

SISTEMA NEURAL DE DETECÇÃO DE PLACAS DE VEÍCULOS AUTOMOTORES EM IMAGENS DIGITALIZADAS

LUÍS VICTOR CASCAO*, JOSÉ M. DE SEIXAS*

**Laboratório de Processamento de Sinais
COPPE/Poli - Universidade Federal do Rio de Janeiro
Rio de Janeiro, RJ, Brasil*

Emails: victor.cascao@poli.ufrj.br, seixas@lps.ufrj.br

Abstract— Computer vision techniques may be helpful in systems aiming at identifying vehicles. This paper focuses on the the automatic localization of the Brazilian license plate in scanned images, which is required for a classification task. This is achieved finding the region of interest (ROI) based on the spectral analysis of lines of the image and neural networks for identifying which ROI actually has the plate in it. Results show an identification efficiency above 86 % for images acquired in a practical situation and without light control.

Keywords— Plates Detection System, Patterns Recognition, Principal Components Analysis, Neural Networks.

Resumo— Técnicas de visão computacional podem ser úteis em sistemas com o objetivo de identificar veículos. Este artigo enfoca a localização automática da licença veicular brasileira em imagens digitalizadas, o que é necessário para uma tarefa de classificação. Isto é conseguido encontrando a região de interesse (ROI) baseado na análise espectral da imagem de linhas e redes neurais para identificar se a ROI realmente contém a placa. Os resultados mostram uma eficiência de identificação acima 86 % para imagens adquiridas em uma situação prática e sem controle de iluminação.

Palavras-chave— Sistema de Detecção de Placas, Reconhecimento de Padrões, Análise de Componentes Principais, Redes Neurais.

1 Introdução

Com o aumento da quantidade de veículos nas cidades, é cada vez mais imperativo desenvolver novas formas de identificação automática das placas dos veículos automotores particulares para fins de controle de fluxo, cobrança de pedágios, segurança, entre outros fatores.

Atualmente, a tecnologia de reconhecimento da licença veicular através de imagens digitalizadas já está bem difundida. Contudo, o crescimento da demanda de fiscalização e segurança eletrônica faz com que novos produtos e tecnologias surjam para suprir a demanda de fiscalização e controle eletrônico. Os sistemas de reconhecimento de licença veicular são compostos de 4 etapas: reconhecimento da imagem da placa numa imagem, recorte da placa da imagem, recorte dos caracteres contidos na imagem da placa e identificação da licença veicular.

Para o mesmo banco de dados usado neste trabalho, os trabalhos recentes como (Gesuladi, A.R., 2001), sobre segmentação e classificação, através de redes neurais, dos caracteres contidos na placa e (Leite, B.B., 2005), baseado na técnica de filtragem ótima, obtendo uma eficiência de 81% da placa na imagem, foram referência o trabalho a ser detalhado a seguir.

O algoritmo desenvolvido neste trabalho busca detectar placas de veículos particulares especificada

pelo artigo 1º da resolução 045/98 do CONTRAN¹, que determina as dimensões de 130mm de altura, 400mm de comprimento e cor cinza. Para isso, foi utilizado um banco de dados que consiste de 1250 imagens de automóveis obtidas no pedágio da ponte Rio - Niterói durante quatro dias e sob condições não controladas de luminosidade. Técnicas de pré-processamento foram utilizadas com o objetivo de melhorar a qualidade das imagens, visto que as mesmas foram obtidas em condições reais e sem estruturação do ambiente.

2 Pré-Processamento

O pré-processamento se faz necessário para facilitar a localização da placa na imagem, pois o ruído é uma das principais fontes de erro de classificação da rede neural. Com a fase de pré-processamento, evita-se a possível detecção de áreas da imagem que não possuam placa, na verdade.

Diferentes formas de pré-processamento da imagem foram utilizadas, conforme detalhado a seguir.

2.1 Filtro Passa Baixa

As imagens utilizadas neste trabalho apresentam informação de ruído de diferentes naturezas. A filtragem Passa-Baixa surge como uma possível so-

¹Conselho Nacional de Trânsito

lução para a eliminação do ruído, que, nesse caso, é toda e qualquer informação que não seja a placa alvo. Como, por exemplo, o ruído introduzido pela câmera na aquisição das imagens em forma de rajadas, elementos dos automóveis em formato parecido com uma placa, como faróis e pára-choques e elementos com muitas bordas, como grades frontais. A opção foi por Filtro Passa-Baixa Gaussiano (Gonzalez, 1993), de forma que se retirasse o ruído em rajada e suavizasse a imagem, rejeitando os elementos de alta frequência como pode ser visto na figura 1



Figura 1: Imagem sem pré-processamento (superior) e após filtragem passa-baixa (inferior).

2.2 Filtro de Mediana

A utilização do filtro da mediana se dá em casos que a imagem apresenta ruídos impulsivos (pixels brancos e pretos aleatórios) ou quando se quer fazer uma filtragem passa-baixa preservando os contornos e bordas da imagem (Phillips D., 1994). Para esse trabalho, foi escolhido um filtro de mediana com tamanho 5x5 devido a sua habilidade em filtrar os ruídos da imagem e de conservar as informações de contornos da placa.

2.3 Áreas de Interesse

O pré-processamento descrito em (Leite, B.B., 2005) consiste, basicamente, em filtrar os componentes de frequência indesejados, formando uma nova imagem que mostre apenas as linhas que estão numa faixa de frequência característica horizontal da placa, entre 0,03 e 0,06, como visto na figura 2.

A eficiência do pré-processamento consiste em localizar regiões que contenham completamente a placa. Se existir uma região, dentro daquelas retornadas pelo algoritmo, que contenha por completo a placa, considera-se um acerto, caso contrário, é

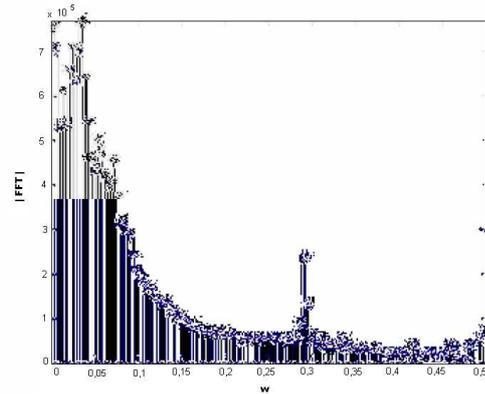


Figura 2: Gráfico acumulativo para validação da faixa de frequência característica.



Figura 3: Imagem original I (esquerda) e a imagem F resultante (direita), formada pelo módulo da resposta em frequência de cada linha.

considerado um erro. Baseado nesse critério, segundo (Leite, B.B., 2005), existem duas configurações do conjunto de parâmetros ² que se destacam, conforme a tabela 1 ³

Tabela 1: Eficiências de localização do pré-processamento

ll	ul	T	num Resp	dist Resp	tam Recorte	efi- ciência
0,03	0,06	18	3	35	60	0,92
0,03	0,06	18	4	35	60	0,928

O conjunto de parâmetros de entrada, que gera 92,8% de eficiência, tem um número de respostas na etapa de pré-processamento igual a 4, o que faz com que a área de atuação da rede seja maior, aumentando a probabilidade de erro da classificação pela rede neural. Para o conjunto que apresenta a eficiência de 92,0%, a rede neural atuará numa área menor, uma vez que, para essa eficiência de pré-processamento, o número de respostas é igual a 3.

²Os parâmetros de configuração do pré-processamento de áreas de interesse são: Limite inferior do filtro com frequência normalizada (ll); Limite superior do filtro com frequência normalizada (ul); Tamanho da janela deslizante em pixels (T); Número de respostas de áreas da imagem (numResp); Distância entre as respostas em pixels (distResp); Número de linhas recortada da imagem em pixels (tamRecorte).

³Os resultados da tabela 1 apresentam incerteza de 0,008.

3 Projeto da Redes Neurais

Nesta seção, será abordada a construção da rede neural MLP identificadora, baseada nas análises estatísticas dos padrões das placas existentes nas imagens dos veículos automotores. Nessa etapa, a base de dados foi dividida em dois conjuntos, o conjunto de desenvolvimento, com 1000 imagens, e o conjunto de teste, com as 250 imagens restantes. Antes de passar para o projeto da rede neural em si, é necessário fazer uma análise dos dados disponíveis, buscando estimar as flutuações dos parâmetros relevantes ao projeto da rede.

3.1 Análise Estatística

Com as informações de posicionamento da base de dados, foi estimada a flutuação nas dimensões das placas. A partir das medidas de cada placa, obteve-se a distribuição da largura e da altura das mesmas, calculando-se a média e o desvio padrão.

Constatou-se que a largura da placa apresenta uma distribuição bi-modal, com o limiar de separação em 237 *pixels* e com valor médio da largura da placa de 247,14 com um desvio padrão de 11,58 *pixels*. Já a distribuição da altura das placas apresentou uma distribuição, com valor médio de 42,53 e com desvio padrão de 2,62 *pixels*.

A partir das informações obtidas, foi determinado que o tamanho da área a ser pesquisada e submetida à rede neural será do tamanho da *placa média* com 248 *pixels* de largura por 44 *pixels* de altura. Sabendo o tamanho da área, chegamos ao montante de 10912 *pixels* a serem analisados, tornando o processamento proibitivo para uma rede neural. Portanto, foi feita a redução da dimensão do espaço original de dados através da Análise de Componentes Principais.

3.2 Análise de Componentes Principais (PCA)

A representação dos dados a ser utilizada consiste em dividir a imagem em blocos de $m \times m$ *pixels*, e considerar cada um desses blocos como uma amostra de um vetor aleatório x (Jolliffe, 1990). Por exemplo, no caso em que $m = 8$, então $x \in R^{64}$.

Para saber o quanto de informação pode-se retirar da imagem sem perdas relevantes de informação da placa, foi feito um gráfico acumulativo da influência de cada componente na energia total da placa. A figura 4 mostra a curva de carga, para sub-divisões em 4×4 *pixels*, 8×8 e 16×16

Por inspeção dos gráficos da figura 4, chega-se a conclusão que, com poucos componentes, pode-

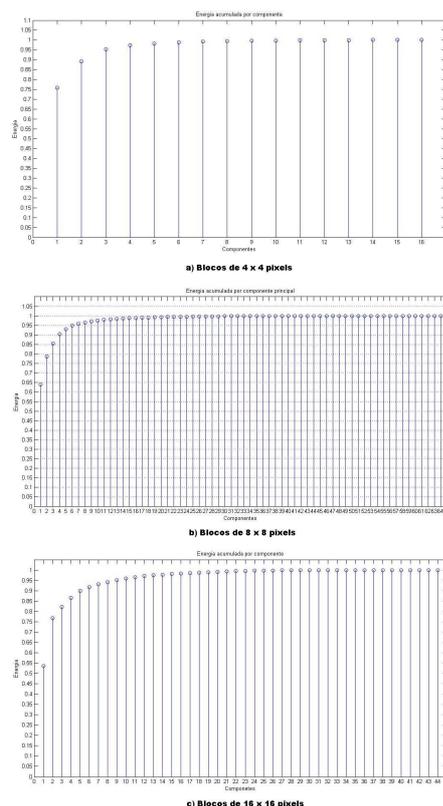


Figura 4: Energia acumulada para imagens das placas. a) subdivisão de 4×4 *pixels*, b) com 8×8 *pixels* e c) 16×16 *pixels*.

se representar com fidelidade as características da placa. Portanto, para esse trabalho, o critério de escolha da quantidade de componentes a serem retidos foi o limiar de reconstrução da energia ao nível de 95%, sendo que, e a escolha do número de *pixels* da subdivisão foi dada considerando-se a taxa de compressão obtida e o tempo de processamento⁴.

Logo a imagem deve ser subdividida em blocos de 8×8 pois com 8 componentes atende-se o limiar imposto e se obtém uma taxa de compressão na razão de 1:8 com o tempo de processamento de 50ms em detrimento a baixa taxa de compressão(1:4) obtida nos blocos de 4×4 e ao grande tempo de processamento (200ms) quando utilizado blocos de 16×16 *pixels*. Portanto após aplicar-se o PCA, a placa será representada por 155 vetores $\in R^8$. Reduzindo assim a dimensão de entrada da rede neural de 10912 entradas para 1240, tornando agora o processamento dessas entradas possível.

3.3 Dados de treinamento

Para cada uma das mil imagens do conjunto de treinamento, foi extraída a placa e seis outros pedaços

⁴Tempos medidos pelo Matlab. Testes realizados num computador pessoal com processador AMD Turion 2GHz e 1GB de memória RAM

da imagem sorteadas com o tamanho da *placa média* (44×448 pixels), que contenham até 50% da placa na imagem. Sendo as projeções nas 8 principais componentes usadas como a entrada da rede neural.

O equilíbrio de populações (Haykin, 1999) se faz necessário e também foi avaliado do ponto de vista do erro de classificação da rede treinada. As proporções testadas foram de 1:6, 1:1 e 3:2. Nestas duas últimas, a população de placas foi replicada cinco vezes e três vezes respectivamente. Esta última proporção apresentou o menor erro de classificação. Devido a proporção escolhida de 3:2, obtemos para o conjunto de treinamento 10.000 vetores (imagens já pré-processadas) que possuem a dimensão de 1240.

3.4 Topologia da Rede Neural

Os testes para definição do número de neurônios na camada escondida foram feitos utilizando os pré-processamentos escolhidos e aprendizagem por gradiente descendente⁵ e gradiente descendente com aprendizado adaptativo⁶ (Bishop, 2007). O treinamento foi repetido cinco vezes para cada configuração e, a partir dos erros de classificação obtidos no teste destas redes treinadas, o erro médio para cada rede foi estimado. Este erro foi utilizado como critério para estipular a topologia da rede neural e o tipo de aprendizado utilizado no treinamento como pode ser verificado nas tabelas 2 e 3.

Tabela 2: Erro de classificação por tipo de treinamento usando imagens sem pré-processamento

Treinamento	Quantidade de Neuronios			
	1	2	3	4
traingda	4,72%	xxx	xxx	3,54%
trainscg	6,00%	5,24%	3,92%	2,90%
Treinamento	5	6	7	
traingda	xxx	2,50%	xxx	
trainscg	2,80%	2,80%	2,76%	

Tabela 3: Erro de classificação por tipo de treinamento usando imagens com filtragem mediana

Treinamento	Quantidade de Neuronios			
	1	2	3	4
traingda	3,70%	xxx	xxx	3,20%
trainscg	4,83%	4,61%	4,20%	2,28%
Treinamento	5	6	7	
traingda	xxx	3,04%	xxx	
trainscg	2,00%	1,68%	1,70%	

Como a incerteza das tabelas 2 e 3 é de 0,04% os resultados apresentados para ambas as topologias são equivalentes. Por parcimônia, foi escolhida a rede com seis neurônios.

⁵Representado pela sigla trainscg

⁶Representado pela sigla traingda

4 Algoritmo de Localização da Placa

A busca pela presença da placa é feita através de uma janela de varredura deslizante, que possui a dimensão do tamanho da *placa média*, e, para cada passo de varredura, é executada a análise dos componentes principais e as projeções da imagem nos componentes são propagados pela rede neural. O passo de varredura da janela pode ser escolhido de forma que um passo menor terá uma maior probabilidade de se localizar a placa.

Na figura 5.a, está ilustrada a janela deslizante numa amostra do banco de imagens. Na figura 5.b, está ilustrado o respectivo vetor de saída, no qual, podemos perceber vários picos, sendo que os mais altos apresentam grande probabilidade de corresponder a uma uma região que contém placa.

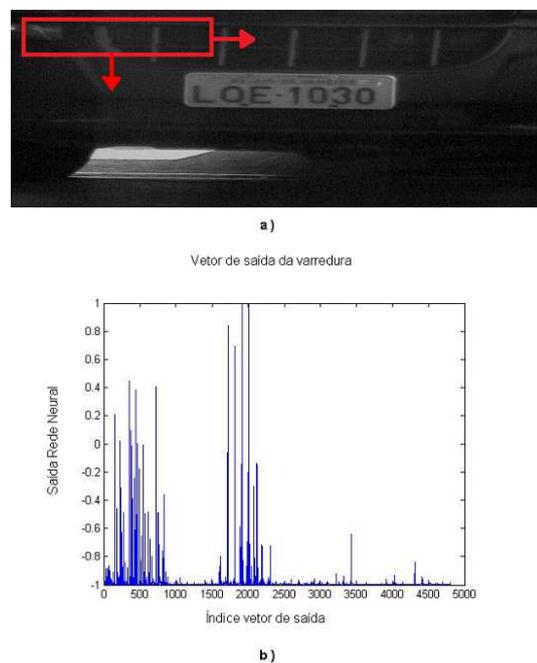


Figura 5: Exemplo vetor de saída da varredura.

5 Resultados

5.1 Medição da Eficiência

Neste trabalho foi utilizado, como referência para a análise de acertos, a informação de posicionamento (x, y) do primeiro *pixel* da placa. Considera-se um acerto quando o ponto inicial do resultado obtido pela rede neural apresentar distância euclidiana menor do que de 15 *pixels* com respeito ao *pixel* inicial da placa;

5.2 Eficiência de detecção

A partir da tabela 4, pode-se verificar a influência de cada parâmetro do algoritmo de localização da placa sobre a eficiência de detecção ⁷. Primeiramente, com o aumento do passo de varredura, perde-se na detecção, já que, para um passo maior que 1 *pixel*, as chances de conter perfeitamente a placa inteira diminuem, entretanto o custo computacional diminui, uma vez que haverá menos áreas a serem investigadas. Sobre o pré-processamento utilizado na imagem, vemos que a escolha de áreas de maior probabilidade de ocorrência da placa influencia em muito a diminuição do custo computacional, pois se reduz a área a ser varrida pela janela deslizante; porém, seu uso limita a eficiência do sistema de detecção a 92%

Tabela 4: Resultados de detecção obtidos.

Passo X [pixels]	Passo Y [pixels]	Pré-Processamento	Eficiência de detecção	Distância média [pixels]
1	1	sem pré-processamento	94,8% s	3,47
1	1	mediana	92,0%	3,78
1	1	áreas de interesse	88,0%	3,79
4	4	sem pré-processamento	90,0%	4,76
4	4	mediana	88,0%	4,95
4	4	áreas de interesse	86,0%	4,8
8	4	sem pré-processamento	75,2%	8,38
8	8	mediana	71,6%	8,23
8	8	sem pré-processamento	63,2%	8,36
8	8	áreas de interesse	64,8%	8,03

6 Conclusão

De acordo com os resultados apresentados, o método proposto para o reconhecimento de placas apresentou um resultado satisfatório, levando-se em conta que as condições de luminosidade das imagens não foi controlada. Uma rede neural MLP foi capaz de identificar placas em imagens de automóveis particulares com eficiência de 94,8% e com resolução de 3,47 *pixels*.

Há boas perspectivas para trabalhos futuros. Os seguintes problemas podem ser atacados, de modo a melhorar o desempenho do sistema. Alterar o

pré-processamento, utilizando uma filtragem mais adequada. Melhorar, via pós-processamento, a eficiência de detecção estimando melhor a região de ocorrência da placa, a partir do vetor de saída da varredura. Pode-se acelerar o processo de varredura, através de uma busca mais rápida com saltos de 8 pixels na vertical e 8 na horizontal por exemplo, seguida de uma busca fina nas regiões com mais chance de ter uma placa conforme a saída da rede neural. Um modelo neural para a identificação da licença a partir do recorte e reconhecimento dos caracteres da placa já detectada, tem sido desenvolvido (Gesuladi, 2002), o que completa assim a solução do problema de reconhecimento da licença veicular do automóvel.

7 Agradecimentos

Os autores agradecem à FAPERJ e CNPq pelo apoio dado a este projeto e aos colegas do CBPF (Centro Brasileiro de Pesquisas Físicas), pelo compartilhamento do banco de dados e pelas discussões frutíferas no desenvolvimento deste trabalho.

Referências

- Bishop, M. (2007). *Pattern Recognition and Machine Learning*, 1 edn, Springer.
- Gesuladi, A. da R. ; Seixas, J. M. d. . A. M. P. . A. M. P. . (2002). Character recognition in car license plates based on principal component analysis and neural processing., *Brazilian Symposium on Neural Networks*, Recife, pp. 01-06.
- Gesuladi, A.R. (2001). Sistema Neuronal de Reconhecimento de Placas em Veículos Automotores Particulares, *Disertação de m.sc.*, COPPE/UFRJ e CBPF.
- Gonzalez, R.C., W. R. (1993). *Digital Image Processing*, Addison-Wesley Publishing Company.
- Haykin, S. (1999). *Neural Networks. A Comprehensive Foundation*, 2 edn, Prentice-Hall, Inc.
- Jolliffe, I. (1990). *Principal Component Analysis: A beginner's guide - introduction and application*, Weather.
- Leite, B.B. (2005). Detecção de imagens digitalizadas de placas de automóveis por meio de filtragem casada, *Disertação de m.sc.*, COPPE/UFRJ.
- Phillips D. (1994). *Image Processing in C, R D Publications*.

⁷Os resultados da tabela apresentam incerteza de 0,4%