

Reconhecimento e segmentação de caracteres em placas de veículos usando análise de projeção e redes neurais

***Abstract.** This article presents a system for the recognition of vehicles plates based on techniques of Image Processing and Artificial Neural Networks. The system uses projection analysis in order to segment characters in a image of a vehicle license plate. For each segmented character, the system uses neural network to recognize it. In this paper, we implemented a variation in the projection analysis, involving a cut-off threshold. This variation is compared to original algorithm.*

***Resumo.** Este artigo apresenta um sistema de reconhecimento de placas de veículos baseado em técnicas de Processamento de Imagens e Redes Neurais Artificiais. O sistema utiliza análise de projeção para segmentar os caracteres de uma imagem correspondente a placa de um veículo. Para cada caractere segmentado, utiliza-se uma rede neural para reconhecê-lo. Neste trabalho também foi implementada uma variação na análise de projeção, envolvendo um limiar de corte. Esta variação é comparada com o algoritmo original.*

1. Introdução

O reconhecimento automático de caracteres em placas de veículos é um sistema inteligente que utiliza também técnicas de processamento de imagens. De acordo com Guingo, Thomé e Rodrigues (2002), hoje em dia torna-se cada vez mais necessário à utilização desses sistemas na engenharia de tráfego devido ao grande aumento de veículos, pois auxiliam no controle e monitoramento de tráfego com muita eficiência e baixo custo.

O ganho com a implantação de um sistema automático beneficia na redução de custo em diversos setores como fiscalização e controle de tráfego, porém além da redução de custos, o principal ganho é rapidez em que a informação é disponibilizada.

Neste sentido o presente trabalho estuda as técnicas de processamento de imagens e redes neurais artificiais (RNA) [Silva, I. N., Spatti, D. H. e Flauzino, R. A. 2010], a fim de realizar a segmentação e o reconhecimento de caracteres das imagens de placas de veículos através da construção de um protótipo de software que processa e analisa as informações contidas nas fotos. A primeira etapa da construção do sistema consiste em transformar a imagem da placa recortada em escala de cinza e binarizá-la, a segunda etapa compõe a segmentação da placa, por fim, o reconhecimento dos caracteres extraídos da imagem através da RNA. Além disso, pretende-se comparar dois algoritmos de localização de caracteres: um algoritmo proposto com outro já existente.

2. Arquitetura do sistema proposto

Filho e Neto (1999) definem que um sistema de visão artificial (SVA) é capaz de adquirir, processar, e interpretar imagens correspondentes a cenas reais. A Figura 1 representa um diagrama de blocos de uma SVA.

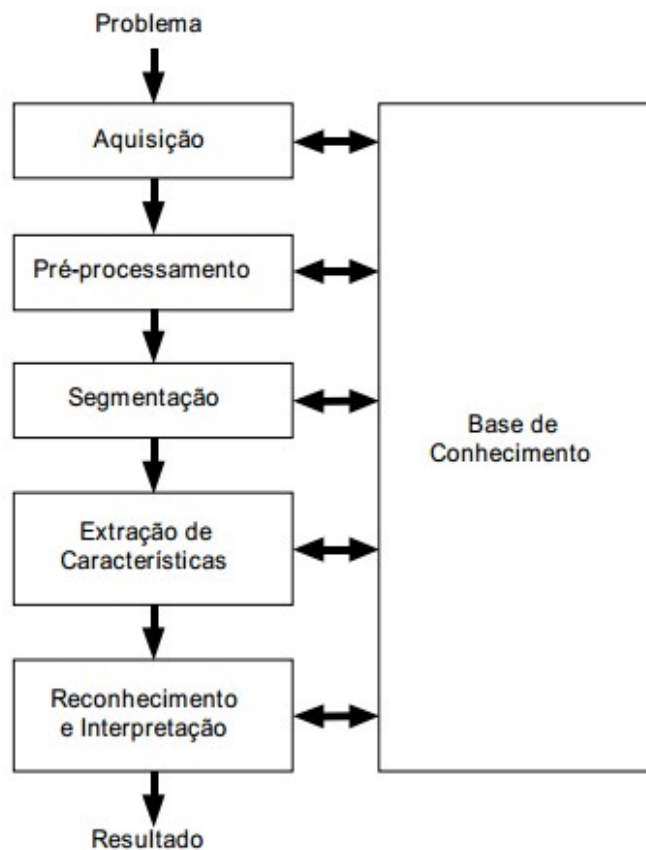


Figura 1. Um Sistema de Visão Artificial (SVA) e suas principais etapas [Filho e Neto 1999, p. 9]

- **Aquisição da imagem:** A primeira etapa nesse processo é a aquisição da imagem, a qual necessita de uma câmera para a captura das fotografias. Esta etapa resulta em uma imagem de veículo.
- **Pré-processamento:** A imagem produzida pela aquisição pode apresentar diversas imperfeições, como ruídos, contraste inadequado, junção de caracteres. As imagens são realçadas para destacar detalhes e restauradas quando estão deterioradas. A função do pré-processamento é aumentar a qualidade das imagens para facilitar todo o processo restante.
- **Segmentação:** Esta fase consiste em dividir a imagem em partes significativas, ou seja, os objetos e o fundo. Uma das operações mais simples de segmentação é a separação por cor. Os algoritmos de segmentação tentarão localizar retângulos e a partir deles identificar o que contém o mesmo padrão que as placas de automóveis, a partir da localização da placa será efetuada a localização de cada caractere.
- **Extração de características:** A extração de características consiste em buscar atributos como: medir nível de intensidade médio ou cor, textura, formas geométricas, dimensões. Nesse passo são extraídas com precisão as características de cada dígito, a fim de diferenciar dígitos parecidos, por exemplo, os números “1” e “7”.
- **Reconhecimento e Interpretação:** O reconhecimento é o processo para adicionar rótulos em objetos, já a interpretação, tem a função de adicionar significado aos objetos reconhecidos. Nesta última etapa, o processo de interpretação irá verificar se o conjunto de objetos pertence a uma sequência válida de dígitos.

- Base de conhecimento: Armazena conhecimento sobre o problema, essa base guia cada etapa para a solução. Na etapa de segmentação, a base de conhecimento tem função de orientar a pesquisa por objetos com formato de retângulo.

O trabalho proposto aplica este modelo de sistema de visão artificial usando uma arquitetura modular como é vista na Figura 2. Para efetuar o reconhecimento de placas de veículos, são necessários os seguintes passos:

- O primeiro passo é a aquisição da imagem, a captura correta da fotografia é fundamental para o desempenho do sistema.
- Logo após, as fotografias são processadas a fim de corrigir a inclinação, ruídos, brilho ou contraste, etc.
- O passo seguinte é analisar quais informações são necessárias e quais informações da imagem podem ser descartadas. É nessa etapa que a região da placa é recortada e o restante da foto é descartado.
- Após a localização da placa, a imagem é transformada em tons de cinza e binarizada. Geralmente em 0 (preto) e 1 (branco) para destacar os objetos do fundo.
- Feito esses passos, é realizada a extração dos caracteres da placa.
- Por fim, é feito o reconhecimento dos caracteres através de uma rede neural treinada.

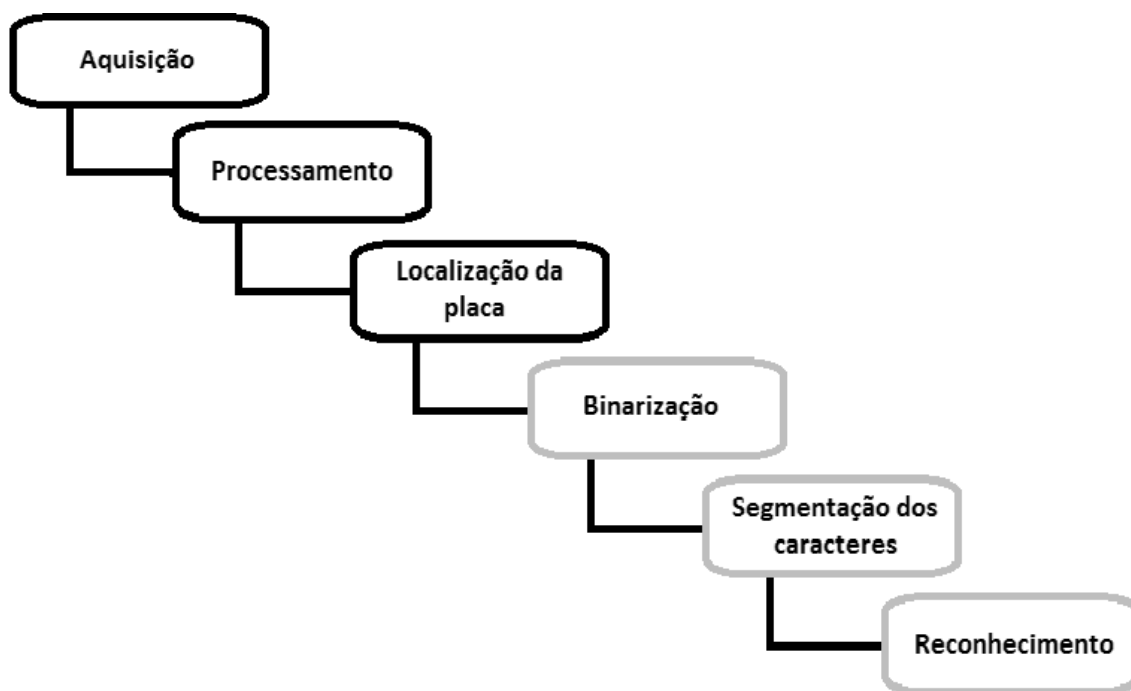


Figura2. Processo para reconhecimento de placas

O objetivo deste trabalho é realizar as últimas três fases da Figura 2: Binarização, Segmentação dos caracteres e Reconhecimento. Supõe-se que uma placa foi corretamente localizada pela fase de Localização de Placa e somente uma imagem contendo a placa é fornecida como entrada para o módulo de Binarização. A seguir são explicadas as técnicas usadas para as fases propostas neste trabalho.

2.1 Binarização

Para binarização foi utilizado o método de Otsu. Ele é baseado na análise discriminante, tem finalidade de definir um limiar ótimo de uma dada imagem. O ponto de corte é

obtido supondo que os *pixels* da imagem podem ser divididos em duas classes, o fundo (c_1 : intervalo $[1, \dots, t]$ de níveis de cinza) e o objeto (c_2 : intervalo $[t+1, \dots, L]$) ou vice-versa. O número de pixels do nível i é denotado por n_i e o número total de pixels é dado por $N = n_1 + n_2 + \dots + n_L$. (Otsu, 1979). A probabilidade para cada classe sendo $p_i = n_i/N$ é representada pela seguinte equação:

$$\omega_1 = \Pr(c_1) = \sum_{i=1}^t p_i$$

$$\omega_2 = \Pr(c_2) = \sum_{i=t+1}^L p_i$$

Após o cálculo da probabilidade das classes c_1 e c_2 é necessário definir suas respectivas média:

$$\mu_1 = \sum_{i=1}^t i \Pr(i|c_1) = \sum_{i=1}^t \frac{i \cdot p_i}{\omega_1}$$

$$\mu_2 = \sum_{i=t+1}^L i \Pr(i|c_2) = \sum_{i=t+1}^L \frac{i \cdot p_i}{\omega_2}$$

A variância entre as classes é representada por:

$$\sigma_b^2 = \omega_1 * \omega_2 * (\mu_2 - \mu_1)^2$$

O limiar ótimo é determinado pela máxima variância encontrada.

2.2 Segmentação por análise de projeção

De acordo com Mao, Rosenfeld, Kanungo (2003), Casey e Lecolinet (1996) a segmentação por análise de projeção consiste em projetar horizontalmente a quantidade de *pixels* pretos (a projeção será feita de acordo com a cor dos objetos, neste caso, os objetos são *pixels* pretos) de cada linha e também projetar verticalmente a quantidade de *pixels* pretos de cada coluna.



Figura 3. Projeção de placa de veículo: (a) projeção horizontal; (b) projeção vertical [Belvisi, Freitas, Figueiredo, Marcovith 1999]

A projeção horizontal da Figura 3(a) é feita para identificar a altura de cada objeto, já na Figura 3(b) a projeção vertical serve pra separar cada letra e cada número. Esse algoritmo possui falha em imagens com ruídos, quando um caractere possui ligação com o caractere vizinho fica impossível definir um ponto de corte entre os objetos.

2.2.1 Algoritmo proposto para segmentação

O algoritmo de análise de projeção pode apresentar falha em imagens com ruídos, pois quando um caractere possui ligação horizontal com o caractere vizinho torna-se difícil separá-los apenas com esse algoritmo.

O algoritmo proposto possui o mesmo funcionamento que o algoritmo de análise de projeção, porém existe a implementação do ponto de corte que consiste em separar os objetos interligados. Para definir o ponto de corte entre os caracteres é projetada uma linha horizontalmente na projeção vertical da imagem, como mostrado na Figura 4. Esta linha equivale ao valor da média das projeções verticais das imagens das letras contidas na Figura 5 menos o desvio padrão dos valores obtidos para a média de cada imagem.

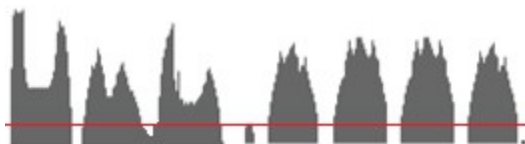


Figura 4. Exemplo de projeção com algoritmo proposto [Belvisi, Freitas, Figueiredo, Marcovith 1999]

Para calcular essa média foi utilizado somente as 36 amostras de caracteres segmentados da Figura 5. Em seguida todas as imagens foram ajustadas para um tamanho padrão, somando suas respectivas projeções verticais e por fim, foi calculada a média do resultado dessa soma e seu desvio padrão.

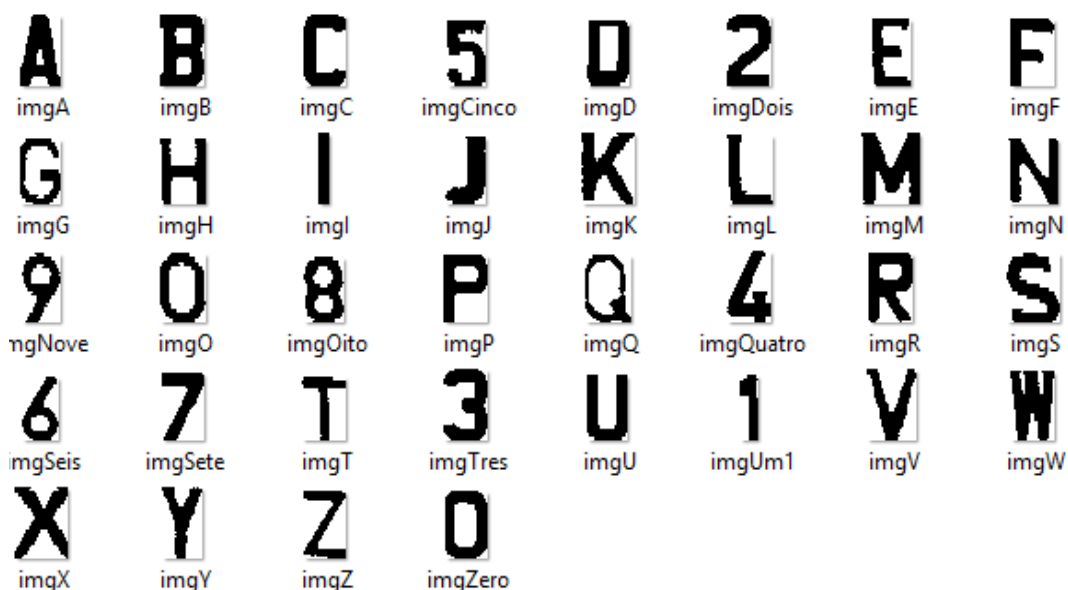


Figura 5. Imagem contendo imagens de letras e números usadas para calcular o limiar de aceitação na análise de projeção.

O Código 1 ilustra o pseudocódigo que calcula este limiar. Neste pseudocódigo, a função *mediaVetorial* retorna a média de um vetor passado como parâmetro. A função *projecaoVertical* retorna o vetor contendo a projeção vertical de uma imagem. O vetor *imagem* representa as imagens da base utilizada para calcular este limiar (Figura 5). Os valores de média para cada imagem são armazenados no vetor *valoresMediosProjecao*. A função *desvioPadraoVetorial* calcula o desvio padrão de um vetor.

```
Para i = 1 a 36 faça
    m = mediaVetorial(projecaoVertical(imagem[i]));
    valoresMediosProjecao[i] = m;
Fim
media = mediaVetorial(valoresMediosProjecao);
desvio = desvioPadraoVetorial(valoresMediosProjecao);
limiar = media - desvio;
```

Código 1. Pseudocódigo para cálculo do limiar

2.3 Reconhecimento

Para realizar o reconhecimento dos caracteres foi criada uma rede neural para que ela possa aprender a classificar cada padrão. A rede neural utilizada foi o perceptron de múltiplas camadas (MLP – *Multilayer Perceptron*), pois conseguem resolver problemas complexos, por exemplo, reconhecimento de padrões e otimização de sistemas (Russel e Norvig, 2004). O treinamento foi feito utilizando o algoritmo de retropropagação (*backpropagation*).

3. Metodologia

Um sistema de reconhecimento automático de caracteres em placas de automóveis possui todas as etapas descritas previamente. Neste trabalho algumas etapas foram feitas manualmente, visto que foi suposta a existência dos módulos anteriores às etapas propostas neste trabalho (como visto na Figura 2). As seções seguintes descrevem como foram testadas as etapas descritas.

3.1 Obtenção de dados

Para a realização deste trabalho, foram utilizadas duas bases de imagens sendo que uma das bases contém 674 imagens com condições reais e foi cedida pelo Detran de Sergipe, a segunda base contém 100 imagens, que pertence ao LPDSI/CBPF, disponível no site: <<http://www.cbpf.br/cat/lpdsi>>. Para a realização dos testes foram utilizadas 468 imagens da base, pois algumas imagens foram descartadas por não ser humanamente identificáveis, possuem placas parcialmente recortadas ou por não possuírem placas nos veículos.

3.2 Localização da placa

Todas as placas das imagens foram localizadas e recortadas manualmente gerando uma nova imagem contendo apenas a placa, logo após, foram redimensionadas para o tamanho 230x80 para seguir o mesmo padrão.

3.3 Binarização

Após a localização, realizou-se a transformação da placa em escala de cinza tornando o processamento simples. Após a transformação em tons de cinza, a imagem foi binarizada utilizando o método de Otsu, os objetos foram representados por *pixels* pretos enquanto o fundo por *pixels* brancos.

3.4 Segmentação dos caracteres

O processo de segmentação é fundamental para o sucesso do sistema, a realização incorreta dessa etapa influencia diretamente no processo de reconhecimento, para realizar a segmentação foram aplicados o método de análise de projeção (Figura 6) e o método do algoritmo proposto.



Figura 6. Resultado da segmentação de caracteres pelo algoritmo de análise de projeção

As 468 imagens binárias foram utilizadas tanto na segmentação por análise de projeção quanto pelo algoritmo desenvolvido. No primeiro algoritmo, 200 placas foram segmentadas corretamente, enquanto que no segundo, os caracteres foram extraídos corretamente de 248 placas. No processo seguinte foram utilizados os 1736 caracteres (744 letras e 992 números) das 248 placas resultantes do processo do algoritmo proposto, que, como consequência, resultou em um maior número de imagens para teste da etapa de reconhecimento de caracteres.

3.5 Reconhecimento

Para realizar o reconhecimento de cada caractere foram criadas duas redes neurais, uma para as letras e outra para os números, evitando problemas como a classificação errada dos caracteres, por exemplo, seria difícil distinguir o número 0 da letra O ou o número 1 da letra I.

A estrutura da rede neural das letras é a seguinte: 1184 entradas (correspondentes a cada pixel de uma imagem binário de letra contendo 32 x 37 pixels), 60 neurônios na camada intermediária (essa quantidade de neurônios foi usada pois apresentou um resultado satisfatório no processo de aprendizagem) e 26 neurônios na camada de saída (cada saída corresponde a uma das letras do alfabeto). A rede neural de números é formada por 1184 entradas, 60 neurônios na camada oculta e 10 neurônios de saída (os 10 neurônios representam as saídas de 0 a 9).

A rede de letras foi treinada e testada respectivamente com 597 e 147 imagens de letras. A rede de números foi treinada e testada respectivamente com 796 e 196 imagens

de números, ou seja, 80% dos caracteres foram utilizados para o treinamento e 20% para a realização dos testes.

4. Resultados

Com o objetivo de avaliar experimentalmente os algoritmos de segmentação de caracteres e reconhecimento, foram realizados testes envolvendo as imagens em condições reais descritas na Seção 3.1. A Tabela 1 apresenta os resultados para a segmentação e a Tabela 2, os resultados para o reconhecimento de letras e números.

Tabela1. Resultadoda segmentação

Algoritmo	Placas testadas	Sucesso	Percentual de sucesso
Análise de projeção	468	200	42,7%
Proposto	468	248	53,0%

Tabela2. Resultadodo reconhecimento de caracteres

Reconhecimento	Total testado	Sucesso	Percentual de sucesso
Letras	147	122	82,9%
Números	195	165	84,6%

Observa-se na Tabela 1 que o algoritmo proposto obteve desempenho melhor, porém em alguns casos o algoritmo de análise de projeção detectou os sete caracteres corretamente, enquanto o algoritmo proposto não obteve sucesso, por exemplo, em alguns casos quando o caractere é um “H”, o algoritmo detectou que a linha horizontal da letra seria um ruído e descartou, dividindo em dois objetos.

A grande maioria dos erros do algoritmo de análise de projeção é causada pelos ruídos, pois o algoritmo não é capaz de identificar um ponto de corte entre dois caracteres que possuem junção, já o algoritmo proposto conseguiu encontrar corretamente o ponto de corte.

A Figura 8 mostra a classificação dos erros encontrados na análise de projeção. Nota-se que houve uma taxa maior de erros no algoritmo proposto em relação à última classificação (Letra não extraída corretamente) isso devido ao problema de separação de caractere equivocado, pois como já foi explicado em alguns casos o algoritmo detectou que letras como “H”, “J” e “U” deveriam ser separadas. Já na segunda classificação (Junção de caracteres) é apresentada a maior taxa de erros pelo algoritmo de análise de projeção, pois esse algoritmo não implementa técnicas de detecção de ruídos.

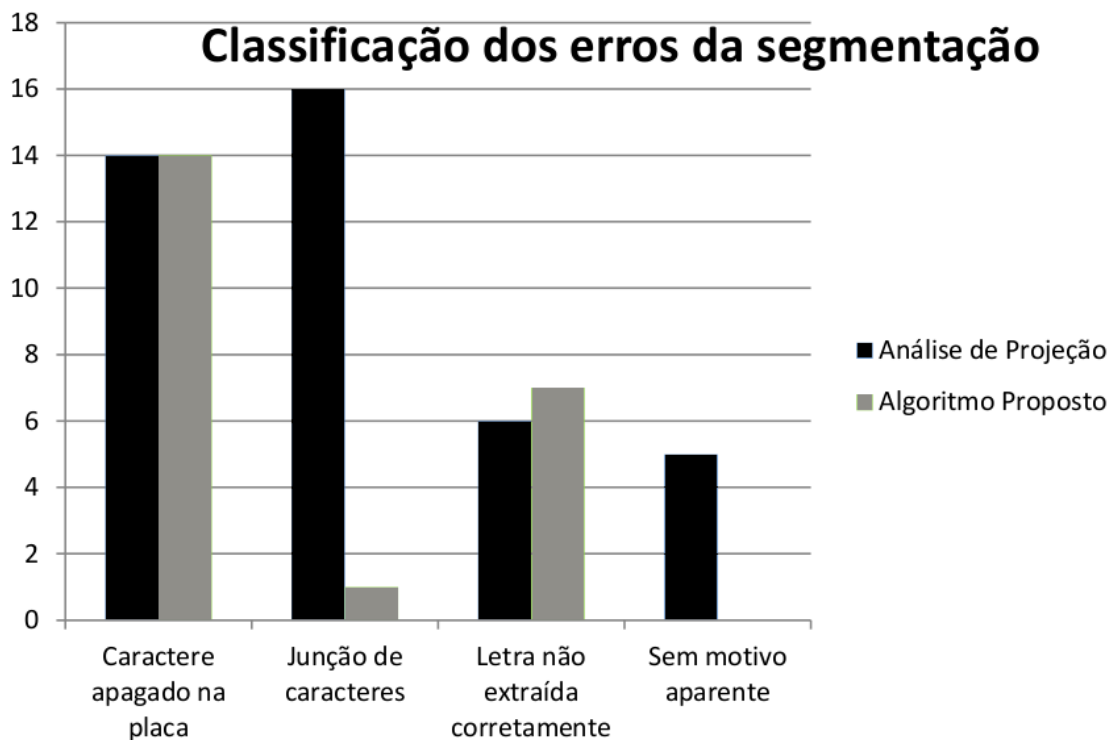


Figura 7. Classificação de erros da segmentação

Uma possível solução para esse problema é a eliminação de ruídos utilizando erosão. A erosão é uma operação morfológica que reduz o tamanho geométrico de objetos contidos em uma imagem, ou seja, o processo retira informações contidas nas imagens gerando erosão nas áreas [Gonzalez e Woods 2002].

Quanto ao resultado da Tabela 2, observa-se que devido a este problema com ruídos a taxa de acerto não foi elevada, pois o pré-processamento retiraria imperfeições da imagem facilitando todo o processo de reconhecimento. Outro problema foi uma pequena inclinação encontrada em algumas placas. Além disso, o treinamento não foi feito com um número suficiente de amostras visto que algumas letras possuem maior quantidade de exemplos que outras tornando a rede irregular, ou seja, a rede se torna especialista em reconhecimento de apenas algumas letras.

5. Conclusão

Esse trabalho apresentou aspectos relacionados à área de processamento de imagens digitais e reconhecimento de padrões, com intuito de definir as etapas necessárias para ser desenvolvido o sistema de reconhecimento de placas de veículos.

O estudo comparou dois algoritmos de segmentação de imagens, um baseado na projeção de histograma vertical e horizontal e o outro uma modificação deste, ou seja, o algoritmo proposto. Foram treinadas e testadas duas redes: uma para reconhecer cada letra e outra para cada número, a fim de classificar as imagens resultantes da etapa de segmentação. Dessa forma, o algoritmo de análise de projeção obteve uma taxa de acerto de 42,7%. O algoritmo proposto obteve resultados melhores em relação ao algoritmo de análise, pois apresentou uma taxa de acerto de 53,0%.

Como trabalhos futuros, pretende-se testar valores de limiar para o algoritmo de segmentação proposto que envolvam em seu cálculo mais imagens de caracteres e letras.

Também, pretende-se usar o método de erosão da morfologia matemática no processo de segmentação dos caracteres para eliminação de ruídos.

Referências

- Belvisi, R., Freitas, R., Figueiredo, Raifran, M., D., Matumoto, G. e Bianchi, R. A. C. (1999) “Um sistema de reconhecimento automático de placas de automóveis”, <http://fei.edu.br/~rbianchi/publications/enia99.pdf>, Novembro.
- Casey, R. e Lecolinet, E. (1996) “*A survey of methods in strategies in character segmentation*”, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 18.
- Filho, O. M. e Neto, H. V. (1999) “Processamento Digital de Imagens”, Rio de Janeiro: Brasport.
- Gonzalez, R. C. e WOODS, R. E. (2002) “*Digital Image Processing*”. 2. ed. Upper Saddle River, NJ, USA, Prentice-Hall.
- Guingo, B. C., Thomé, Antonio C. G. e Rodrigues, R. J. (2002) “Reconhecimento automático de placas de veículos automotores através de redes neurais artificiais”, UFRJ, RJ.
- Mao, S., Rosenfeld, A. e Kanungo, T. (2003) “*Document structure analysis algorithms: a literature survey*”, Proceeding of SPIE/IS&T, vol. 5010.
- Otsu, N. (1979) “*A threshold selection method from gray-level histograms*”, IEEE Transactions on System Man Cybernetics, Vol. SMC-9, nº1.
- Russell, S. e Norvig, P. (2004) “Inteligência Artificial”, Rio de Janeiro: Elsevier.
- Silva, I. N., Spatti, D. H. e Flauzino, R. A. (2010) “Redes Neurais Artificiais: para engenharia e ciências aplicadas”, São Paulo: Artliber Editora Ltda.