6 O RECONHECEDOR INTELIGENTE DE SINAIS DE TRÂNSITO BRASILEIROS (RIST-Br)

Neste capítulo, será apresentado o desenvolvimento do Reconhecedor Inteligente de Sinais de Trânsito Brasileiros, o RIST-Br, implementado pelo autor desta monografia. O RIST-Br foi projetado para analisar imagens digitais do ambiente, obtidas a partir de uma câmera fotográfica digital, e indicar se existem nessas imagens determinados sinais de trânsito brasileiros. O sistema do RIST-Br é composto por três softwares desenvolvidos pelo autor. Esses softwares foram elaborados com o auxílio da biblioteca de processamento de imagens "Sistema IMG" [Kim 2004], apresentada na seção 3.3 do Capítulo 3, e do simulador de redes neurais artificiais SNNS [SNNS 1998], apresentado na seção 4.3 do Capítulo 4. A metodologia adotada para o desenvolvimento do RIST-Br baseia-se nos passos fundamentais para o processamento de imagens (seção 3.1.2 do Capítulo 3) e também na metodologia apresentada no Capítulo 5, com algumas adaptações com relação ao experimento apresentado na seção 5.2.

6.1 Objetivos e Relevância

No campo da Visão Computacional (VC), o Reconhecimento Inteligente de Sinais de Trânsito (RIST) preocupa-se com a identificação da sinalização de trânsito presente em uma imagem do ambiente, com o objetivo de auxiliar o motorista em sua tarefa de navegação [Johansson 2002].

Conforme mencionado na introdução deste trabalho (Capítulo 1), a densidade do trânsito de veículos nas rodovias do Brasil - especialmente nos grandes centros urbanos - tem causado, entre outros problemas, um crescente número de acidentes com vítimas [DENATRAN 2001]. Sistemas computacionais e dispositivos de apoio ao motorista [Paromtchik 1997] , como os apresentados na seção 2.2 do Capítulo 2, têm potencial para reduzir o número de acidentes nas rodovias do país [Paromtchik 1997]. Dessa forma, esses sistemas de apoio poderão contribuir para o bem estar social e econômico da população, na medida em que vidas seriam poupadas e perdas econômicas seriam evitadas.

Como exposto na seção 2.3.2, do Capítulo 2, um sistema de RIST oferece diversas vantagens para o motorista cumprir com segurança a tarefa de dirigir um veículo. Sendo essa tarefa essencialmente dependente da visão, um sistema de RIST tem condições para ajudar o motorista a: (i) permanecer alerta para os sinais de trânsito; (ii) evitar e resolver situações de risco (p.ex.: presença de obstáculos na pista; estar trafegando a uma velocidade acima do permitido em determinado local; tentar tomar o sentido oposto ao determinado para a via); (iii) manter-se informado sobre rotas a seguir.

Dessa forma, este trabalho tem como objetivo o desenvolvimento de um sistema de RIST com potencial para apoiar o motorista na tarefa de estar atento às informações apresentadas

pelos sinais de trânsito e, assim, contribuir para o aumento da segurança do tráfego de veículos nas rodovias do Brasil.

6.2 O Projeto do RIST-Br

O RIST-Br foi projetado para utilizar técnicas de processamento de imagens e de inteligência artificial. As técnicas de processamento de imagens utilizadas durante o desenvolvimento do RIST-Br serão apresentadas com maiores detalhes nas seções a seguir. Os fundamentos de processamento de imagens foram apresentados no Capítulo 3. No que diz respeito à inteligência artificial, foram utilizadas no RIST-Br técnicas envolvendo as Redes Neurais Artificiais (RNA), cujos fundamentos foram apresentados no Capítulo 4. O papel das redes neurais - e os detalhes de seu projeto - no sistema RIST-Br será descrito nas próximas seções, conforme são apresentadas as demais técnicas e metodologias adotadas.

6.1.1 A Captura das Imagens de Entrada

Como o sistema RIST-Br deveria analisar imagens do ambiente em busca de sinais de trânsito, optou-se pela utilização de uma câmera fotográfica digital - gentilmente cedida pelo orientador desta monografia - para a captura das imagens a serem analisadas pelo RIST-Br. A câmera utilizada pode ser visualizada na Figura 6.1, abaixo.

Figura 6.1 Câmera Fotográfica Digital Sony DSC-P72 [Sony 2004]

A base de imagens foi constituída por um total de 148 imagens capturadas, em cores, com resolução (quantidade de pixels da imagem) de 3,1 megapixels (3,1 milhões de pixels), arquivadas em formato JPEG (*Joint Photographic Experts Group* - Grupo de Especialistas em Fotografia) [Murray 1996]. Todas as imagens foram capturadas no período diurno. As condições climáticas variavam de ensolarado a nublado.

Todas as fotografias foram capturadas da seguinte maneira: (i) no modo de exposição automática (em câmeras fotográficas, o modo de exposição automática controla, sem a interferência do fotógrafo, a quantidade de luz e o tempo durante o qual a mesma irá incidir sobre o sensor, ou sobre o filme); (ii) sem a utilização da aproximação (*zoom*) ótica ou digital; (iii) sem a utilização de iluminação auxiliar (p.ex.: *flash*); (iv) sem a utilização de quaisquer efeitos de imagem; (v) na resolução máxima permitida pela câmera digital: 2048 colunas x 1536 linhas de pixels (3.145.728 pixels, ou aproximadamente 3,1 megapixels); (vi) em arquivo de formato JPEG [Murray 1996], modelo de cores RGB (apresentado no Capítulo 3, seção 3.2.1), com 8 bits por componente de cor (total de 24 bits de cor, aproximadamente 16,8 milhões de cores - Capítulo 3, seção 3.1.1); (vii) com o fotógrafo imóvel, posicionado em pé, a uma distância de aproximadamente sete metros do sinal de trânsito de interesse. Além disso, todas as fotografias foram capturadas de maneira que o plano do sinal de trânsito estivesse aproximadamente paralelo ao plano do sensor digital da câmera (Capítulo 3, Figura 3.5).

Nenhuma imagem sofreu edição de qualquer natureza após a captura, antes de ser submetida ao sistema RIST-Br. Contudo, algumas imagens selecionadas para publicação através desta monografia foram editadas para que as placas de licenciamento de alguns veículos, e os rostos de alguns pedestres e motoristas, pudessem ser apagados - preservando assim sua privacidade.

Os sinais de trânsito presentes nas imagens foram amostrados aleatoriamente. Esses sinais de trânsito correspondem a uma amostra dos sinais encontrados: nas vias da região central da cidade de São Leopoldo (localizada na Grande Porto Alegre - Rio Grande do Sul); ao longo da BR-116 (trecho entre as cidades de Canoas e Novo Hamburgo, na Grande Porto Alegre); e em Novo Hamburgo.

Tipos de Sinais	Total de Imagens
Parada Obrigatória	33
Área Escolar	24
Sentido Obrigatório	53
Velocidade Máxima	23
Outros Sinais	15
TOTAL	148

Tabela 6.1 Tipos de Sinais de Trânsito Fotografados para a Base de Imagens

Nas figuras abaixo, são apresentadas algumas imagens da base construída, conforme o tipo de sinal de trânsito fotografado. O código abaixo de cada imagem é a identificação criada pelo autor com o único propósito de organizar as imagens na base.

(a) p276

(b) p290

(c) p292

(d) p312

(e) p369

(f) p370

Figura 6.2 Exemplos de Sinais de Parada Obrigatória: (a) sinal oculto pelos galhos da árvore; (b) sinal parcialmente à sombra da árvore; (c) sinal iluminado pelo Sol; (d) sinal à sombra da árvore; (e) sinal parcialmente oculto; (f) sinal em dia parcialmente nublado

(a) e275

(b) e286

(c) e355

(d) e410

(e) e425

(f) e431

Figura 6.3 Exemplos de Sinais de Área Escolar: (a) sinal à sombra da àrvore; (b) sinal iluminado pelo Sol; (c), (d) e (e) mostram sinais em dia nublado;(f) sinal em dia parcialmente nublado

(a) sd334

(b) sd358

(c) sd426

(d) se265

(e) se299

(f) se337

Figura 6.4Exemplos de Sinais de Sentido Obrigatório: (a) sinal iluminado pelo Sol; (b) e (c) mostram sinais em dia nublado; (d) sinal iluminado pelo Sol; (e) sinal sob a sombra; (f) sinal em dia nublado

(a) v378

(b) v402

(c) v403

(d) v406

(e) v408

(f) v409

Figura 6.5 Exemplos de Sinais de Velocidade Máxima Permitida: de (a) até (d) mostram sinais em dias parcialmente nublados; (e) e (f) mostram sinais em dia nublado

(a) oe287

(b) oe304

(c) os314

(d) os296

(e) oe293

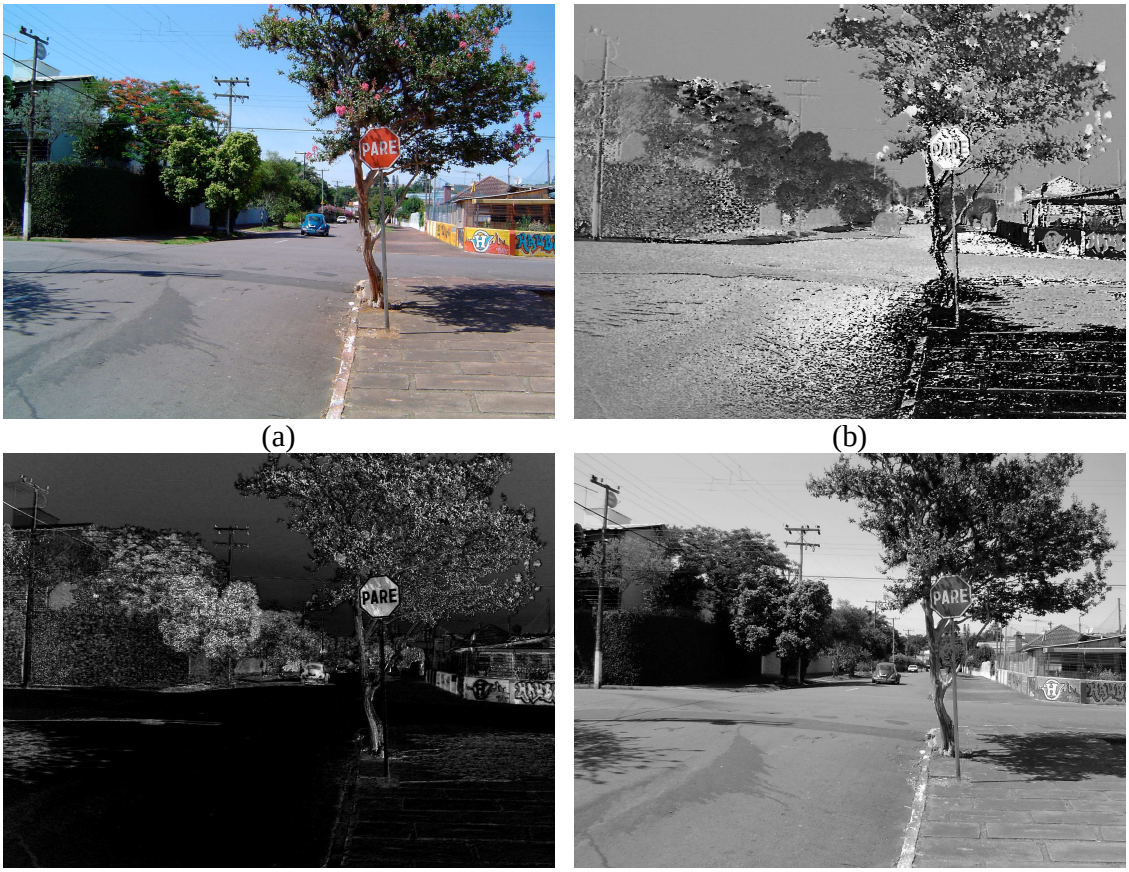
(f) oe294

Figura 6.6 Exemplos de Outros Sinais de Trânsito: (a) e (b) mostram sinal de Saliência ou Lombada iluminado pelo Sol e à sombra, respectivamente; (c) sinal de Proibido Acionar Buzina ou Sinal Sonoro, localizado à sombra; (d) sinal de Proibido Virar à Esquerda, iluminado pelo Sol; (e) e (f) mostram sinal de Semáforo à Frente, à sombra e iluminado pelo Sol, respectivamente

6.2.1 O Pré-processamento das Imagens de Entrada

O primeiro passo do pré-processamento das imagens de entrada foi a sua conversão para o modelo de cores HSI (adotado devido às suas características semelhantes às apresentadas pelo modelo HSV no Capítulo 3, seção 3.2.2), a partir de seu modelo RGB original.

Essa conversão foi efetuada através do programa de conversão "RGB2HSIs", presente na biblioteca "Sistema IMG", descrita na seção 3.3 do Capítulo 3. O programa "RGB2HSIs" recebe como entrada a imagem a ser convertida (em cores, modelo RGB, arquivo no formato JPEG) e gera três imagens de saída, todas em arquivos no formato TGA (formato da empresa Truevision) [Murray 1996]. Cada uma das imagens de saída representam, respectivamente, as informações das componentes H, S e I desse modelo de cores, e todas as três são visualizadas em tons de cinza. Nas figuras a seguir, pode ser observado um exemplo deste primeiro passo de pré-processamento.



(a) (b)

(c) (d)

Figura 6.7 Primeiro Passo do Pré-Processamento - Conversão de RGB para HSI (Imagem p370): (a) imagem original; (b) componente de cor (H); (c) componente de saturação (S); componente de intensidade (I)

Apenas o arquivo correspondente à componente H (*hue* - cor) de cada imagem é utilizado para o passo seguinte do pré-processamento: a binarização da imagem segundo o mecanismo atencional (esse mecanismo atencional será apresentado com maiores detalhes na próxima seção - a seção 6.2.2 - que trata da segmentação das imagens de entrada). As imagens que correspondem às componentes de cor S e I são descartadas.

Nesta fase, o usuário do sistema deve fornecer como entrada para o RIST-Br a imagem original e a imagem que corresponde à componente de cor H, e informar qual sinal de trânsito que deseja que o sistema procure. Essa informação é relevante para o pré-processamento porque o RIST-Br irá efetuar uma binarização da imagem com relação à cor que deverá ser encontrada. Essa cor dependerá do tipo de sinal de trânsito (p.ex.: para a Parada Obrigatória e outras Placas de Regulamentação (ver seção 6.2.5.1, logo adiante, neste Capítulo) , a cor procurada será a vermelha; para a Área Escolar, a cor será a amarela) que o usuário deseja que o sistema encontre na cena.

Desse modo, o usuário fornecerá para o sistema apenas as imagens da Figura 6.7 (a) e (b). Abaixo, na Figura 6.8, pode ser visualizada uma ampliação da Figura 6.7 (b).

Figura 6.8 Imagem Correspondente à Componente de Cor (H)

do Modelo HSI para a Imagem de Entrada

Conforme mencionado na seção 3.2.2, no Capítulo 3 - sobre as vantagens do modelo de cores HSI - a componente H é invariante às mudanças na iluminação da imagem. A partir do uso de uma imagem que representa apenas a componente H (Figura 6.8) de uma imagem original de entrada (Figura 6.7 (a)), o sistema RIST-Br rastreia, da esquerda para a direita e de cima para baixo (ver Capítulo 3, Figura 3.1), pixel a pixel, a imagem da Figura 6.8 e a binariza.

A binarização realizada pelo RIST-Br, que é o segundo passo do pré-processamento, gera uma nova imagem, formato BMP (*bitmap* - mapa de bits, de duas cores (1 bit de cor/pixel) [Murray 1996]) cujos pixels "ligados" (pretos) corresponderão aos pixels da Figura 6.8 que estiverem dentro de uma faixa de valores (limiares) correspondentes à cor vermelha (tal faixa de valores foi determinada previamente através de uma análise da imagem da Figura 6.8 com o auxílio de um editor gráfico, e da comparação dessa figura com a imagem original). Esse procedimento de binarização equivale à aplicação de um filtro sobre a imagem.

Assim, podemos visualizar na Figura 6.9, abaixo, o resultado da binarização da Figura 6.8.

Figura 6.9 Resultado da Binarização da Figura 6.8, em Busca do Sinal de Parada Obrigatória (Filtro para a Cor Vermelha)

6.2.2 A Segmentação das Imagens de Entrada

De acordo com a seção 5.3.3, do Capítulo 5, que trata da segmentação das imagens de entrada, foi adotada para o mecanismo atencional do sistema RIST-Br a técnica da computação das característicias visuais primitivas [Rodrigues 2002].

Esse mecanismo atencional utiliza características básicas da imagem, como cores ou texturas. Para o sistema atencional do RIST-br foi selecionada a cor porque a mesma é um elemento característico muito importante na sinalização do trânsito (além de ser um elemento relativamente independente com relação à resolução, ao tamanho e à orientação da imagem capturada [Bender 2003]), servindo para destacar o sinal dos outros elementos da imagem e atrair a atenção do sistema atencional humano.

O resultado da binarização da imagem da Figura 6.8, mostrado na Figura 6.9, demonstra o potencial da técnica atencional adotada. Na imagem da Figura 6.9, podem ser observados os pixels que ficaram situados dentro de uma faixa de valores correspondentes à cor vermelha na Figura 6.8. Assim pode-se notar, através na Figura 6.9, que o sinal de trânsito que está sendo procurado pelo sistema RIST-Br encontra-se em destaque na imagem. Esse destaque é explorado nesta etapa de segmentação pois os sinais de trânsito caracterizam-se, na imagem binarizada (e também nas imagens em cores), por uma região de vários pixels da mesma cor - o que facilita a sua detecção.

A segmentação da imagem de entrada é efetuada a partir da imagem binarizada da Figura 6.9. O sistema RIST-Br rastreia a imagem da Figura 6.9, utilizando uma "cabeça de leitura" de 200x200 pixels. Essa "cabeça de leitura" faz a leitura dos 40 mil pixels sob a mesma e, se o resultado do somatório de pixels "ligados" (pretos) sob a mesma for maior ou igual a dez mil pixels (a exceção ocorre para os sinais de Sentido Obrigatório, para os quais o limiar é de três mil pixels), então o RIST-Br marca essa região sob a "cabeça de leitura" como sendo uma subimagem candidata a ser o sinal de trânsito procurado (isto é, que poderá vir a ser reconhecida como o sinal de trânsito procurado pelo RIST-Br).

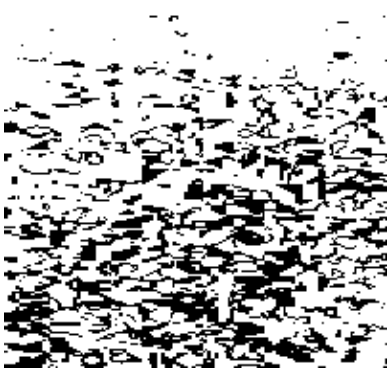
Essa região marcada como subimagem candidata é extraída da imagem binarizada e armazenada em memória secundária (p.ex.: disco magnético rígido). Também é gravado na memória secundária um arquivo em formato texto com as informações de cor (modelo RGB) de todos os 40 mil pixels da imagem original (Figura 6.7 (a)) que encontram-se sob a mesma área em que a subimagem candidata foi encontrada.

Em seguida, a resolução da subimagem candidata é reduzida de 200x200 pixels para 40x40 pixels através de uma reamostragem. Nessa reamostragem, cada pixel de uma coluna da subimagem 40x40 equivale a um pixel extraído a cada cinco colunas da subimagem 200x200, sendo que o mesmo procedimento foi adotado para as linhas: cada pixel de uma linha da subimagem 40x40 equivale a um pixel extraído a cada cinco linhas da subimagem 200x200. Isso resultou em uma miniatura 40x40 pixels da subimagem candidata binária original, conforme pode ser visualizado na Figura 6.10.

A "cabeça de leitura" efetua o rastreamento de toda a imagem em busca de subimagens candidatas, da esquerda para a direita e de cima para baixo, a passos de 50 pixels (tanto na horizontal quanto na vertical), enquanto os procedimentos descritos nos últimos dois parágrafos são executados. Esses procedimentos foram definidos para a análise de imagens com uma resolução fixa de 2048x1536 pixels (2048 colunas x 1536 linhas).

(a) 200x200 pixels

(b) 40x40 pixels



(c) 200x200 pixels

(d) 40x40 pixels

Figura 6.10 Exemplos de Subimagens Candidatas: (a) e (c) de 200x200 pixels; e Suas Respectivas Miniaturas de 40x40 pixels (b) e (d)

Conforme as subimagens candidatas (Figura 6.10 (a) e (c)) são extraídas da imagem binarizada (e as miniaturas são criadas - Figura 6.10 (b) e (d)), suas localizações são assinaladas na imagem original da cena (Figura 6.7 (a)), que foi fornecida como entrada para o RIST-Br juntamente com a imagem correspondente à componente H do modelo HSI

(Figura 6.8). Como resultado, temos a imagem apresentada na Figura 6.11, abaixo, que constitui a saída da detecção atencional do RIST-Br.

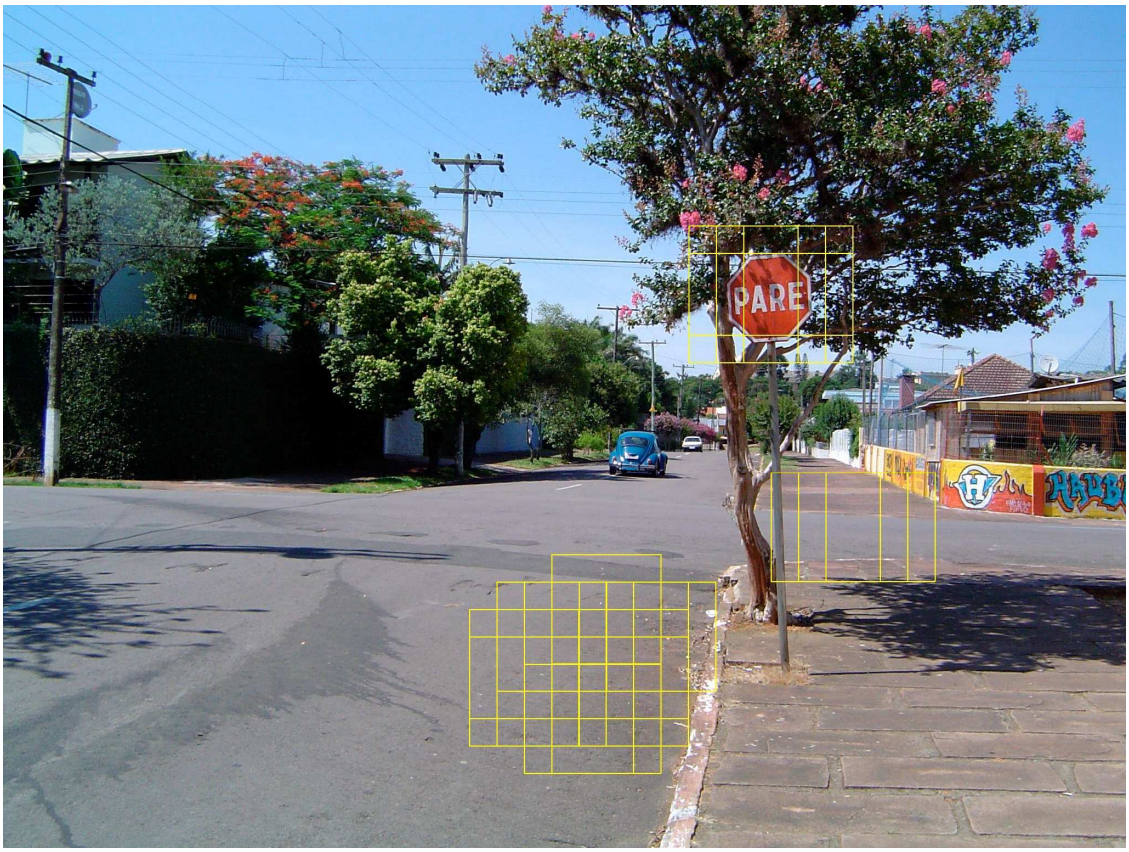


Figura 6.11 Resultado da Detecção de Subimagens Candidatas pelo Mecanismo Atencional do RIST-Br

Na Figura 6.11, acima, podem ser observadas, marcadas por quadrados de contornos na cor amarela, as regiões da imagem original onde o mecanismo atencional do RIST-Br encontrou as subimagens candidatas a serem um sinal de Parada Obrigatória. Notar que a "cabeça de leitura" possui uma área de 200x200 pixels, e move-se a passos de 50 pixels (na horizontal e na vertical). Os pequenos quadrados que podem ser vistos ocorrem, portanto, devido à sobreposição de subimagens candidatas encontradas na imagem original pelo sistema. Para o exemplo da Figura 6.11, o sistema atencional detectou 27 subimagens candidatas.

6.2.3 A Representação e a Descrição das Subimagens Candidatas

Neste passo do processamento das imagens, a representação das imagens para o sistema RIST-Br é feita através da miniatura 40x40 pixels (p.ex.: Figura 6.10 (b)) obtida para cada uma das subimagens candidatas, e através de um cálculo que levará em conta as cores da imagem original na área correspondente à da subimagem candidata de 200x200 pixels.

O sistema utiliza, da imagem original (Figura 6.7 (a)), as cores dos pixels presentes sob a mesma área de 200x200 pixels onde foi encontrada a imagem candidata (p.ex.: Figura 6.10 (a)). O cálculo é efetuado sobre um arquivo em formato texto que contém todas as informações de cor (modelo RGB) de cada um dos 40 mil pixels da área correspondente à da subimagem candidata.

Nesse arquivo em formato texto, a informação de cor é encontrada na forma de valores no intervalo [0,255]. Cada três valores do arquivo correspondem à cor de um pixel da imagem (p.ex.: (255,255,255)) no modelo de cores RGB. Um pequeno trecho desse arquivo, representando o cabeçalho e os cinco primeiros pixels da área (da imagem original - Figura 6.7 (a)) sob a subimagem candidata (Figura 6.10 (a)) pode ser visualizado na Figura 6.12 abaixo.

```
200 200

 89  88  86      77  77  77     133 133 133     142 144
143    106 106 108
```

Figura 6.12 Trecho Inicial do Arquivo em Formato Texto da Área, na Imagem Original, Correspondente à Subimagem Candidata - Notar o Cabeçalho (200 200) e os valores para as Componentes R, G e B para os Cinco Primeiros Pixels

O cálculo dos valores para a representação (que é o primeiro passo desta fase de representação e descrição) da cor dessa subimagem são efetuados da seguinte maneira:

Passo 1: Primeiro, o RIST-Br calcula o somatório S_n do número de pixels da subimagem para cada uma das 256 tonalidades (n) possíveis de uma componente de cor. Este é o histograma de uma das componentes de cor da subimagem - que mostra como todos os pixels da subimagem estão distribuídos em cada uma das 256 tonalidades. Esse histograma tem a aparência da Figura 6.13, abaixo.

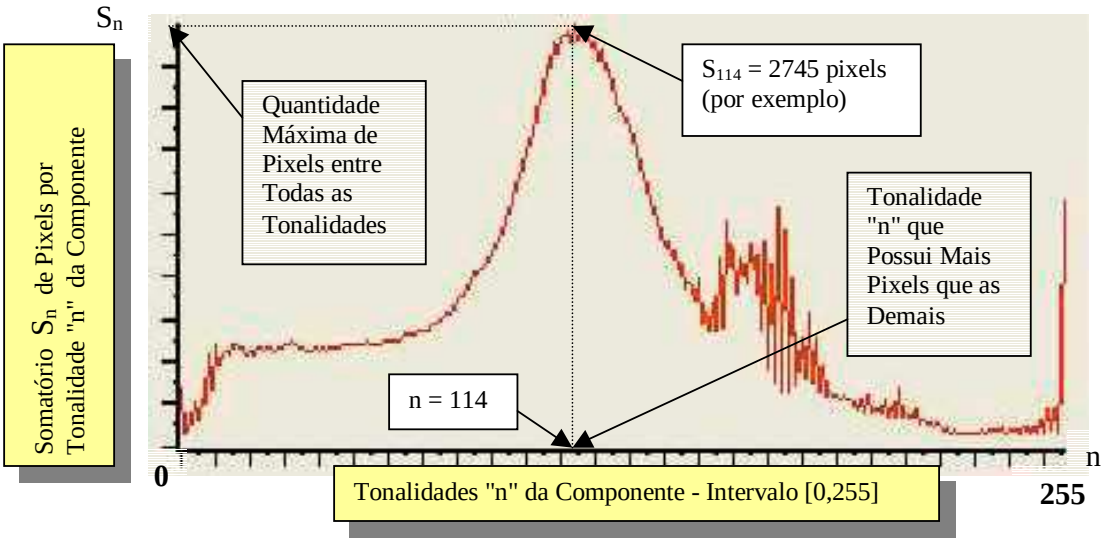


Figura 6.13 Exemplo de Histograma de Cores de Uma Componente (R, G ou B)

Passo 2: Em seguida, o RIST-Br calcula a média ponderada com relação ao número de pixels que estão distribuídos em cada uma das 256 tonalidades de cor da componente (isto é, a média ponderada com relação ao histograma da componente de cor) - o que é realizado através da seguinte equação:

$$\textit{Média do Histograma} = \frac{\sum_{i=0}^{n}\left(n \cdot S_n \right)}{n+1} \tag{6.1}$$

onde:
n é a tonalidade do pixel [0,255] em uma das componentes RGB;
S_n é o somatório dos pixels da subimagem que possuem a tonalidade n.

Passo 3: o RIST-Br encontra a tonalidade da componente de cor, entre as 256 possíveis, que possua o maior número de pixels da subimagem com relação às outras tonalidades - esse número máximo de pixels será o valor utilizado, juntamente com a Média do Histograma (Equação 6.1), pelo sistema RIST-Br (ver Figura 6.13).

Assim, seguindo os passos de um a três acima, para cada uma das componentes RGB dos pixels da subimagem serão obtidos dois valores: uma média ponderada (a Média do Histograma) e um máximo para a componente R (vermelho); uma média ponderada e um máximo para a componente G (verde); e uma média ponderada e um máximo para a componente B (azul);

Esses seis valores representativos para a cor da área correspondente à da subimagem candidata, juntamente com a miniatura 40x40 pixels, representam a subimagem candidata.

A descrição da subimagem candidata, por sua vez, exige um passo adicional, que leva em conta a miniatura de 40x40 pixels. É efetuado um somatório, linha a linha e coluna a coluna, dos pixels "ligados" (pretos) na subimagem candidata miniaturizada (40x40 pixels). Esses somatórios resultam em 80 valores (os 40 somatórios das colunas mais os 40 somatórios das linhas), que são usados em conjunto com os seis valores correspondentes às cores da subimagem candidata. Um exemplo desse somatório pode ser observado na Figura 6.14, abaixo.

Σ	**4**	**2**	**1**	**1**
3				
1				
2				
2				

Figura 6.14 Exemplo dos Somatórios dos Pixels de Cada Linha e Coluna de uma Imagem de 4x4 Pixels

Assim, temos os 86 valores que irão descrever cada uma das subimagens candidatas detectadas pelo mecanismo atencional do RIST-Br, e que serão as entradas a serem fornecidas para a RNA. Isso equivale a uma redução de 99,93% no volume de informações necessário para descrever a subimagem original, pois esse volume, que era de 120.000 valores (200 x 200 pixels x 3 bytes (um byte por cor)), passou a ser de apenas 86 valores.

6.2.4 O Reconhecimento dos Sinais de Trânsito nas Subimagens Candidatas

Para o reconhecimento dos sinais de trânsito nas subimagens candidatas (e, conseqüentemente, na própria imagem original), foi escolhida a técnica de aprendizado de máquinas que utiliza as redes neurais artificiais (cujos fundamentos são apresentados no Capítulo 4). O projeto adotado para a rede neural artificial que realizará o reconhecimento dos sinais de trânsito será apresentado nas seções 6.2.5.2 e 6.2.5.3, abaixo.

A rede neural artificial desenvolvida - conforme explicam os fundamentos do Capítulo 4 - tem a capacidade de armazenar, por meio de seus pesos sinápticos, informações suficientes para - dada uma certa entrada - responder de maneira que sua resposta fique o mais próximo possível da resposta ideal para a entrada fornecida.

Esse armazenamento das informações nos pesos sinápticos da rede é obtido através de um treinamento supervisionado, *off-line*, onde são apresentados para a rede neural (a cada época do aprendizado - ver seção 4.2.5.3, no Capítulo 4) todos os pares de padrões de entrada e padrões de saída - isto é, pares de vetores de entrada e vetores de saída. Essa apresentação é feita através de um algoritmo de aprendizado para a rede neural artificial. Neste trabalho, o algoritmo de aprendizado adotado foi o de propagação resiliente *(resilient propagation* - RPROP - apresentado na seção 4.2.5.2)

Nas próximas seções, serão apresentadas: as Classes de Sinais de Trânsito Brasileiros reconhecidas pelo RIST-Br; o tipo e a topologia da rede neural artificial projetada; e também como transcorreu o treinamento dessa rede neural artificial.

6.2.4.1 As Classes de Sinais de Trânsito Brasileiros Reconhecidas pelo RIST-Br

O Departamento Nacional de Estradas de Rodagem (DNER), cuja denominação atual é Departamento Nacional de Infra-Estrutura de Transportes (DNIT) , do Brasil, divide a sinalização de trânsito brasileira em seis classes: Placas de Advertência; Placas Educativas; Placas de Indicação; Placas de Obras; Placas de Regulamentação; e Placas de Serviços Auxiliares [DNIT 2004].

Abaixo, na Tabela 6.1, podem ser observadas algumas ilustrações que exemplificam cada uma dessas classes (as ilustrações não estão em suas proporções reais).

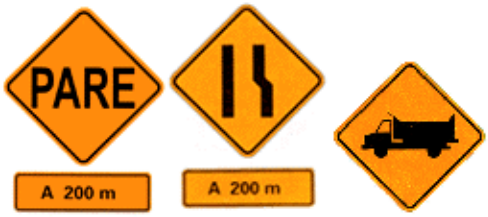
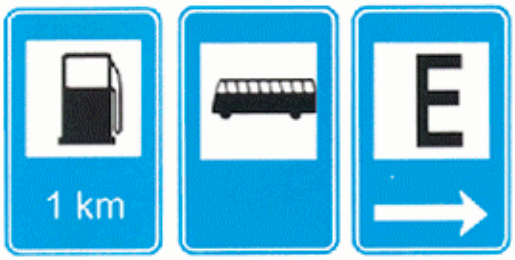
Classe	Exemplos
Placas de Advertência	
Placas Educativas	
Placas de Indicação	
Placas de Obras	
Placas de Regulamentação	
Placas de Serviços Auxiliares	

Tabela 6.2 Exemplos de Sinais de Cada Classe de Sinais de Trânsito Brasileiros

A rede neural artificial (RNA), do RIST-Br, projetada para a fase de reconhecimento das subimagens candidatas, foi treinada para reconhecer os sinais mostrados nas ilustrações da Figura 6.15, abaixo.

(a) Área Escolar (b) Parada Obrigatória (c) Sentido Obrigatório

Figura 6.15 Tipos de Sinais de Trânsito Utilizados para Treinamento da RNA

Os sinais apresentados na Figura 6.15 foram selecionados pelo autor devido à freqüência com que podem ser encontrados nos centros urbanos brasileiros.

Embora os sinais (b) e (c), da Figura 6.14, pertençam à classe das Placas de Regulamentação, ambos foram tratados, desde o pré-processamento, como sinais de classes diferentes - isto é, foi criada uma classe especial, para o RIST-Br, específica para o sinal de Parada Obrigatória. Tal procedimento foi adotado devido às diferenças visuais e de forma entre o sinal de Parada Obrigatória e os demais sinais da classe de Placas de Regulamentação (os quais seguem o padrão de um contorno representado por um anel vermelho, com um fundo branco sobre o qual está a imagem regulamentadora, em preto).

Já o sinal de Área Escolar, tal como ocorre como o sinal de Sentido Obrigatório (para a classe de Placas de Regulamentação), é representativo da classe de Placas de Advertência.

Portanto, embora o RIST-Br tenha sido projetado para reconhecer especificamente os sinais de Área Escolar, de Parada Obrigatória, e de Sentido Obrigatório (Figura 6.15 (a), (b) e (c), respectivamente), o RIST-Br é capaz de reconhecer também os demais sinais pertencentes às classes das Placas de Advertência e das Placas de Regulamentação (observar Tabela 6.1).

6.2.4.2 O Tipo e a Topologia da RNA

A rede neural artificial (RNA) projetada para o RIST-Br é um perceptron de múltiplas camadas totalmente conectado (apresentado na seção 4.2.5 do Capítulo 4, e ilustrado pela Figura 4.5 do mesmo capítulo).

Essa RNA foi inicialmente projetada com 86 neurônios na camada de entrada, 22 neurônios na única camada oculta, e 4 neurônios na camada de saída (usarei, neste trabalho, a notação 86:22:4 para descrever a topologia da RNA, conforme as quantidades de neurônios em sua camada de entrada, em sua camada oculta, e em sua camada de saída, respectivamente). Porém, sua topologia foi modificada, e a RNA final apresenta 172 neurônios na camada oculta (86:172:4). Essa ampliação do número de neurônios na camada oculta teve como objetivo melhorar a capacidade de aprendizado da RNA [Gori 1998].

Os 86 neurônios da camada de entrada receberão os 86 valores definidos na seção 6.2.4, acima, que descrevem uma subimagem candidata com a finalidade de submetê-la ao reconhecimento da RNA. Os 4 neurônios da camada de saída fornecerão, cada um, um valor no intervalo $[0,1] \in \Re$.

Dado que cada neurônio da camada de saída representa uma das classes projetadas pelo autor - (i) Não-Sinal; (ii) Sinal de Parada Obrigatória; (iii) Sinal de Área Escolar; (iv) Sinal de Sentido Obrigatório - o neurônio com o valor de resposta maior que o valor apresentado pelos demais indicará a qual sinal de trânsito corresponde a subimagem candidata apresentada para a RNA.

Cabe lembrar, conforme exposto na seção anterior, que os sinais de Área Escolar e de Sentido Obrigatório podem ser substituídos por quaisquer outros de suas respectivas classes (Placas de Advertência e Placas de Regulamentação (exceto a Parada Obrigatória), respectivamente).

6.2.4.3 O Treinamento da RNA

Para o treinamento da rede neural artificial do RIST-Br, foi utilizado o simulador SNNS, apresentado na seção 4.3 do Capítulo 4, através da ferramenta "Batchman", que permite a criação de um roteiro (a ser executado pelo SNNS) para o treinamento da RNA. Conforme mencionado, o algoritmo de aprendizado utilizado foi o RPROP. A rede foi treinada de maneira off-line, utilizando a técnica de validação cruzada (seção 4.2.5.3 do Capítulo 4), a qual exige dois conjuntos de dados: um conjunto de dados de treinamento, e outro de dados de validação.

Os conjuntos de dados, de treinamento e validação, são representados por arquivos texto (no formato padrão para o SNNS) contendo os padrões de entrada (os 86 valores descritores para cada subimagem candidata) e os padrões de saída para cada entrada. Os padrões de saída correspondem a quadro valores (um para cada neurônio da camada de saída), cada um dos quais pertencentes ao conjunto {0,1}. Por exemplo: para um padrão de entrada correspondente a uma subimagem candidata "Não-Placa", o padrão de saída será "1 0 0 0".

A utilização desses pares (padrão de entrada, e padrão de saída esperada) correspondem à técnica do treinamento supervisionado (apresentado na seção 4.2.4, do Capítulo 4).

A primeira rede neural artificial (RNA) projetada, com 22 neurônios na única camada oculta (o valor 22 foi escolhido por ser aproximadamente igual ao número de neurônios da camada de entrada dividido por quatro), foi treinada com um conjunto de treinamento de 392 padrões (2/3 dos dados, dos quais 300 padrões correspondiam à classe "Não-Placa";) e um conjunto de validação de 196 padrões (1/3 dos dados, dos quais 139 padrões correspondiam à classe "Não-Placa"). Os padrões para os sinais a serem reconhecidos foram inseridos de maneira aleatória a esses conjuntos - a fim de facilitar o aprendizado da

RNA [Tafner 1996], porém as quantidades de padrões para cada sinal não foi controlada pelo autor.

Os resultados para o treinamento dessa primeira RNA (que serão apresentados na seção 6.4, abaixo), nessas condições, mostram que, embora a proporção de "Não-Placas" nos conjuntos de treinamento e validação tenha enfraquecido a resposta da RNA aos padrões de sinais de trânsito, o sistema RIST-Br ainda foi capaz de interpretar as respostas (saídas) da RNA. Essa interpretação foi possível porque o RIST-Br verifica apenas o valor do neurônio correspondente ao sinal de trânsito selecionado pelo usuário para a busca na imagem. Como todas as saídas dos neurônios correspondentes aos sinais de trânsito ficaram com valores muito baixos, com relação à saída do neurônio "Não-Placa", o RIST-Br foi ajustado para que uma saída (de um dos outros três neurônios) acima de um limiar de 0,10 (de um intervalo $[0,1] \in \Re$) fosse interpretada como sendo o respectivo sinal de trânsito.

A segunda rede neural artificial (RNA) projetada (e adotada para a versão final do RIST-Br) com 172 neurônios na única camada oculta (o valor 172 foi escolhido por ser duas vezes maior do que o número de neurônios da camada de entrada, e também devido às colocações de [Gori 1998]), foi treinada com um conjunto de treinamento de 100 padrões e um conjunto de validação de 25 padrões. Os padrões para os sinais a serem reconhecidos foram inseridos de maneira ordenada (como será explicado no próximo parágrafo) a esses conjuntos - a fim de facilitar o aprendizado da RNA [Tafner 1996], porém as quantidades de padrões para cada sinal não foi controlada pelo autor.

A maneira escolhida para ordenar os padrões nos arquivos de treinamento e validação foi a de colocar os padrões na seguinte ordem: sinal de Área Escolar; sinal de Parada Obrigatória; sinal de Sentido Obrigatório; Não-Placa; Não-Placa (repetição). Essa ordem foi repetida ao longo dos dois arquivos, até que o total de padrões de treinamento e validação fosse atingido. Assim, temos:

Padrões	**Quantidade de Padrões**	
	Arquivo de Treinamento	**Arquivo de Validação**
Não-Placa	40	10
Parada Obrigatória	20	5
Área Escolar	20	5
Sentido Obrigatório	20	5
TOTAL	**100**	**25**

Tabela 6.3 Dados dos Arquivos de Treinamento e Validação para a RNA 86:172:4

Na Figura 6.16, abaixo, pode ser visualizado um trecho do arquivo de treinamento de 100 padrões, para leitura pelo SNNS.

```
SNNS pattern definition file V4.2
generated at Tue Apr 27 15:00:00 2004


No. of patterns : 100
No. of input units : 86
No. of output units : 4



# 001
# Somatorios das linhas do minibloco 40x40 numero 1 e272:
1 0 0 1 2 0 2 1 3 1 1 2 2 3 7 6 9 11 9 12 8 13 18 19 11 12 11 16 14 13 11 14 18 17 18 16 18
18 21 0

# Somatorios das colunas do minibloco 40x40 numero 1:
0 1 0 1 2 1 4 2 0 1 3 1 0 1 3 6 5 7 9 13 14 16 15 17 16 14 15 14 16 14 16 15 14 19 20 18 17
11 18 0
# Valores de Histograma para o bloco RGB 200x200 [gravado como ".txt"] numero 1:

# Media e Maximo para o Vermelho:
48 2510
# Media e Maximo para o Verde:
51 1816
# Media e Maximo para o Azul:
38 5024
# SAIDAS (apenas uma deve estar ativa!):

# NAO-PLACA / PARADA OBRIGATORIA / AREA ESCOLAR / SENTIDO OBRIGATORIO.
0 0 1 0

# ------------------------------------------------------------


# 002
# Somatorios das linhas do minibloco 40x40 numero 11 p271:
1 3 3 4 2 0 3 1 2 1 2 3 0 1 12 15 17 18 18 18 19 20 21 20 13 8 10 11 9 10 9 10 9 11 17 18 18
19 20 0

# Somatorios das colunas do minibloco 40x40 numero 11:
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 4 2 14 14 8 17 18 16 18 24 19 16 19 14 19 22 19 16 16 18 18 18
21 26 0
# Valores de Histograma para o bloco RGB 200x200 [gravado como ".txt"] numero 11:

# Media e Maximo para o Vermelho:
178 644
# Media e Maximo para o Verde:
184 970
# Media e Maximo para o Azul:
175 978
# SAIDAS (apenas uma deve estar ativa!):

# NAO-PLACA / PARADA OBRIGATORIA / AREA ESCOLAR / SENTIDO OBRIGATORIO.
0 1 0 0

# ------------------------------------------------------------
```

Figura 6.16Trecho do Arquivo de Padrões de Treinamento para a RNA 86:172:4

Ambas as redes neurais artificiais (tanto a RNA 86:22:4 quanto a RNA 86:172:4) foram treinadas de acordo com as seguintes informações fornecidas no roteiro para o "Batchman":

- Inicialização randômica dos pesos sinápticos (e do *bias*) da rede, no intervalo [α,β] (onde, pela sintaxe do roteiro, α>β), com α = 1,0 e β = -1,0;
- Parâmetros para o algoritmo RPROP iguais a:
  - Parâmetro Δ_0 = 1 (valor inicial para todos os valores de atualização (Δ_{ij}) dos pesos);
  - Parâmetro Δ_{max} = 30 (limite máximo para os valores de atualização dos pesos);
  - Parâmetro α = 1 (decaimento dos pesos - este parâmetro foi inserido no algoritmo do RPROP pelos desenvolvedores do SNNS. O decaimento dos pesos determina a relação entre os dois seguintes objetivos: a redução do erro na saída da rede; e a redução na magnitude dos pesos [SNNS 1998]).
- O treinamento deveria continuar até que o Erro Médio Quadrado (*Mean Square Error* - MSE - ver seção 4.2.5.3, do Capítulo 4) fosse maior que 0,001 ou enquanto o número de épocas de treinamento transcorridas fosse menor que 2000.

A Figura 6.17, abaixo, ilustra o treinamento efetuado com as duas redes neurais artificiais, e também com os conjuntos de padrões de treinamento e validação. Foram efetuados os seguintes treinamentos.

Treinamento 1: RNA 86:22:4 com conjunto de treinamento de 392 padrões e conjunto de validação de 196 padrões (total de padrões diferentes: 588).

Treinamento 2: RNA 86:22:4 com conjunto de treinamento de 100 padrões e conjunto de validação de 25 padrões (total de padrões diferentes: 125). Todos estes padrões são diferentes daqueles usados no Treinamento 1.

Treinamento 3: RNA 86:172:4 com conjunto de treinamento de 100 padrões e conjunto de validação de 25 padrões (total de padrões diferentes: 125). Todos estes padrões são diferentes daqueles usados no Treinamento 1, e são exatamente os mesmos do Treinamento 2.

As curvas do erro médio quadrado (Mean Square Error - MSE) podem ser observadas, para cada um dos treinamentos, nas figuras abaixo.

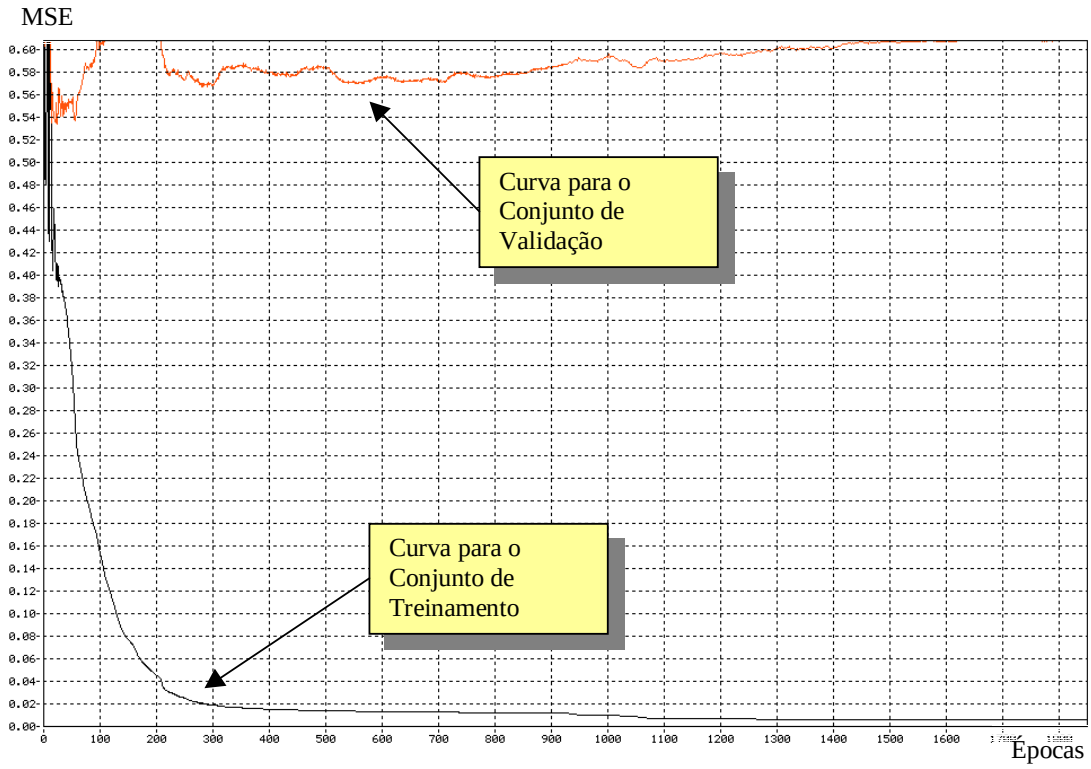


Figura 6.17 Curvas (MSE x Épocas) para o Treinamento 1

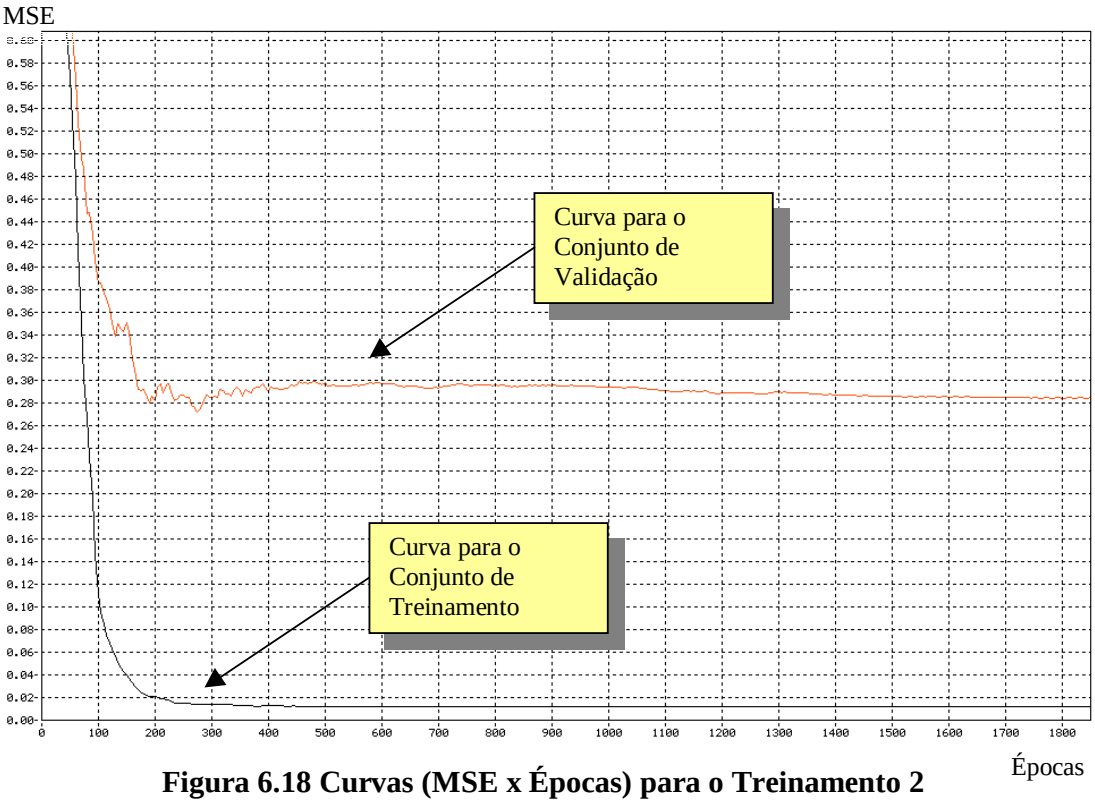


Figura 6.18 Curvas (MSE x Épocas) para o Treinamento 2

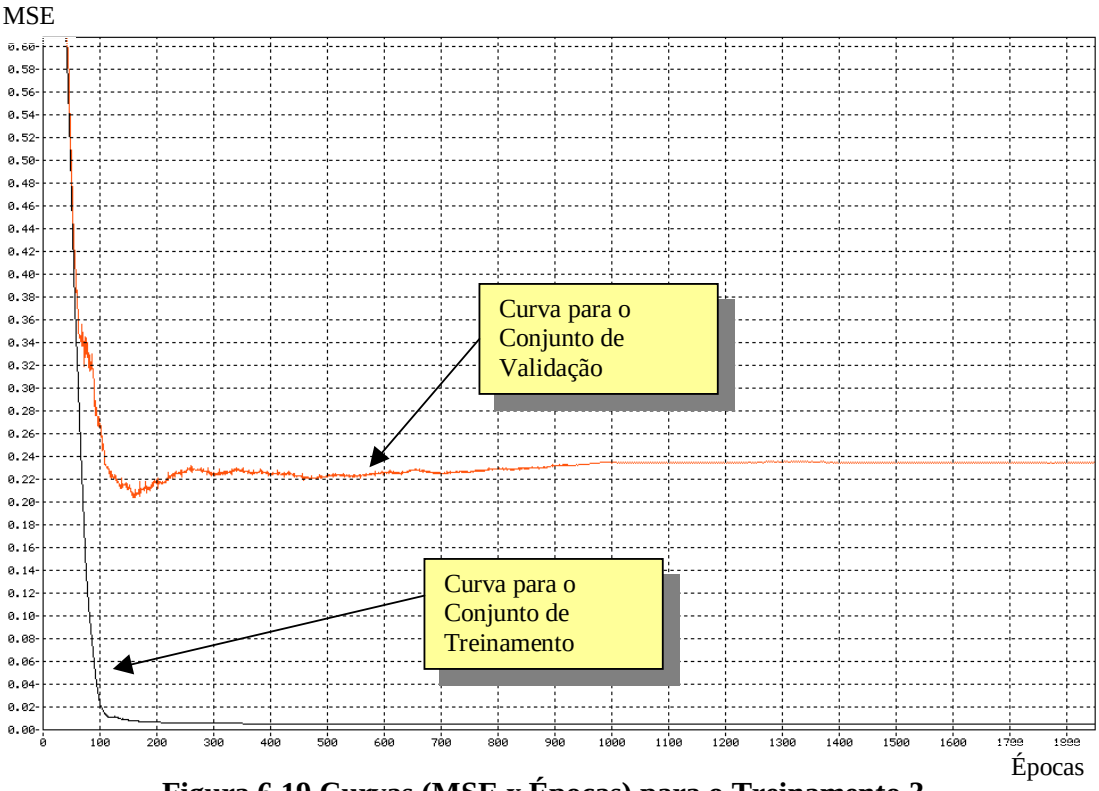


Figura 6.19 Curvas (MSE x Épocas) para o Treinamento 3

Abaixo, na Figura 6.20, podem ser vistas, lado a lado e numa escala menor, as curvas de aprendizagem dos Treinamentos 2 e 3.

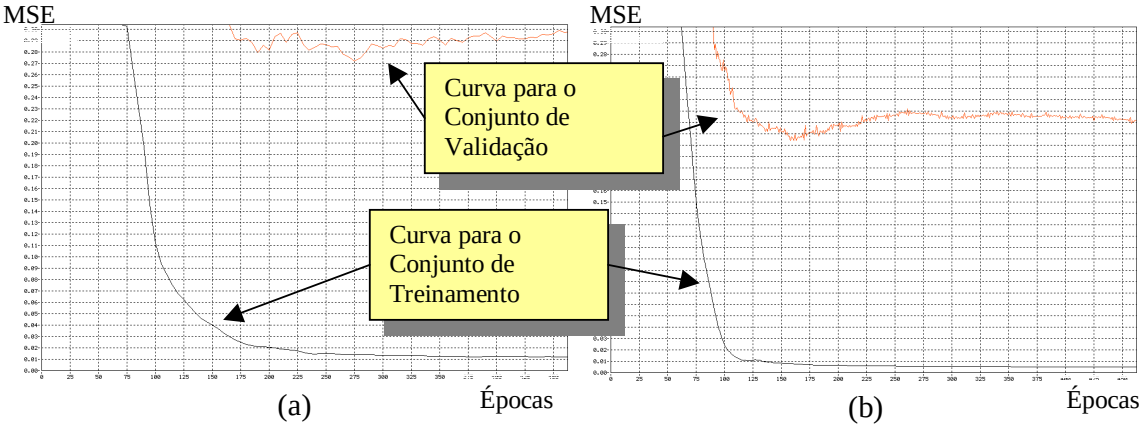


Figura 6.20 Comparação da Aprendizagem: (a) Treinamento 2; e (b) Treinamento 3

A RNA artificial resultante do Treinamento 3 foi a escolhida para o sistema RIST-Br, pois apresentou um aprendizado com menor erro do que as demais, o que é justificado também pela curva de erro dos dados de validação - que também mostra valores de erros mais baixos que as curvas de erro dos dados de validação dos outros treinamentos. A Figura 6.21, abaixo, mostra uma ampliação da Figura 6.20 (b).

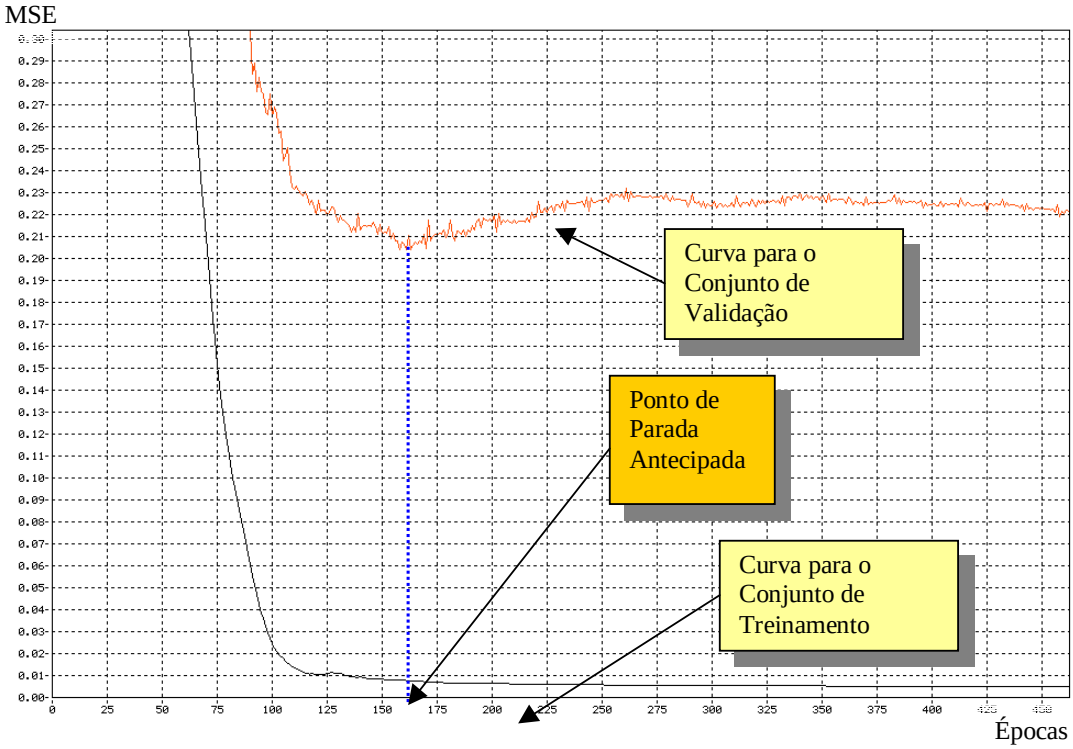


Figura 6.21 Curvas de Aprendizado do Treinamento 3

O ponto de parada antecipada do treinamento - que é a época do treinamento definida através da heurística do ponto mínimo da curva de erro dos dados de validação (ver seção 4.2.5.3, do Capítulo 4) - para o Treinamento 3 seria aproximadamente a época de número 162.

A Matriz de Confusão para os dados do conjunto de validação do Treinamento 3, apresentada na Tabela 6.3 abaixo, permite afirmar que a RNA 86:172:4, escolhida para ser utilizada pelo RIST-Br, conseguiu reconhecer 100% das vezes cada um dos diferentes sinais de trânsito dentre os 25 padrões apresentados durante a validação. Também podemos afirmar que a RNA 86:172:4 identificou corretamente, 100% das vezes todos os padrões "Não-Placa", apresentados entre os 25 padrões do conjunto de validação. A RNA também

foi capaz de identificar, 100% das vezes, quando um padrão não pertencia a uma determinada classe.

Matriz de Confusão Treinamento 3 RNA 86:172:4 n = 25 padrões (validação)			Previstas							
			Não-Placa		Parada Obrigatória		Área Escolar		Sentido Obrigatório	
			SIM	NÃO	SIM	NÃO	SIM	NÃO	SIM	NÃO
Encontradas pela RNA	Não-Placa	SIM	**10**	0	0	0	0	0	0	0
		NÃO	0	**15**	0	0	0	0	0	0
	Parada Obrigatória	SIM	0	0	**5**	0	0	0	0	0
		NÃO	0	0	0	**20**	0	0	0	0
	Área Escolar	SIM	0	0	0	0	**5**	0	0	0
		NÃO	0	0	0	0	0	**20**	0	0
	Sentido Obrigatório	SIM	0	0	0	0	0	0	**5**	0
		NÃO	0	0	0	0	0	0	0	**20**

Tabela 6.4 Matriz de Confusão do Conjunto de Validação do Treinamento 3

A Matriz de Confusão foi construída com base nos resultados do Treinamento 3. Dois exemplos para a construção da Matriz de Confusão da Tabela 6.3 podem ser vistos a seguir:

(i) se foi previsto que o padrão era SIM uma Área Escolar, e foi encontrado pela RNA que o padrão era SIM uma Área Escolar, então adicionamos 1 (um) ao valor que está na intersecção dessas duas variáveis, da tabela.

(ii) se foi previsto que o padrão NÃO era uma Não-Placa, e foi encontrado pela RNA que o padrão NÃO era uma Não-Placa, então adicionamos 1 (um) ao valor que está na intersecção dessas duas variáveis, da tabela. Para este exemplo, na Tabela 6.3, podemos

perceber que a RNA identificou, 100% das vezes, todos os padrões que correspondiam a uma Não-Placa (10 padrões Não-Placa), e também identificou corretamente, 100% das vezes, todos os padrões que NÃO correspondiam a uma Não-Placa (que eram os 15 sinais de trânsito presentes entre os 25 padrões de validação).

6.2.4.4 A Visualização do Reconhecimento do RIST-Br

Finalmente, após todos os passos - da captura da imagem digital original (seção 6.2.1) até o reconhecimento, do qual trata esta seção 6.2.5 - podemos visualizar sobre a imagem original as regiões correspondentes àquelas subimagens candidatas (a serem um sinal de trânsito) que foram reconhecidas pela RNA do RIST-Br como sendo um sinal de trânsito.

Essa visualização pode ser observada na Figura 6.22, abaixo.

Figura 6.22 Visualização do Resultado do Reconhecimento de Sinais de Trânsito Pelo RIST-Br (Reconhecimento de um Sinal de Parada Obrigatória)

A Figura 6.23, abaixo, mostra uma ampliação da região onde foi reconhecido o sinal de Parada Obrigatória da Figura 6.22.

Figura 6.23 Ampliação da Região onde foi Reconhecido o Sinal de Parada Obrigatória

Na Figura 6.23, acima, pode ser observado que, das 27 subimagens candidatas (Figura 6.11), apenas as três subimagens que aparecem marcadas na Figura 6.23 foram reconhecidas como um sinal de Parada Obrigatória. Como as subimagens reconhecidas realmente correspondem a um sinal de Parada Obrigatória (podemos verificar isso visualmente), isso demonstra o sucesso do reconhecimento para esse exemplo.

6.3 A Implementação do RIST-Br

Esta seção apresentará uma visão geral do funcionamento e relacionamento dos três programas, ou módulos, implementados para o RIST-Br, uma vez que seu projeto já foi detalhado na seção 6.2 deste Capítulo.

A implementação computacional do RIST-Br foi efetuada na linguagem de programação C [Kernighan 1986]. Para a geração do código executável (ou do código objeto, como ocorre na seção 6.3.2, logo adiante), foi utilizado o compilador GCC [GCC 2004] sob o sistema operacional Windows^{XP} Home Edition [Microsoft 2004]. O microcomputador utilizado para a implementação do RIST-Br, assim como para a execução de todos os testes sobre as imagens da base construída (e também para a redação desta monografia) - possui a seguinte configuração de microprocessamento e memória: microprocessador Pentium 4 [Intel 2004] - freqüência de *clock* de 2,80 GHz; memória primária (RAM) DDR de 512 MB (megabytes), com freqüência de 400 MHz; e memória secundária (disco magnético rígido) de 80 GB (gigabytes), com velocidade de 7200 rotações por minuto.

A interface com o sistema RIST-Br é feita através da linha de comando do prompt do MS-DOS, de onde são executados os módulos apresentados a seguir.

6.1.1 O Módulo de Localização de Subimagens Candidatas

O módulo de localização de subimagens candidatas (módulo de detecção, ou módulo do mecanismo atencional) é responsavel pelos passos de: pré-processamento da imagem de entrada; segmentação da imagem de entrada; e representação, e descrição, das subimagens candidatas (a serem um sinal de trânsito) que serão submetidas ao módulo de identificação (reconhecimento).

Este módulo recebe como entrada duas imagems: a imagem original, capturada através da câmera digital (Figura 6.7 (a)); e a imagem em tons de cinza correspondente à componente H do modelo de cor HSI (Figura 6.7 (b)).

Como saída, este módulo gera: (i) para cada subimagem candidata é gerado e gravado na memória secundária: um arquivo no formato texto com as informações sobre a cor (modelo RGB) de todos os pixels da subimagem candidata; uma imagem no formato BMP (mapa de bits - profundidade de cor de 1 bit, isto é, apresenta apenas pixels pretos e brancos) de 200x200 pixels, que corresponde à subimagem candidata (Figura 6.10 (a)); e uma imagem no formato BMP de 40x40 pixels, que corresponde à miniatura da subimagem candidata

(Figura 6.10 (b)); (ii) um arquivo no formato texto contendo a descrição (que servirá como entrada para a RNA) de cada subimagem candidata, juntamente com as respectivas coordenadas do local onde a subimagem localiza-se na imagem original; (iii) uma imagem no formato BMP (Figura 6.9) que mostra o resultado da binarização da imagem que representa a componente H do modelo HSI (essa imagem foi utilizada para a busca de subimagens candidatas); e (iv) a imagem que mostra a localização das subimagens candidatas, isto é, a imagem que apresenta o resultado do mecanismo atencional do RIST-Br (Figura 6.11).

6.3.1 O Módulo de Identificação dos Sinais de Trânsito

O módulo de identificação dos sinais de trânsito é responsável por enviar para RNA as descrições de cada uma das subimagens candidatas, e de receber as respectivas respostas da RNA. Dessa maneira, este módulo é responsável pelo reconhecimento dos sinais de trânsito nas subimagens candidatas. Para a inclusão da RNA (criada no SNNS) neste módulo de identificação, foi utilizada a ferramenta "snns2c", que acompanha o SNNS. Essa ferramenta permite gerar, a partir do arquivo da RNA criado pelo SNNS, um arquivo de código fonte em linguagem C que implementa a RNA. Esse código fonte da RNA, em C, é compilado em um arquivo objeto pelo compilador GCC, e posteriormente ligado ao arquivo executável do módulo de identificação durante a compilação deste último.

O módulo de identificação recebe como entrada um arquivo texto (gerado pelo módulo de localização descrito na seção anterior) contento as descrições de cada subimagem candidata, juntamente com as respectivas coordenadas do local onde a subimagem localiza-se na imagem original. Como mencionado na seção 6.2.4, apenas as descrições (os 86 valores de entrada, que descrevem a subimagem candidata e que correspondem ao padrão de entrada para a RNA) de uma subimagem são fornecidas como entradas para a RNA.

Como saída, o módulo de identificação gera um arquivo no formato texto contendo os quatro valores de saída da RNA para cada uma das descrições das subimagens candidatas, juntamente com as respectivas coordenadas do local onde a subimagem localiza-se na imagem original.

6.3.2 O Módulo de Visualização dos Resultados

O módulo de visualização dos resultados é responsável por indicar, sobre a imagem original, as áreas da mesma onde foram encontradas subimagens candidatas cujo módulo de identificação confirmou a presença de um sinal de trânsito.

Este módulo recebe como entrada: o arquivo texto com os valores de saída da RNA para cada uma das subimagens, e também as coordenadas de cada subimagem (gerado pelo módulo de identificação, da seção anterior); e o arquivo, em formato JPEG, da imagem original.

Um limiar, definido no próprio programa (por exemplo, 0,900), servirá para o módulo de visualização decidir se a resposta de um dos quatro neurônios da camada de saída da RNA para determinada subimagem candidata será indicada sobre as respectivas coordenadas (posição) na imagem original, na forma de uma marcação (como na Figura 6.22).

Por exemplo, se o usuário está procurando pelo sinal de Parada Obrigatória, e o neurônio de saída correspondente à essa classe apresentou um valor de 0,053 , então - como o limiar vale 0,900 (e 0,053 < 0,900) - a subimagem candidata não é considerada um sinal de Parada Obrigatória, e assim não é mostrada sobre a imagem original. Ao contrário, se o mesmo neurônio apresentou como saída um valor de 0,974 , então a subimagem candidata é considerada um sinal de Parada Obrigatória (pois 0,974 ≥ 0,900) e é marcada sobre a imagem original (como na Figura 6.22).

Como pode ser observado, apenas a resposta do neurônio correspondente ao sinal procurado pelo usuário é analisada. Tal procedimento de separação das respostas de cada um dos neurônios da camada de saída foi adotado devido ao fato de existir a diferença para a segmentação da imagem original - no caso do sinal ser de Sentido Obrigatório (ver seção 6.2.3).

Essa diferença para a segmentação diz respeito ao seguinte: se fosse utilizado o mesmo limiar para a segmentação - dez mil pixels sob a "cabeça de leitura" (como é feito para os sinais de Parada Obrigatória e Área Escolar) - em busca dos sinais de Sentido Obrigatório, os mesmos não seriam encontrados, pois o total de pixels que compõem a borda vermelha desse sinal seria insuficiente para superar, ou igualar-se a, esse limiar de dez mil pixels. Adicionalmente, quando se está procurando pelo sinal de Área Escolar, ainda há a diferença na cor que está sendo procurada durante a binarização da imagem (procura-se a cor amarela, ao invés da cor vermelha). Assim, optou-se por separar todos os passos para cada sinal de trânsito (desde o pré-processamento até o reconhecimento) a ser procurado na imagem conforme o desejo do usuário.

Finalmente, a saída gerada pelo módulo de visualização é um arquivo com a imagem original, em formato JPEG, mostrando as áreas das subimagens candidatas que foram reconhecidas como sendo o sinal de trânsito procurado (Figura 6.22).

6.3.3 Visão Geral do RIST-Br

Esta seção apresenta a Figura 6.22, abaixo, onde pode ser observada uma visão geral da implementação computacional do RIST-Br.

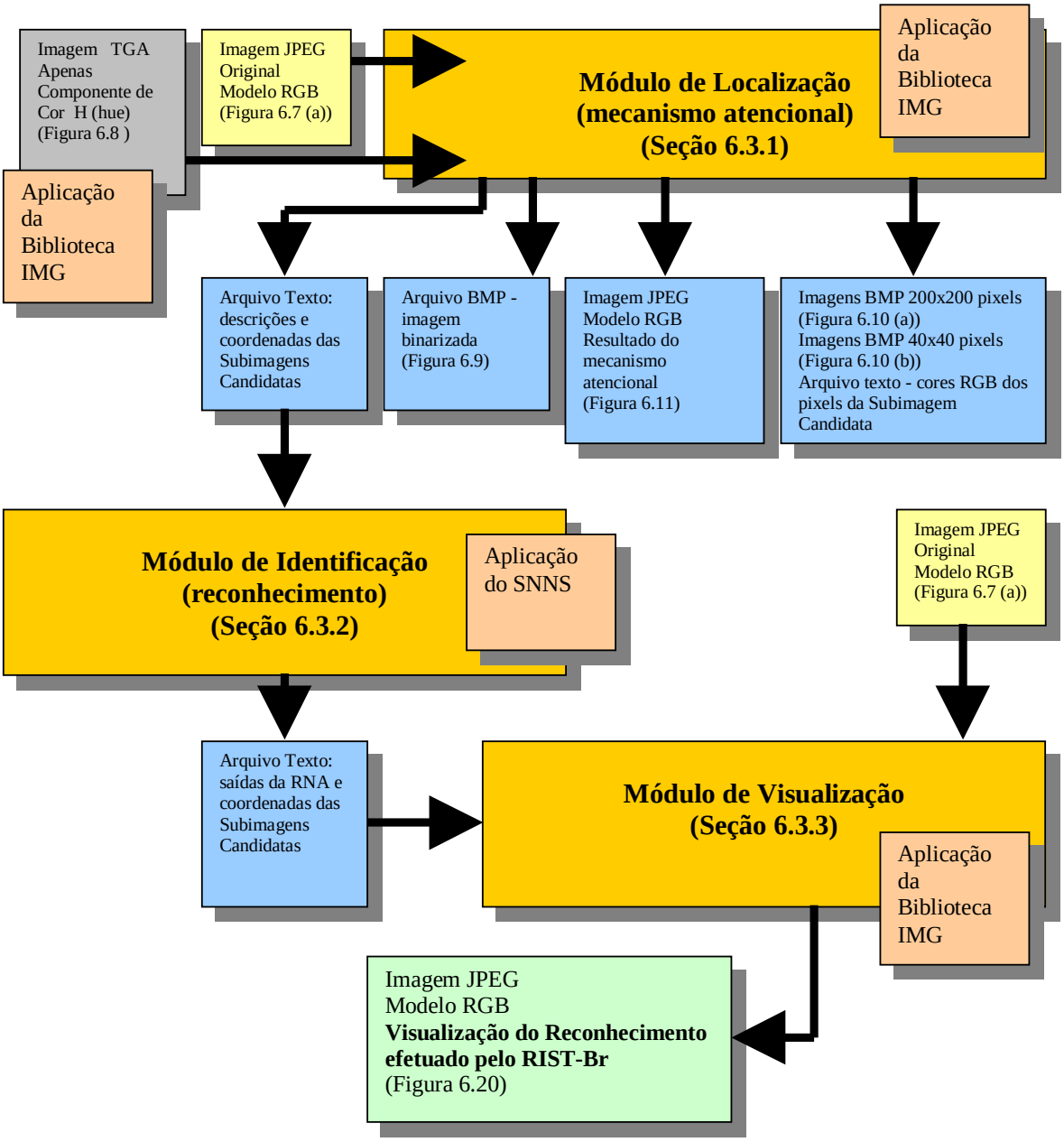


Figura 6.24 Visão Geral do RIST-Br

6.4 Resultados Obtidos

Nesta seção, serão apresentados os resultados obtidos pelo RIST-Br sobre a base de 148 imagens construída para este trabalho. Também serão apresentadas algumas imagens de exemplo dos testes efetuados.

Nas tabelas apresentadas abaixo, "Verdadeiro Positivo" (VP) indica que um sinal foi reconhecido corretamente; "Falso Positivo" (FP) indica que uma subimagem candidata foi reconhecida incorretamente como sendo um sinal de trânsito.

Para o mecanismo atencional, é considerado "Sinal Detectado" se ao menos uma subimagem candidata coincide com a região onde localiza-se o sinal de trânsito na imagem - essa verificação foi efetuada visualmente pelo autor.

Para o reconhecimento, é considerado:

- "Apenas Verdadeiro Positivo" - se apenas subimagens correspondentes ao sinal de trânsito foram reconhecidas;
- "Verdadeiro Positivo e Falso Positivo" - se, além das subimagens correspondentes ao sinal de trânsito reconhecido, outras subimagens forem incorretamente reconhecidas incorretamente, pelo RIST-Br, como sendo um sinal de trânsito;
- "Apenas Falso Positivo" - se apenas subimagens incorretamente reconhecidas pelo RIST-Br como sendo sinais de trânsito são assinaladas na imagem;
- "Sem Reconhecimento" - se o RIST-Br não encontrou subimagem alguma correspondente a um sinal de trânsito (não ocorreram Verdadeiros Positivos e nem Falsos Positivos entre as subimagens candidatas, ou ainda não foram detectadas subimagens candidatas pelo mecanismo atencional).

	Mecanismo Atencional (Localização de Subimagens Candidatas)		
Tipos de Sinais	**Total de Imagens**	**Quantidade de Imagens - Sinal Detectado**	**Porcentagem de Detecção**
Parada Obrigatória	33	19	57,6%
Área Escolar	24	19	79,2%
Sentido Obrigatório	53	34	64,2%
Velocidade Máxima	23	17	73,9%
Outras Placas	15	10	66,7%
TOTAL	148	99	66,9%

	Reconhecimento (Identificação dos Sinais)		
Tipos de Sinais	**Total de Imagens**	**Quantidade de Imagens - Apenas Verdadeiro Positivo**	**Porcentagem de Apenas Verdadeiro Positivo**
Parada Obrigatória	33	10	30,3%
Área Escolar	24	11	45,8%
Sentido Obrigatório	53	2	3,8%
Velocidade Máxima	23	0	0,0%
Outras Placas	15	1	6,7%
TOTAL	148	24	16,2%
	Total de Imagens	**Quantidade de Imagens - Verdadeiro Positivo e Falso Positivo**	**Porcentagem de Verdadeiro Positivo e Falso Positivo**
Parada Obrigatória	33	7	21,2%
Área Escolar	24	8	33,3%
Sentido Obrigatório	53	7	13,2%
Velocidade Máxima	23	2	8,7%
Outras Placas	15	3	20,0%
TOTAL	148	27	18,2%
	Total de Imagens	**Quantidade de Imagens - Apenas Falso Positivo**	**Porcentagem de Apenas Falso Positivo**
Parada Obrigatória	33	6	18,2%
Área Escolar	24	1	4,2%
Sentido Obrigatório	53	32	60,4%
Velocidade Máxima	23	3	13,0%
Outras Placas	15	7	46,7%
TOTAL	148	49	33,1%
	Total de Imagens	**Quantidade de Imagens - Sem Reconhecimento**	**Porcentagem Sem Reconhecimento**
Parada Obrigatória	33	10	30,3%
Área Escolar	24	4	16,7%
Sentido Obrigatório	53	12	22,6%
Velocidade Máxima	23	18	78,3%
Outras Placas	15	4	26,7%
TOTAL	148	48	32,4%

Tabela 6.5 Resultados do RIST-Br: RNA 86:22:4, Treinamento 1 (ver Seção 6.2.5.3), limiar para a visualização de 0,10 (ver seções 6.2.5.3 e 6.3.3)

Limiar para Identificação = 0,10	Parada Obrigatória	Área Escolar
Apenas VP	30,3%	45,8%
VP e FP	21,2%	33,3%
TOTAL	**51,5%**	**79,2%**

Limiar para Identificação = 0,10	Sentido Obrigatório	Velocidade Máxima
Apenas VP	3,8%	0,0%
VP e FP	13,2%	8,7%
TOTAL	**17,0%**	**8,7%**

Limiar para Identificação = 0,10	Outras Placas	TOTAL RIST-Br
Apenas VP	6,7%	**16,2%**
VP e FP	20,0%	**18,2%**
TOTAL	**26,7%**	**34,5%**

Limiar para Identificação = 0,10	Parada Obrigatória	Área Escolar
Apenas FP	18,2%	4,2%
Sem Reconhecimento	30,3%	16,7%
TOTAL	**48,5%**	**20,8%**

Limiar para Identificação = 0,10	Sentido Obrigatório	Velocidade Máxima
Apenas FP	60,4%	13,0%
Sem Reconhecimento	22,6%	78,3%
TOTAL	**83,0%**	**91,3%**

Limiar para Identificação = 0,10	Outras Placas	TOTAL RIST-Br
Apenas FP	46,7%	**33,1%**
Sem Reconhecimento	26,7%	**32,4%**
TOTAL	**73,3%**	**65,5%**

Tabela 6.6 Resultados Globais do RIST-Br: RNA 86:22:4, Treinamento 1 (ver Seção 6.2.5.3), limiar para a visualização de 0,10 (ver seções 6.2.5.3 e 6.3.3)

	Mecanismo Atencional (Localização de Subimagens Candidatas)		
Tipos de Sinais	**Total de Imagens**	**Quantidade de Imagens - Sinal Detectado**	**Porcentagem de Detecção**
Parada Obrigatória	33	19	57,6%
Área Escolar	24	19	79,2%
Sentido Obrigatório	53	51	96,2%
Velocidade Máxima	23	21	91,3%
Outras Placas	15	14	93,3%
TOTAL	148	124	83,8%

	Reconhecimento (Identificação dos Sinais)		
Tipos de Sinais	**Total de Imagens**	**Quantidade de Imagens - Apenas Verdadeiro Positivo**	**Porcentagem de Apenas Verdadeiro Positivo**
Parada Obrigatória	33	11	33,3%
Área Escolar	24	12	50,0%
Sentido Obrigatório	53	2	3,8%
Velocidade Máxima	23	3	13,0%
Outras Placas	15	3	20,0%
TOTAL	148	31	20,9%
	Total de Imagens	**Quantidade de Imagens - Verdadeiro Positivo e Falso Positivo**	**Porcentagem de Verdadeiro Positivo e Falso Positivo**
Parada Obrigatória	33	5	15,2%
Área Escolar	24	1	4,2%
Sentido Obrigatório	53	14	26,4%
Velocidade Máxima	23	5	21,7%
Outras Placas	15	3	20,0%
TOTAL	148	28	18,9%
	Total de Imagens	**Quantidade de Imagens - Apenas Falso Positivo**	**Porcentagem de Apenas Falso Positivo**
Parada Obrigatória	33	3	9,1%
Área Escolar	24	0	0,0%
Sentido Obrigatório	53	28	52,8%
Velocidade Máxima	23	3	13,0%
Outras Placas	15	4	26,7%
TOTAL	148	38	25,7%
	Total de Imagens	**Quantidade de Imagens - Sem Reconhecimento**	**Porcentagem Sem Reconhecimento**
Parada Obrigatória	33	14	42,4%
Área Escolar	24	11	45,8%
Sentido Obrigatório	53	9	17,0%
Velocidade Máxima	23	12	52,2%
Outras Placas	15	5	33,3%
TOTAL	148	51	34,5%

Tabela 6.7 Resultados do RIST-Br: RNA 86:172:4, Treinamento 3 (ver Seção 6.2.5.3), limiar para a visualização de 0,900 (ver seções 6.2.5.3 e 6.3.3)

Limiar para Identificação = 0,900	Parada Obrigatória	Área Escolar
Apenas VP	33,3%	50,0%
VP e FP	15,2%	4,2%
TOTAL	**48,5%**	**54,2%**

Limiar para Identificação = 0,900	Sentido Obrigatório	Velocidade Máxima
Apenas VP	3,8%	13,0%
VP e FP	26,4%	21,7%
TOTAL	**30,2%**	**34,8%**

Limiar para Identificação = 0,900	Outras Placas	TOTAL RIST-Br
Apenas VP	20,0%	**20,9%**
VP e FP	20,0%	**18,9%**
TOTAL	**40,0%**	**39,9%**

Limiar para Identificação = 0,900	Parada Obrigatória	Área Escolar
Apenas FP	9,1%	0,0%
Sem Reconhecimento	42,4%	45,8%
TOTAL	**51,5%**	**45,8%**

Limiar para Identificação = 0,900	Sentido Obrigatório	Velocidade Máxima
Apenas FP	52,8%	13,0%
Sem Reconhecimento	17,0%	52,2%
TOTAL	**69,8%**	**65,2%**

Limiar para Identificação = 0,900	Outras Placas	TOTAL RIST-Br
Apenas FP	26,7%	**25,7%**
Sem Reconhecimento	33,3%	**34,5%**
TOTAL	**60,0%**	**60,1%**

Tabela 6.8 Resultados Globais do RIST-Br: RNA 86:172:4, Treinamento 3 (ver Seção 6.2.5.3), limiar para a visualização de 0,900 (ver seções 6.2.5.3 e 6.3.3)

	Mecanismo Atencional (Localização de Subimagens Candidatas)		
Tipos de Sinais	**Total de Imagens**	**Quantidade de Imagens - Sinal Detectado**	**Porcentagem de Detecção**
Parada Obrigatória	33	19	57,6%
Área Escolar	24	19	79,2%
Sentido Obrigatório	53	51	96,2%
Velocidade Máxima	23	21	91,3%
Outras Placas	15	14	93,3%
TOTAL	148	124	83,8%

	Reconhecimento (Identificação dos Sinais)		
Tipos de Sinais	**Total de Imagens**	**Quantidade de Imagens - Apenas Verdadeiro Positivo**	**Porcentagem de Apenas Verdadeiro Positivo**
Parada Obrigatória	33	11	33,3%
Área Escolar	24	8	33,3%
Sentido Obrigatório	53	1	1,9%
Velocidade Máxima	23	2	8,7%
Outras Placas	15	2	13,3%
TOTAL	148	24	16,2%
	Total de Imagens	**Quantidade de Imagens - Verdadeiro Positivo e Falso Positivo**	**Porcentagem de Verdadeiro Positivo e Falso Positivo**
Parada Obrigatória	33	8	24,2%
Área Escolar	24	8	33,3%
Sentido Obrigatório	53	42	79,2%
Velocidade Máxima	23	15	65,2%
Outras Placas	15	9	60,0%
TOTAL	148	82	55,4%
	Total de Imagens	**Quantidade de Imagens - Apenas Falso Positivo**	**Porcentagem de Apenas Falso Positivo**
Parada Obrigatória	33	5	15,2%
Área Escolar	24	3	12,5%
Sentido Obrigatório	53	10	18,9%
Velocidade Máxima	23	1	4,3%
Outras Placas	15	3	20,0%
TOTAL	148	22	14,9%
	Total de Imagens	**Quantidade de Imagens - Sem Reconhecimento**	**Porcentagem Sem Reconhecimento**
Parada Obrigatória	33	9	27,3%
Área Escolar	24	5	20,8%
Sentido Obrigatório	53	0	0,0%
Velocidade Máxima	23	5	21,7%
Outras Placas	15	1	6,7%
TOTAL	148	20	13,5%

Tabela 6.9 Resultados do RIST-Br: RNA 86:172:4, Treinamento 3 (ver Seção 6.2.5.3), limiar para a visualização de 0,600 (ver seções 6.2.5.3 e 6.3.3)

Limiar para Identificação = 0,600	Parada Obrigatória	Área Escolar
Apenas VP	33,3%	33,3%
VP e FP	24,2%	33,3%
TOTAL	**57,6%**	**66,7%**

Limiar para Identificação = 0,600	Sentido Obrigatório	Velocidade Máxima
Apenas VP	1,9%	8,7%
VP e FP	79,2%	65,2%
TOTAL	**81,1%**	**73,9%**

Limiar para Identificação = 0,600	Outras Placas	TOTAL RIST-Br
Apenas VP	13,3%	**16,2%**
VP e FP	60,0%	**55,4%**
TOTAL	**73,3%**	**71,6%**

Limiar para Identificação = 0,600	Parada Obrigatória	Área Escolar
Apenas FP	15,2%	12,5%
Sem Reconhecimento	27,3%	20,8%
TOTAL	**42,4%**	**33,3%**

Limiar para Identificação = 0,600	Sentido Obrigatório	Velocidade Máxima
Apenas FP	18,9%	4,3%
Sem Reconhecimento	0,0%	21,7%
TOTAL	**18,9%**	**26,1%**

Limiar para Identificação = 0,600	Outras Placas	TOTAL RIST-Br
Apenas FP	20,0%	**14,9%**
Sem Reconhecimento	6,7%	**13,5%**
TOTAL	**26,7%**	**28,4%**

Tabela 6.10 Resultados Globais do RIST-Br: RNA 86:172:4, Treinamento 3 (ver Seção 6.2.5.3), limiar para a visualização de 0,600 (ver seções 6.2.5.3 e 6.3.3)

A seguir, podem ser observadas algumas imagens que exemplificam os resultados de reconhecimento de sinais de trânsito do RIST-Br (obtidas a partir do RIST-Br utilizando a RNA do Treinamento 3, com limiar de 0,600).

(a) p267 - Detecção

(b) p267 - Reconhecimento

Figura 6.25 Exemplos de Detecção e Reconhecimento pelo RIST-Br

(a) e286 - Detecção

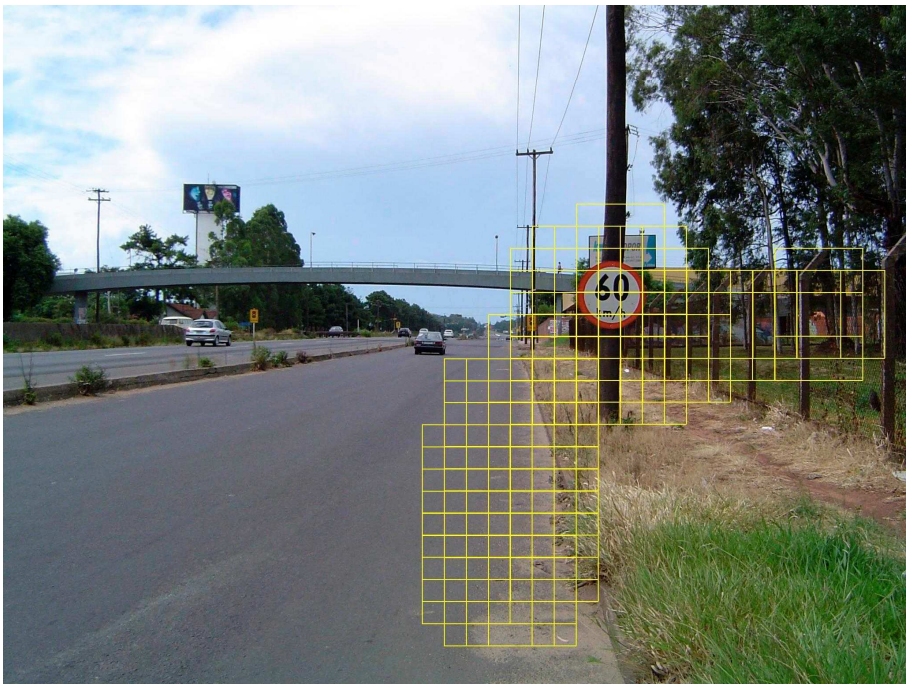
(b) e286 - Reconhecimento

Figura 6.26 Exemplos de Detecção e Reconhecimento pelo RIST-Br

(a) sd358 - Detecção

(b) sd358 - Reconhecimento

Figura 6.27 Exemplos de Detecção e Reconhecimento pelo RIST-Br



(a) v409 - Detecção

(b) v409 - Reconhecimento

Figura 6.28 Exemplos de Detecção e Reconhecimento pelo RIST-Br

6.5 Perspectivas de Trabalhos Futuros

Como perspectivas de trabalhos futuros, o autor desta monografia aponta várias sugestões para aprimoramentos do RIST-Br:

- Treinamento e testes de RNAs de diferentes tipos e topologias, e aplicação da parada antecipada do treinamento da RNA;
- Aumento da base de sinais de trânsito reconhecíveis pelo RIST-Br;
- Reconhecimento de sinais de trânsito em diversas condições climáticas;
- Reconhecimento de sinais de trânsito à noite;
- Reconhecimento de caracteres e figuras presentes nos sinais de trânsito;
- Adaptação do RIST-Br para o reconhecimento de pedestres;
- Tornar o reconhecimento de sinais de trânsito pelo RIST-Br mais eficiente;
- Adaptar o RIST-Br para operação em tempo-real sobre imagens de vídeo;
- Utilização do RIST-Br em conjunto com um veículo de testes [Kelber et al. 2003];
- Implementar o RIST-Br em *hardware*.

6.6 Considerações Finais Sobre Este Capítulo

Neste capítulo foram apresentados o projeto, a implementação, e os resultados do funcionamento do Reconhecedor Inteligente de Sinais de Trânsito Brasileiros - RIST-Br.

O desenvolvimento do RIST-Br envolveu um projeto baseado: (i) no estudo do estado da arte do reconhecimento de sinais de trânsito (Capítulo 2); (ii) nos fundamentos do processamento de imagens (Capítulo 3); (iii) nos fundamentos das redes neurais artificiais (Capítulo 4); e, finalmente, (iv) no desenvolvimento de uma metodologia adequada à resolução do problema (Capítulo 5).

Os resultados apresentados neste Capítulo para o funcionamento do RIST-Br comprovam que o sistema possui potencial para auxiliar os motoristas brasileiros - tornando, assim, o trânsito de veículos mais seguro.

7 CONCLUSÃO

Esta monografia apresentou um sistema de Reconhecimento Inteligente de Sinais de Trânsito Brasileiros - o RIST-Br - projetado e implementado pelo autor.

O objetivo proposto para este trabalho foi o desenvolvimento de um sistema de RIST, através da aplicação de técnicas de inteligência artificial - especificamente, as técnicas envolvendo as redes neurais artificiais.

Para que esse objetivo fosse atingido, foram efetuados estudos que envolveram: (i) a pesquisa sobre o estado da arte do reconhecimento de sinais de trânsito (apresentada no Capítulo 2); (ii) os fundamentos do processamento de imagens (apresentado no Capítulo 3), os quais forneceram uma base importante para o tratamento das imagens e para a metodologia adotada; (iii) os fundamentos das redes neurais artificiais (apresentadas no Capítulo 4), cuja aplicação ao reconhecimento de padrões demonstrou ser adequada à resolução do problema proposto pelo RIST-Br; e, finalmente, os estudos para este trabalho envolveram (iv) o desenvolvimento de uma metodologia adequada à resolução do problema (apresentada no Capítulo 5).

A Tabela 7.1, abaixo, mostra os resultados para a detecção dos sinais de trânsito pelo mecanismo atencional do RIST-Br. A Tabela 7.2, logo em seguida, mostra alguns resultados globais obtidos pelo RIST-Br.

	Mecanismo Atencional (Localização de Subimagens Candidatas)		
Tipos de Sinais	**Total de Imagens**	**Quantidade de Imagens - Sinal Detectado**	**Porcentagem de Detecção**
Parada Obrigatória	33	19	57,6%
Área Escolar	24	19	79,2%
Sentido Obrigatório	53	51	96,2%
Velocidade Máxima	23	21	91,3%
Outras Placas	15	14	93,3%
TOTAL	148	124	83,8%

Tabela 7.1 Resultados do RIST-Br: RNA 86:172:4, Treinamento 3 (ver Seção 6.2.5.3, no Capítulo 6), limiar para a visualização de 0,600 (ver seções 6.2.5.3 e 6.3.3, ambas no Capítulo 6)

A Tabela 7.1, acima, mostra que o mecanismo atencional do RIST-Br conseguiu detectar corretamente, em 83,8% das imagens analisadas, o sinal de trânsito determinado pelo usuário para ser procurado pelo RIST-Br na imagem fornecida.

Limiar para Identificação = 0,600	TOTAL RIST-Br
Apenas VP	16,2%
VP e FP	55,4%
TOTAL	71,6%

Limiar para Identificação = 0,600	TOTAL RIST-Br
Apenas FP	14,9%
Sem Reconhecimento	13,5%
TOTAL	28,4%

Tabela 7.2 Resultados Globais do RIST-Br: RNA 86:172:4, Treinamento 3 (ver Seção 6.2.5.3, Capítulo 6), limiar para a visualização de 0,600 (ver seções 6.2.5.3 e 6.3.3, ambas do Capítulo 6)

A Tabela 7.2, acima, demonstra o funcionamento global do sistema. Em 16,2% das 148 imagens analisadas, houve ao menos uma subimagem candidata que foi corretamente reconhecida como sendo o sinal de trânsito presente, sem que "Falsos Positivos" (subimagens incorretamente identificadas como sinais de trânsito) ocorressem. Em 55,4% das imagens, além da subimagem corretamente reconhecida como sendo um sinal de trânsito, houve a ocorrência de pelo menos uma subimagem incorretamente reconhecida. O RIST-Br foi capaz de reconhecer o sinal de trânsito presente na cena em 71,6% das 148 imagens (isto é, em 106 das 148 imagens). Já as imagens em que apenas "Falsos Positivos" (subimagens incorretamente reconhecidas) foram encontrados, juntamente com as imagens em que as subimagens não foram classificadas (ou não foram encontradas pelo mecanismo atencional), totalizaram 28,4%, ou 42 das 148 imagens analisadas.

Os resultados obtidos pelo RIST-Br, apresentados nas duas tabelas acima e na seção 6.4 do Capítulo 6, comprovam o potencial do sistema, o qual apresenta várias possibilidades de aprimoramento, algumas das quais são citadas pelo autor na seção 6.5, também do Capítulo 6.