

RECONHECIMENTO ÓTICO DE CARACTERES PARA RECONHECIMENTO DAS SINALIZAÇÕES VERTICAIS DAS VIAS DE TRÂNSITO

SAMUEL LUZ GOMES¹; ELIZÂNGELA DE SOUZA REBOUÇAS¹;
PEDRO PEDROSA REBOUÇAS FILHO¹;

1 – INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO CEARÁ (IFCE),
LABORATÓRIO DE PROTÓTIPOS (LPROT), GRUPO DE PESQUISA DE SISTEMAS
INTELIGENTES (GPSI)
pedrosarf@ifce.edu.br

Resumo - O presente artigo apresenta o desenvolvimento de um aplicativo para o sistema ANDROID utilizando um sistema de reconhecimento óptico de caracteres para detectar e reconhecer a sinalização vertical de regulamentação de trânsito, especificamente para o sinal R-19¹ (velocidade máxima permitida) nos padrões brasileiros. Para o desenvolvimento deste aplicativo foram utilizadas técnicas de Processamento Digital de Imagens (PDI) e Reconhecimento Padrões (RP) através de Rede Neural Artificial (RNA). Os resultados finais mostram que o sistema apresenta-se como promissor para identificação e reconhecimento de placas de sinalização, obtendo êxito em todos os vídeos testados, detectando e reconhecendo a placa de sinalização de trânsito de velocidade adequadamente.

Palavras-chave: Processamento Digital de Imagem. Visão Computacional. Aplicações Automotivas.

I. INTRODUÇÃO

De acordo com o DETRAN², em 1904, existiam aproximadamente 55.000 veículos no mundo, hoje somente no estado do Ceará temos uma frota com cerca de dois milhões de veículos, já em São Paulo existem mais de vinte e quatro milhões de veículos. Com o aumento da frota de veículos no país há, consequentemente, um aumento no número de acidentes de trânsito e de problemas com engarrafamentos. Nesse cenário, os sistemas de visão computacional podem auxiliar os condutores dos veículos servindo como mais um item de segurança, como também no monitoramento de tráfego.

Em virtude do constante crescimento da frota de veículos, principalmente nos grandes centros urbanos, surgem problemas cada vez mais difíceis de serem solucionados, exigindo, assim, um maior investimento em engenharia de tráfego para tornar o trânsito mais rápido, seguro e eficiente. Um dos problemas mais graves é a falta de atenção dos condutores à sinalização das vias de trânsito, que ocasiona grande parte dos acidentes.

Os Sistemas de Visão Computacional, também chamados de Visão Artificial, têm sido utilizados em várias aplicações, tais como: detecção, reconhecimento e classificação de objetos; estimativa de parâmetros como tamanho, posição, orientação e velocidade, entre outros. Uma das aplicações mais clássicas da Visão Artificial é o OCR - Optical Character Recognition ou Reconhecimento Óptico de Caracteres. Tal aplicação consiste no reconhecimento de letras, algarismos, enfim, símbolos escritos (ALEXANDRIA, 2005).

Pensando nisto, este trabalho propõe a utilização de um Sistema de Visão computacional utilizando Processamento Digital de Imagem (PDI) e Inteligência Computacional para desenvolver um OCR com o foco em sistema de detecção e reconhecimento de placas de sinalização vertical de regulamentação de trânsito, especificamente para o sinal R-19 – velocidade máxima permitida – nos padrões brasileiros.

II. REVISÃO BIBLIOGRÁFICA

Um sistema de Visão Computacional típico possui 6 etapas: Aquisição de imagens, pré-processamento, segmentação, extração dos atributos, identificação dos padrões previamente estabelecidos e, por fim, é gerado um relatório com as análises e informações obtidas (GONZALEZ and WOODS, 2010).

O pré-processamento se caracteriza pelo uso de funções clássicas de Processamento Digital de Imagens (PDI), onde o principal objetivo é tornar a imagem mais adequada para uma aplicação específica (ALEXANDRIA, 2005). O pré-processamento normalmente envolve rotinas de restauração, realce ou somente a representação apropriada dos dados de entrada. Neste trabalho, o pré-processamento é realizado com filtros de suavização.

¹ R-19: Regulamenta o limite máximo de velocidade em que o veículo pode circular na pista ou faixa, válido a partir do ponto onde o sinal é colocado.

² Disponível em: <http://portal.detran.ce.gov.br/index.php/curiosidades>

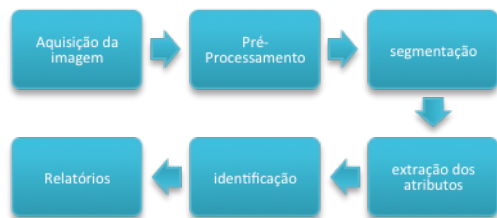


Figura 1- Etapas do SVC desenvolvido para detecção e reconhecimento de sinalização vertical regulamentadora de velocidade

Os Filtros de suavização são utilizados para redução de ruídos, como, remoção de pequenos detalhes da imagem antes da extração de objetos e conexão de pequenas descontinuidades em linhas ou curvas.

O filtro de mediana é um filtro não linear de suavização, o qual, como o nome sugere, substitui os valores do pixel pela mediana dos valores de intensidade na vizinhança desse pixel, incluindo o valor original do pixel. O filtro de mediana é particularmente eficaz na presença de ruído impulsivo, também conhecido como ruído sal e pimenta (GONZALEZ and WOODS, 2010).

Quanto a etapa de segmentação, apresentada na Figura 1, é o processo de separar as várias regiões que compõem uma imagem, de acordo com suas propriedades. A técnica de segmentação a ser utilizada varia com a aplicação do sistema. O processo de segmentação é de fundamental importância em qualquer sistema de Visão Artificial, de tal forma que o desempenho do sistema é essencialmente dependente do desempenho do processo de segmentação. Uma segmentação realizada de forma adequada em um sistema de visão computacional é um passo substancialmente importante para seu correto funcionamento (ALEXANDRIA, 2005). Neste trabalho, as técnicas utilizadas neste trabalho são a detecção de bordas de Canny e Transformada de Hough.

A abordagem do detector de bordas de Canny baseia-se em três objetivos básicos:

- I. Baixa taxa de erro: Todas as bordas devem ser encontradas e devem ser o mais próximo possível das bordas verdadeiras.
- II. Os pontos de bordas devem estar bem localizados: As bordas detectadas devem estar mais próximas possíveis das bordas verdadeiras.
- III. Resposta de um único ponto de borda: O detector não deve detectar múltiplos pixels de borda onde apenas um único ponto de borda existe.

Em geral é difícil (ou impossível) encontrar uma solução fechada que satisfaça todos os objetivos anteriores. Canny (1986) chegou à conclusão de que uma boa aproximação para um detector ótimo de bordas de degrau é a primeira derivada de uma gaussiana, que em uma dimensão pode ser dada por:

$$d = \left(-\frac{x}{\sigma^2} \right) e^{-\frac{x^2}{2\sigma^2}}$$

Onde σ^2 consiste na variância dos dados e x o dado de entrada.

A transformada de Hough é uma técnica de extração de características utilizadas na análise de imagens, visão computacional e processamento de imagem digital. O objetivo da técnica é encontrar pontos imperfeitos do objeto

comparando com classe de formas pelo processo de votação. Este procedimento de votação é levado em um espaço de parâmetros, a partir do qual os candidatos de objetos são obtidos como máximos locais num espaço acumulador assim chamado, que está explicitamente construído pelo algoritmo para o cálculo da transformada de Hough.

A clássica transformada de Hough estava preocupada com a identificação de linhas na imagem, mas depois a transformada de Hough foi estendida para identificar posições de formas arbitrárias, mais comumente círculos ou elipses. Duda e Hart (1972) utilizaram coordenadas polares para representação de uma linha. Sugeriram que linhas poderiam ser completamente parametrizadas usando o comprimento, r , e a orientação, q , do vetor normal para a linha da imagem original. Usando esta parametrização, todo o ponto (x, y) na linha satisfará a equação $r = x \cdot \cos(q) + y \cdot \sin(q)$.

A etapa de identificação, apresentada na Figura 1, neste projeto é realizada utilizando uma Rede Neural Artificial do tipo Multilayer Perceptron (MLP).

A MLP é uma rede supervisionada, ou seja, os pesos de cada camada são reajustados de acordo com o valor da saída da rede comparada com o valor real da saída (ARBIB, 2003). A MLP é subdividida em entrada do sistema, pesos da camada oculta, saída da camada oculta que consiste na entrada da camada de saída, pesos da camada de saída e saída do sistema, podendo ser mais de uma camada oculta (RUCK *et al.*, 1990; HAYKIN, 2008).

Nesse sentido, a fase de treinamento da MLP pode ser dividida em duas etapas, uma chamada de sentido direto na qual o trânsito de informações da entrada até a saída, gerando o valor de saída da rede. Enquanto isso na segunda etapa, chamada de sentido inverso, é utilizado o valor do erro gerado no sentido direto, e este é retro propagado para as camadas intermediárias para ajustar os pesos destas (BISHOP, 2006; WIDROW and WINTER, 1988).

III. METODOLOGIA

Nesta seção estão listados os métodos SVC utilizados para a detecção e o reconhecimento da sinalização de trânsito vertical de regulamentação, especificamente para o sinal R-19, “velocidade máxima permitida”. A Figura 2 mostra o fluxograma do algoritmo utilizado para a resolução do problema proposto, enquanto a Figura 3 demonstra os resultados de cada etapa descrita pelo fluxograma.

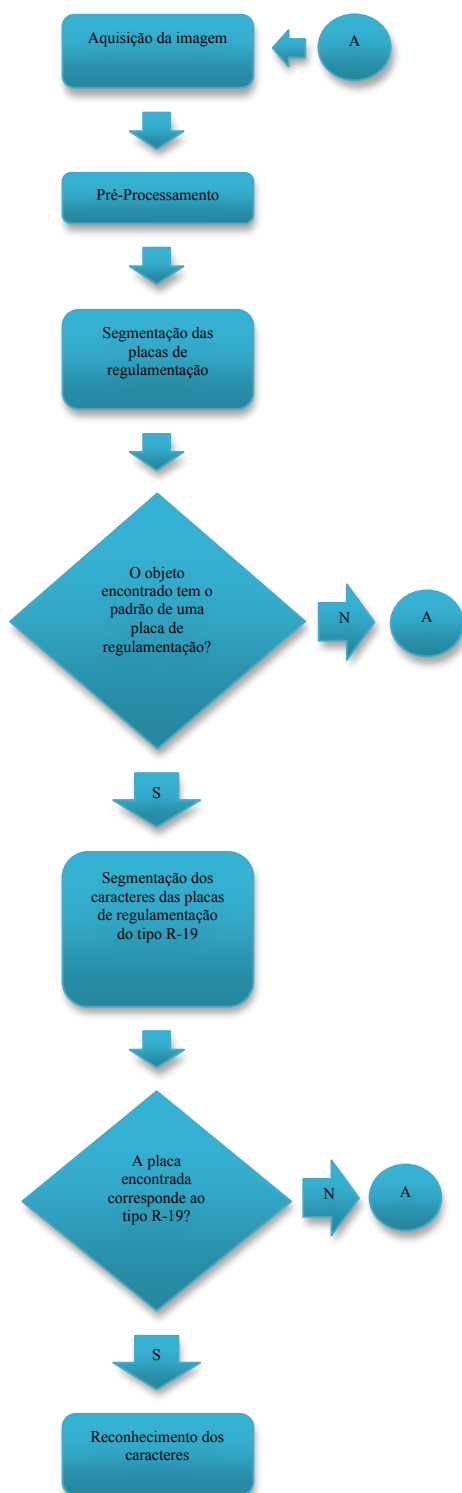


Figura 2 - Fluxograma do algoritmo de segmentação dos caracteres das placas de regulamentação R-19

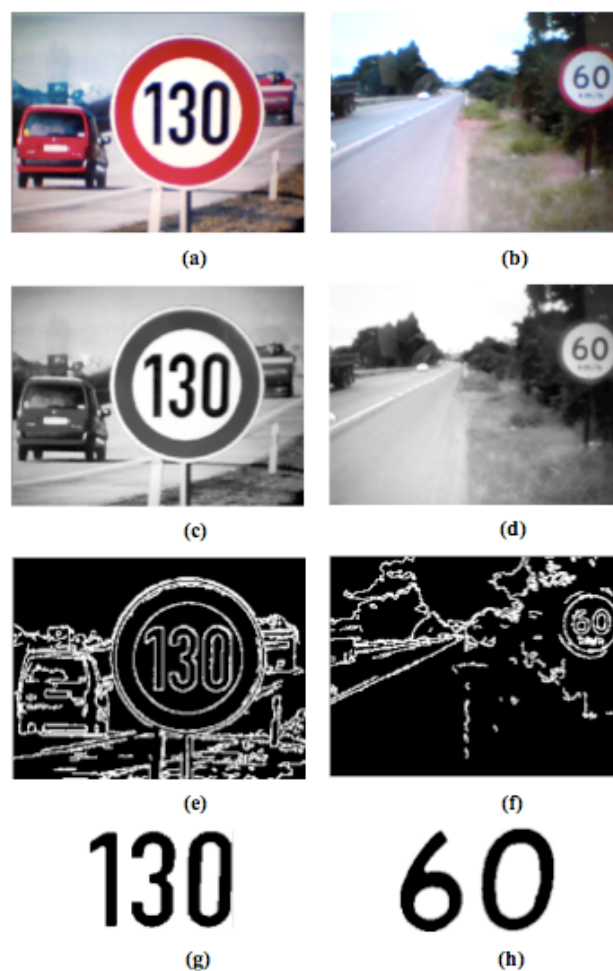


Figura 3 - Etapas do algoritmo de segmentação dos caracteres
a) e b) imagens originais; c) e d) imagens pré-processadas; e) e f) etapa de segmentação das placas; g) e h) etapa de segmentação dos caracteres.

3.1 Aquisição das imagens

As imagens foram adquiridas por uma câmera de um smartphone com a resolução de 5 Mpx, onde foi gerado um banco de dados de várias placas, em diferentes situações de iluminação, de distância e de angulação para o melhor treinamento do sistema de reconhecimento.

3.2 Pré-processamento

Nesta etapa foram utilizadas técnicas para retirar os ruídos e outras imperfeições da imagem tornando, assim, as etapas seguintes mais eficientes.

A imagem de entrada, que estava inicialmente colorida, foi convertida para tons de cinza, com a finalidade de diminuir o custo computacional nos processos seguintes, posteriormente, aplicamos um filtro passa-baixa (Mediana) para reduzir os ruídos, equalizamos o histograma da imagem para melhorar o contraste.

3.3 Segmentação

A segmentação foi dividida em duas etapas, onde a primeira etapa detecta e confirma a presença de uma placa de regulamentação na imagem e a segunda etapa detecta os caracteres presentes na placa e confirma que a placa encontrada é de fato do tipo R-19.

Foi utilizado o filtro operador Canny com a finalidade de detectar os contornos dos objetos na imagem. Posteriormente foi aplicada uma técnica de morfologia matemática para corrigir pequenas falhas geradas no pré-processamento e melhorar os resultados obtidos pelo Canny. Após a estas operações foi realizada uma análise dos objetos encontrados para a identificação das placas baseada em uma relação entre a largura, altura e área dos mesmos, detectando, assim, os contornos com o padrão circular. Os objetos validados como placas foram armazenados e processados posteriormente e os demais foram descartados.

Os objetos armazenados na etapa anterior foram analisados com o objetivo de identificar os caracteres contidos neles. Sendo encontrados caracteres de acordo com o padrão estabelecido, a placa foi validada como do tipo R-19 e os seus caracteres foram preparados para a etapa de reconhecimento.

Para adequar os caracteres aos padrões do sistema de reconhecimento utilizado no projeto os mesmos foram redimensionados para o tamanho 35x35 pixels com o fundo preto e o caractere branco e centralizado em relação à largura do mesmo, este redimensionamento têm por objetivo tornar o algoritmo invariante à distância.

3.4 Reconhecimento dos caracteres

Para o reconhecimento dos caracteres foi utilizado uma Rede Neural Artificial (RNA) do tipo perceptrons de múltiplas camadas (MLP, *multilayer perceptron*) com o treinamento supervisionado conhecido como retropropagação de erro (*error back-propagation*).

IV. RESULTADOS

Para efetuar os teste foi gerado um banco de dados contendo vídeos 21 vídeos de placas de regulamentação de velocidade máxima (20, 30, 40, 60 e 80Km/h). As imagens foram coletadas nas ruas e avenidas da cidade de Fortaleza. A partir dos vídeos foram extraídas 12520 imagens contendo placas em diferentes inclinações e distâncias.

Das 12520 imagens de placas existentes no banco de dados, 5677 foram localizadas pela etapa de detecção do software, obtendo assim, um total de 45,3% na taxa de detecção das placas.

Das 5677 placas detectadas pelo software, 5483 placas foram reconhecidas corretamente na etapa posterior, onde é efetuado o reconhecimento dos caracteres das placas detectadas, esta etapa apresentou uma taxa de acerto de aproximadamente 96,6%.

Analisando o sistema de uma forma completa obtivemos uma taxa de detecção e reconhecimento de aproximadamente 43,8%, reconhecendo corretamente 5474 placas de um banco de dados com 12520.

A tabela 1 mostra a taxa de detecção e acerto referente a cada valor da placa.

Tabela 1- Resultados obtidos pelo sistema proposto.

	TOTAL	20 km	30 km	40 km	60 km	80 km
Taxa de detecção	45,3%	26,8%	36,8%	60,8%	31,4%	62,2%
Taxa de reconhecimento	91,1%	100,0%	100,0%	99,7%	92,8%	100,0%
Total de placas						
Detectadas e reconhecidas	41,0%	26,8%	36,8%	60,6%	21,06%	62,2%

A Tabela 1 analisa cada frame do vídeo separadamente, entretanto vale ressaltar que as placas foram detectadas e reconhecidas adequadamente em todos os vídeos, e em todos os casos, a maior parte do reconhecimento é obtida com sucesso. Caso a análise deste artigo fosse realizada com vista no acerto da detecção e do reconhecimento do vídeo, a taxa de acerto seria de 100%.

V. CONCLUSÃO

Os resultados finais mostram que o sistema apresenta-se como promissor para identificação e reconhecimento de placas de sinalização, em que todos os vídeos tiveram a placa detectada e reconhecida adequadamente.

Entretanto, analisando cada frame separadamente, constata-se que a principal deficiência do software é claramente a etapa de detecção das placas, onde se obteve uma média de 45,3%. Já na etapa de reconhecimento o software apresentou resultados bastante satisfatórios, com 96,6% de acerto.

Como sugestão a trabalhos futuros é recomendada a utilização de técnicas de detecção mais eficientes para este caso da detecção de placas, onde o presente artigo mostrou maior deficiência com apenas 45,3% das placas detectadas. Recomenda-se também aplicar outras técnicas de reconhecimento dos caracteres para aumentar o rendimento da rede neural MLP.

VI. REFERÊNCIAS BIBLIOGRÁFICAS

- Arbib, M. A. (2003). The handbook of Brain Theory and Neural Networks, M.I.T Press.
- ALEXANDRIA, A. R. (2005). Sistema de reconhecimento óptico de algarismos para medidores convencionais de energia.
- Bishop, C. M. (2006). Pattern Recognition and Machine Learning, Springer.
- DUDA, R., HART, P. (1972). Use of the Hough transformation to detect lines and curves in pictures. Comm. Of ACM 15, 1. 11-15.
- Gonzalez, R. C. and Woods, R. (2010). Digital Image Processing, 3a edn, Pearson Prentice Hall, New Jersey.
- Gonzalez, R. C., Woods, R. E. and Eddins, S. I. (2009). Digital Image Processing, 3rd edn, Gatesmark Publishing, Knoxville.
- Haykin, S. O. (2008). Neural Networks and Learning Machines, Pearson Prentice Hall.
- Ruck, D. W., Rogers, S. K., Kabrisky, M., Oxley, M. E. and Suter, B. W. (1990). The multilayer perceptron as an approximation to a bayes optimal discriminant function, IEEE TRANSACTIONS ON NEURAL NETWORKS 1.
- Widrow, B. and Winter, R. (1988). Neural nets for adaptive filtering and adaptive pattern recognition, IEEE Computer 21: 25-39.

VII. COPYRIGHT

Direitos autorais: Os autores são os únicos responsáveis pelo material incluído no artigo.