito à imagem    Figura 6b: polígono circunscrito ao círculo

$$r = \frac{1}{2} \sqrt{h^2 + w^2} \quad (1)$$

$$R = \frac{r}{\cos\left(\frac{P}{n}\right)} \quad (2)$$

$$R = \frac{1}{2} \sqrt{h^2 + w^2} * \cos\left(\frac{P}{n}\right)^{-1} \quad (3)$$

$$s = \sqrt{h^2 + w^2} * \tan\left(\frac{P}{n}\right) \quad (4)$$

Onde,

$r$  = raio do círculo circunscrito (metade da diagonal)

$h$  = altura da imagem

$w$  = comprimento da imagem

$n$  = número de lados do polígono

$R$  = raio do círculo circunscrito (em relação ao hexágono) – lado do hexágono

$s$  = tamanho do lado do hexágono.

Usando um algoritmo de reconstrução a partir dos vetores de características gerados (Figura 7) foi possível perceber que a fidelidade resultante não se mostrou favorável com polígonos de número ímpar de lados. Desta forma, ficou decidido concentrar a investigação em cima de polígonos com número par de lados, mais especificamente o quadrado (Figura 8).



Figura 7: reconstrução do processo de vetorização gerado.

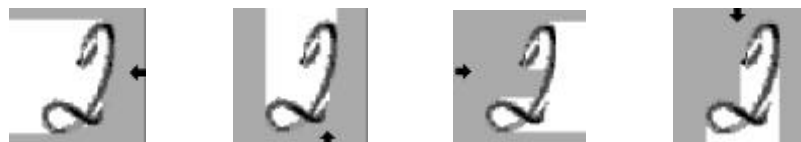


Figura 8: Seqüência de procedimentos de projeção para o quadrado

Como descrito anteriormente o vetor de entrada possui uma dimensão predefinida e as características são as distâncias computadas em número de pixels e tomadas a partir de cada lado do polígono até o contorno da imagem. Este método não só provou ter um bom poder de discriminação como também resolver problemas relacionados com escala, rotação e translação da imagem original. Uma vez que o *bitmap* esteja inicialmente centralizado dentro de um retângulo antes de ser circundado pelo polígono desejado, o problema da translação é automaticamente eliminado. A aplicação de um esquema de interpolação na tomada dos pontos para o cômputo da distância ao contorno resolve o problema de escala e, o problema da rotação, pode ser solucionado aplicando-se uma rotação circular no vetor resultante para a direita ou para esquerda, conforme for a inclinação do caractere.

O desempenho da técnica foi avaliado através do estudo de diferentes polígonos como mostrado na Fig.9 abaixo:

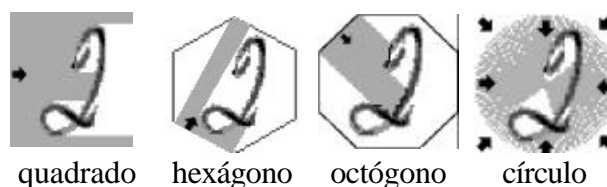


Fig.9: Métodos de extração de características

Neste trabalho, está sendo utilizado o método do quadrado, apontado em [6] como o mais adequado para este tipo de aplicação. O tamanho adotado para o quadrado é de 32 pixels para cada lado, o que totaliza um vetor de características com 128 medidas de distância.

Estes vetores foram gerados através de um programa especialmente desenvolvido para ler uma imagem bitmap e produzir os valores para todos os tipos de polígonos, assim como para diversos tamanhos dos respectivos lados.

Do conjunto de placas disponíveis para estudo gerou-se uma base de dados com letras e dígitos numéricos, que apresenta a seguinte distribuição.



Tabela 1: Distribuição das letras em cada classe

Letra	%	Letra	%	Letra	%	Letra	%
A	5.38	H	2.04	O	3.36	V	1.74
B	6.32	I	4.76	P	2.56	W	0.79
C	10.24	J	4.09	Q	2.68	X	2.14
D	1.83	K	10.13	R	3.06	Y	1.59
E	1.78	L	15.95	S	2.56	Z	1.44
F	0.98	M	3.13	T	2.23	<b>Total</b>	<b>7.70</b>
G	3.67	N	3.48	U	2.07		
<b>Total</b>	<b>30.2</b>	<b>Total</b>	<b>43.58</b>	<b>Total</b>	<b>18.52</b>		

Tabela 2: Distribuição dos dígitos numéricos em cada classe.

Dígito	%	Dígito	%	Dígito	%
0	9.80	4	10.30	8	7.80
1	9.35	5	9.65	9	11.20
2	10.40	6	10.90	<b>Total</b>	<b>19.00</b>
3	9.80	7	10.80		
<b>Total</b>	<b>39.35</b>	<b>Total</b>	<b>41.65</b>		

Pode-se verificar que a frequência de distribuição das letras não é homogênea, havendo uma incidência maior das letras “C”, “K” e “L” e muito menor das letras “F” e “W”. Já no conjunto de dígitos, a distribuição ficou mais homogênea.

## 5.5 Reconhecimento dos caracteres

A partir das características extraídas do conjunto de caracteres de uma placa, é dado início a fase de reconhecimento. A entrada da rede para cada placa é composta por uma matriz de tamanho 128x7. Os três primeiros para a rede de letras e os quatro restantes para a rede de dígitos. A saída para cada caractere foi configurada como um vetor ortogonal de dimensão 10 para os dígitos e 26 para as letras.[tabela 3 ].

As duas redes neurais são do tipo Feedforward Multi-Layer Perceptron [2] e [8], treinadas com o algoritmo *backpropagation*. O software utilizado foi o MatLab versão 5.2 [7]. A função de ativação utiliza um somatório simples das entradas ajustadas pelos seus respectivos pesos. Como função de propagação utilizou-se a função logística sigmoidal (“*logsig*” no MATLAB) tanto na camada escondida quanto na camada de saída.

Diversas topologias e arquiteturas foram testadas, onde os “melhores” resultados obtidos foram através de redes compostas de duas camadas, uma oculta e outra de saída. Ambas as redes operam com uma entrada de 128 características, obtidas pelo “método do quadrado”. A primeira camada, a oculta, é composta por 47 neurônios e a saída é por 10 neurônios, no caso da

rede especializada no reconhecimento de dígitos numéricos, e por 26 neurônios na rede especializada em letras.

A saída da rede neural foi selecionada de acordo com a estratégia “The Winner Takes All”, onde o maior valor entre todas as saídas foi alterado para um (1) e o restante foi alterado para zero (0). Dessa forma tem-se índice zero (0) de rejeição. O vetor de saída foi escolhido como sendo ortogonal para facilitar o processo de treinamento. A tabela 3 abaixo, mostra a codificação utilizada na saída da rede de dígitos numéricos e letras.

Tabela 3. Codificação da saída dos dígitos numéricos e letras.

Dígitos	Saída	Letras	Saída
0	0000000001	A	10000000000000000000000000000000
1	0000000010	B	01000000000000000000000000000000
...	...	...	...
9	1000000000	Z	00000000000000000000000000000001

- **Pré-processamento dos dados**

O conjunto de dados foi dividido em dois subconjuntos distintos: um para treinamento com 80% das amostras e outro para teste com os 20% restantes [3]. Todas as informações da entrada (treinamento e teste) foram normalizadas por um fator, onde este fator representa a distância entre os lados opostos do polígono usado. A finalidade da utilização deste fator é colocar todos os valores de entrada da rede dentro do intervalo compreendido entre 0 e 1. A Fig.10 apresenta como este fator é extraído para o caso do quadrado.

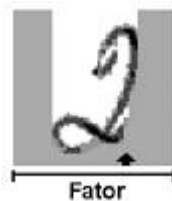


Fig.10: Representação do fator de normalização

Um outro procedimento adotado para a preparação do conjunto de dados foi à eliminação de todas as linhas com variância zero. Este procedimento bem simples elimina algumas variáveis do espaço de entrada, e desta forma, reduz a sua dimensionalidade. Por exemplo: a entrada de tamanho 128 para a rede especializada em letras foi reduzida para 109; e na rede especializada em números foi reduzida de 128 para 99.

Conforme descrito anteriormente, o experimento foi desenvolvido utilizando-se o *toolbox* de redes neurais do MATLAB versão 5.2 e o treinamento da rede foi realizado usando a opção “*traingdx*”, com momento e taxa de aprendizado adaptável.

- **Resultados**

No conjunto de testes obteve-se os seguintes resultados:

Tabela 4. Resultado das redes

Atributos	Dígitos	Letras
Nº epoch para treinamento	6000	10000
Nº epochs realizados	6000	10000
% Classificação Correta	79.75	72.26
% Classificações Incorretas	20.25	27.74
% Rejeições	0.00	0.00

## 6 Conclusões

Os resultados obtidos se mostraram bastante abaixo do desejado, especialmente o da rede responsável pelo reconhecimento de letras, porém há alguns fatores determinantes para este resultado final. O mais importante foi a baixa qualidade geral das fotos à disposição. A baixa resolução, aliada a dificuldades como iluminação, ruído gerado pelo ambiente (chuva, neblina, etc.) e problemas com as câmeras em si (falta de foco, localização inadequada, ângulo de visão impróprio, etc.), causaram problemas que com certeza, foi a fonte das dificuldades enfrentadas por toda a execução do trabalho.

No módulo de segmentação o resultado foi de 76.61% de acerto, considerando os sete caracteres de cada placa. Desempenho este, mais uma vez, longe do esperado, reflexo da baixa qualidade das fotos originais. No módulo de reconhecimento, foram detectadas algumas confusões mais frequentes que ocorreram entre “B”, “D” e “O” para a rede de letras e entre “0” e “8” na rede de dígitos. Isso acontece porque o exterior desses caracteres, seguindo a fonte utilizada na fabricação das placas, tem um desenho muito semelhante.

Tendo em vista os baixos resultados obtidos, principalmente, em casos como a distinção entre “0” e “8”, tentou-se diversas soluções alternativas. Uma delas envolveu o uso de uma rede especializada na diferenciação entre “0” e “8”. No entanto, mesmo tal tentativa foi mal-sucedida, tendo em vista a grande influência da má qualidade dos dados. Esta rede teve como melhor resultado apenas 57% de acerto. Para uma rede de apenas 2 saídas, isso equivale a dizer que o treinamento não surtiu efeito nenhum, pois a rede estava literalmente dando respostas aleatórias.

Em resumo, pode-se concluir que para uma melhoria no reconhecimento, faz-se necessário uma melhor qualidade de imagens, que devem ser obtidas por um equipamento que tenha uma taxa de resolução mais alta. O investimento para a aquisição de tal equipamento não é tão alto, levando-se em consideração, no caso de aplicação de multas, a quantidade de veículos que não podem ser notificados, devido à má qualidade da imagem. Lembrando que 3.846 fotos

não puderam ser aproveitadas. Para o prosseguimento deste trabalho, estão sendo obtidas imagens com maior resolução e ainda, está sendo desenvolvida uma nova técnica de extração de características, cuja finalidade é reduzir as confusões que ocorreram no reconhecimento das letras “B”, “D” e “O” e os algarismos “0” e “8”.

### **... – Referências Bibliográficas**

- [1] Rodrigues, R. J.; Thomé, A. C. G.; Cursive character recognition - a character segmentation method using projection profile-based technique, The 4th World Multiconference on Systemics, Cybernetics and Informatics SCI 2000 and The 6th International Conference on Information Systems, Analysis and Synthesis ISAS 2000 - Orlando, USA - August 2000.
- [2] Bishop, C. M.; Neural Networks for Pattern Recognition, Oxford University Press, 1995.
- [3] Duda, R. O; Hart, P. E. Pattern Classification and Scene Analysis, John Wiley & Sons, 1973.
- [4] Kupac, G. V.; Rodrigues, R. J.; Thomé, A. C. G.; Extração de características para reconhecimento de dígitos cursivos, VIth Brazilian Symposium on Neural Networks - SBRN2000, Rio de Janeiro, november 2000.
- [5] Rodrigues, R. J.; Kupac, G. V.; Thomé, A. C. G.; Character Feature Extraction Using Polygonal Projection Sweep (Contour Detection), accepted for presentation at IWANN2001. Granada - Spain, June 2001.
- [6] Rodrigues, R. J.; Silva, E.; Thomé, A. C. G.; Feature Extraction Using Contour Projection, accepted for presentation at SCI2001. Orlando - USA, July 2001.
- [7] Demuth, H.; Beale, M.; Neural Network Toolbox User's Guide Version 3.0, MathWorks, Inc, 1997.
- [8] Haykin, S., Neural Networks: A Comprehensive Foundation, Prentice Hall, 1999.
- [9] Justiça, Conselho Nacional de Trânsito-CONTRAN, <http://www.mj.gov.br/contran/>, 1997.