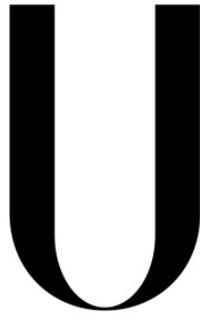


Universidade de Lisboa
Faculdade de Ciências
Departamento de Engenharia Geográfica, Geofísica e Energia



LISBOA

UNIVERSIDADE
DE LISBOA

**Reconhecimento automático de sinalização vertical de trânsito a partir de
dados vídeo de um sistema de mapeamento móvel**

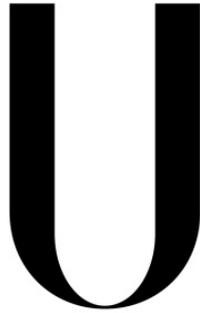
Francisco Miguel Pontes dos Santos

Dissertação

Mestrado em Engenharia Geográfica

2013

Universidade de Lisboa
Faculdade de Ciências
Departamento de Engenharia Geográfica, Geofísica e Energia



LISBOA

UNIVERSIDADE
DE LISBOA

**Reconhecimento automático de sinalização vertical de trânsito a partir de
dados vídeo de um sistema de mapeamento móvel**

Francisco Miguel Pontes dos Santos

Dissertação

Mestrado em Engenharia Geográfica

Orientador:

Professor Doutor João Catalão Fernandes (DEGGE - FCUL)

2013

Resumo

Este trabalho consiste no estudo e desenvolvimento de um método de Reconhecimento Automático de Sinalização Vertical de Trânsito (RASVT), nomeadamente da sinalização de perigo e alguma da sinalização de regulamentação (cedência de passagem, proibição e obrigação), toda de cor vermelha ou azul. O ponto de partida para o seu desenvolvimento são os dados vídeo de um Sistema de Mapeamento Móvel, obtidos em condições reais (não controladas).

A metodologia desenvolvida de reconhecimento de sinalização vertical de trânsito funciona em modo pós-aquisição, ou seja, após aquisição e processamento dos dados vídeo em bruto (*raw*) e dos dados de posicionamento (GPS e IMU). O método desenvolvido é constituído por três fases. A fase de detecção, que consiste em detectar Regiões de Interesse da imagem (RdI) que possam conter sinalização. Esta detecção é feita utilizando um processo de segmentação por cor (vermelha e azul) e utilizando cinco critérios de selecção: dimensões da RdI, rácio altura-comprimento da RdI, proximidade da RdI aos limites da imagem, posicionamento relativo entre o centróide da RdI e centróide da área segmentada e rácio de preenchimento da RdI. A fase de classificação, que tem por base as RdI obtidas na fase anterior e consiste no reconhecimento da forma geométrica de cada região, bem como na agregação, por classes, de cada uma dessas RdI analisadas. As classes utilizadas baseiam-se na cor e forma geométrica. A fase de reconhecimento, consiste na identificação da sinalização que tenha sido detectada e classificada e baseia-se na correspondência dos pictogramas presentes nos sinais. Através de um processo de segmentação por cor é extraído o pictograma para de seguida ser feito o seu reconhecimento, através da correspondência entre o pictograma extraído e pictogramas *template*. Esta correspondência é realizada utilizando a correlação simples.

A fase de detecção apresentou uma taxa de sucesso de 32 % (que excluindo os resultados falsos positivos, apresenta uma taxa de sucesso de 89 %), a fase de classificação teve uma taxa de sucesso de 93 % e a fase de reconhecimento teve uma taxa de sucesso de 91 %. A taxa de sucesso global obtido pelo método RASVT implementado é de 81 %, ou seja, da sinalização presente na amostra analisada, 81 % foi correctamente detectada, classificada e reconhecida.

Palavras-Chave:

Reconhecimento automático; Sinalização vertical de trânsito; Segmentação de imagens.

Abstract

This work aims to develop a method for Automatic Traffic Sign Recognition, namely, danger signs and some regulation signs (giving way, prohibition and obligation signs, all of them red or blue). The starting point for developing the method was a Mobile Mapping System's video data, obtained in real conditions (uncontrolled conditions).

The Automatic Traffic Sign Recognition method performs data post-processing (not real time processing) and comprises three stages. The detection stage consists in detecting the image Regions Of Interest (ROI) that may contain signs. This detection is performed using a color segmentation process (red and blue) and using five selection criteria (ROI dimensions, length-height ratio of the ROI, ROI distance to the limits of the image, the relative positioning between ROI's centroid and a targeted area's centroid and finally the filling ratio of ROI). The step of classification is based on the ROI obtained in the previous stage and consists in recognizing each region's geometric shape, as well as the aggregation per classes of each one of these ROI. The classes used are based on color and geometric shapes. The recognition stage consists in the identification of the traffic signs that has been detected and classified, and is based on the correspondence of the pictograms present in the signals. Through a process of color segmentation the pictogram is extracted from the ROI and then its recognition is done by matching with template pictograms. This matching is performed using simple correlation.

The detection stage showed a success rate of 32 % (if we exclude false positive results, the success rate increases to 89 %), the classification stage had a success rate of 93 % and the final recognition stage had a success rate of 91 %. The overall success rate obtained by the implemented method is 81 %, i.e., from the totality of traffic signs present in the sample, 81 % were correctly detected, classified and recognized.

Keywords:

Automatic Recognition; Traffic signs; Image segmentation.

Índice

1.	Introdução.....	1
1.1	Enquadramento.....	1
1.2	Objectivos.....	1
1.3	Estrutura	2
2.	Sinalização Vertical de Trânsito em Portugal	3
2.1	Sinais de Perigo	4
2.2	Sinais de Regulamentação.....	5
3.	Estado de arte do Reconhecimento Automático de Sinalização Vertical de Trânsito.....	7
3.1	Métodos de Detecção da Sinalização	8
3.1.1	Métodos Baseados na Detecção por Cor	8
3.1.2	Métodos Baseados na Detecção de Forma Geométrica.....	11
3.1.3	Métodos Baseados na Detecção por Cor e Forma Geométrica	13
3.2	Métodos de Reconhecimento e Classificação da Sinalização	14
3.2.1	Métodos Baseados em Redes Neurais	14
3.2.2	Métodos Baseados em <i>Template Matching</i>	15
4.	Método Desenvolvido & Resultados Obtidos	18
4.1	Detecção da Sinalização	19
4.1.1	Segmentação por Cor	23
4.1.2	CrITÉrios de Aceitação / Rejeição de “Clusters”	25
4.1.3	Resultados Obtidos.....	32
4.2	Classificação da Sinalização	38
4.2.1	Identificação da Forma Geométrica	40
4.2.2	Resultados Obtidos.....	43
4.3	Reconhecimento da Sinalização	45
4.3.1	Correspondência de imagens (<i>image matching</i>).....	46
4.3.2	Resultados Obtidos.....	50
5.	Conclusão	53

6. Referências	55
7. Anexos.....	60
Anexo A - Sinalização presente na amostra utilizada para testar o método implementado	60
Anexo B – Documentação do código implementado para reconhecimento automático de sinalização vertical de trânsito num vídeo.	62

Índice de figuras

Figura 1 – Mapa esquemático dos países que assinaram e adoptaram a Convenção sobre a Sinalização Rodoviária de Viena de 8 de Novembro de 1968 [3], retirado de [4].	3
Figura 2 – Exemplos de sinalização de perigo. (a) Sinal com as características típicas de sinal de perigo. (b) Sinais com características que se diferenciam da generalidade da sinalização de perigo.	5
Figura 3 – Exemplos de sinais de cedência de passagem que apresentam formas e cores distintas.	5
Figura 4 – Exemplos de sinais de proibição. (a) Sinais com características típicas de sinal de proibição. (b) Sinal de paragem e estacionamento proibidos. (c) Sinal de fim de proibição.	6
Figura 5 – Exemplos de sinalização de obrigação. (a) Sinais de obrigação. (b) Sinais de fim de obrigação.	6
Figura 6 – Diferentes fases constituintes dos sistemas de reconhecimento automático de sinalização vertical de trânsito, de acordo com [8].	7
Figura 7 – Representação gráfica do espaço cor RGB, imagem retirada de [10].	9
Figura 8 – Representação gráfica dos espaços cor HSL e HSV (cilindro e cone), imagem retirada de [18] e [19].	11
Figura 9 – Distribuição e pesos de cada uma das seis regiões pré-determinadas de busca de vértices referidas em [24].	12
Figura 10 – Esquema explicativo das diferentes fases integrantes do método de reconhecimento automático de sinalização vertical de trânsito (RASVT) implementado.	19
Figura 11 – Esquema dos processos que constituem a fase de detecção de sinalização do método de reconhecimento automático de sinalização implementado.	20
Figura 12 – (a) <i>frame</i> extraído do vídeo original. (b) imagem binária resultante do processo de segmentação por cor descrito no subcapítulo 4.1.1.	21
Figura 13 – (a) exemplo de imagem binária de um triângulo resultante do processo de segmentação por cor. (b) a mesma imagem após ter sido aplicado o processo de dilatação morfológica com o intuito de tornar o triângulo num <i>cluster</i> único e fechado.	22
Figura 14 – Identificação e enumeração dos nove <i>clusters</i> presentes na imagem. Cada <i>cluster</i> corresponde a uma área delimitada, constituída por <i>pixels</i> brancos conectados entre si.	22
Figura 15 – Paleta de cores usada para determinar as gamas de variação da componente H (Tonalidade (<i>Hue</i>)) para as cores vermelha (à direita em cima) e azul (à direita em baixo).	23
Figura 16 – Nesta figura é apresentado um <i>frame</i> (em cima ao centro) e as segmentações realizadas a esse <i>frame</i> para a cor azul (à esquerda em baixo) e para a cor vermelha (à direita em baixo) sem restrição das componentes Saturação (S) e Valor (V) da imagem original.	24
Figura 17 – Nesta figura é apresentado o mesmo <i>frame</i> que o apresentado na Figura 16 (em cima ao centro) e as segmentações realizadas a esse <i>frame</i> para a cor azul (à esquerda em baixo) e para a cor vermelha (à direita em baixo) com os mesmos parâmetros para a componente Tonalidade (H) que os usados na Figura 16, mas com restrição dos parâmetros de Saturação (S) e Valor (V).	25

Figura 18 – <i>Cluster</i> criado a partir de um <i>frame</i> (à esquerda). O mesmo <i>cluster</i> com o rectângulo delimitador de dimensões mínimas, utilizado nos diferentes critérios usados para aceitar ou rejeitar <i>clusters</i> (à direita).....	26
Figura 19 – Sinais captados numa sequência consecutiva de <i>frames</i> com a indicação do rectângulo do <i>cluster</i> delimitador do sinal e das suas dimensões (em <i>pixels</i>) em cada fragmento de <i>frame</i>	26
Figura 20 – Fragmentos de <i>frames</i> nos quais estão presentes sinais, com a indicação do respectivo rácio altura-comprimento (Rac) do rectângulo do <i>cluster</i> para cada situação.....	28
Figura 21 – Exemplo de um sinal/ <i>cluster</i> rejeitado devido a estar em contacto com os limites da imagem.	30
Figura 22 – Posicionamento esperado para o centróide dos sinais circulares e triangulares (normais e em posição inversa).	30
Figura 23 – Esquema exemplificativo do critério de aceitação/rejeição de um <i>cluster</i> baseado no posicionamento do seu centróide numa área circundante do centro geométrico do seu rectângulo para quatro situações. Nas situações (a), (b) e (c) o <i>cluster</i> é aceite. Na situação (d) o <i>cluster</i> é descartado.	31
Figura 24 – <i>Clusters</i> fechados, com "buracos" no seu interior (identificados com a seta amarela) e respectivo resultado obtido depois de efectuado processo de preenchimento dos "buracos".	31
Figura 25 – Sinalização de cor azul não detectada devido às condições de iluminação (contra luz)...	34
Figura 26 – Sinalização de cor vermelha não detectada. (a) Devido ao desrespeito dos critérios I e II de aceitação de <i>clusters</i> . (b) devido ao desrespeito pelo critério IV de aceitação de <i>clusters</i> . (c) devido às condições de iluminação.	35
Figura 27 – Exemplo da influência que diferentes condições de iluminação têm na coloração da sinalização.	36
Figura 28 – Exemplo da influência que o posicionamento do sol tem relativamente à coloração da sinalização.	36
Figura 29 – Situações em que a sinalização se encontra oclusa por diversos elementos presentes na via, impossibilitando a sua detecção.	37
Figura 30 – Situações de falsos positivos provocados por <i>clusters</i> com características semelhantes às da sinalização presentes em: (a) terreno / vegetação, (b) veículo ou partes integrantes dos mesmos, (c) sinais publicitários ou outros sinais que não se pretendiam detectar.	37
Figura 31 – Esquema dos processos que constituem a fase de classificação de sinalização do método de reconhecimento automático de sinalização implementado.....	39
Figura 32 – Exemplos de sinalização com características de cor e forma geométrica semelhantes, mas que pertencem a diferentes classes na legislação nacional. (a) sinais pertencentes às classes de regulamentação (à esquerda) e de perigo (à direita). (b) sinais pertencentes às classes de perigo (à esquerda) e regulamentação (à direita).....	40

Figura 33 – Imagens binárias com as respectivas RdI, provenientes da fase de detecção, assinaladas. Imagem binária resultante da detecção de cor azul (em baixo à esquerda) e da cor vermelha (em baixo à direita).....	41
Figura 34 – Posicionamento, dimensão e peso de cada uma das oito regiões pré determinadas a analisar em cada RdI. O peso encontra-se indicado no centro de cada região.	42
Figura 35 – Exemplo do método de identificação de forma geométrica para: (a) sinal circular; (b) sinal de “stop”; (c) sinal triangular. Com indicação do ParRdI e respectivos pesos considerados no seu cálculo.	42
Figura 36 – As quatro ocorrências consideradas como falsas positivas na fase de classificação.	44
Figura 37 – Esquema dos processos que constituem a fase de reconhecimento de sinalização do método de reconhecimento automático de sinalização implementado.....	45
Figura 38 – A correspondência de imagens (<i>image matching</i>) é realizada somente com <i>templates</i> que tenham a mesma classe que a RdI.	46
Figura 39 - Esquema dos processos que foram implementados por forma a realizar a identificação da sinalização presente nas RdI, para RdI pertencentes à classe “Círculo Azul” (à esquerda) e para RdI pertencentes às restantes clases (à direita).....	48
Figura 41 – Erros encontrados na fase de reconhecimento. Em cima encontram-se as RdI e em baixo as respectivas imagens binárias utilizadas na correspondência com os <i>templates</i> . As situações (a) e (b) são RdI que não foram correctamente identificadas, enquanto (c) e (d) são RdI que não foram identificadas.	51
Figura 42 – Esquema ilustrativo do <i>input</i> e <i>output</i> das funções implementadas de modo a se proceder ao reconhecimento automático de sinalização vertical de trânsito.....	63

Índice de tabelas

Tabela 1 – Exemplo da variação das dimensões (em <i>pixels</i>) mínimas e máximas da sinalização presente em zonas distintas do vídeo. Apenas foram consideradas dimensões maiores do que 25 <i>pixels</i> para sinalização de cor azul e 35 <i>pixels</i> para sinalização de cor vermelha. Os sinais encontram-se identificados de acordo com a nomenclatura utilizada em [5].	27
Tabela 2 – Exemplo de rácios altura-comprimento (<i>Rac</i>) de sinalização presente no vídeo. Os sinais encontram-se identificados de acordo com a nomenclatura utilizada em [5].	29
Tabela 3 – Rácios de preenchimento do rectângulo de <i>cluster</i> (<i>Rprc</i>) esperados para sinalização do tipo circular, triangular e triangular invertida.	32
Tabela 4 – Resultados obtidos para a fase de detecção.	33
Tabela 5 – Resultados obtidos para a fase de classificação tendo por base os resultados da fase de detecção (subcapítulo 4.1.3).	43
Tabela 6 – Resultados obtidos para a fase de reconhecimento, tendo por base os resultados da fase de detecção (subcapítulo 4.1.3) e da fase de classificação (subcapítulo 4.2.2).	50

Lista de acrónimos

CCN – Correlação Cruzada Normalizada

CS – Correlação Simples

HSI – Espaço cor *Hue Saturation Intensity* (Tonalidade Saturação Intensidade)

INIR – Instituto Nacional de Infra-Estruturas Rodoviárias

ParRdI – Parâmetro criado para uma Região de Interesse

Rac – Rácio altura-comprimento do rectângulo de um *cluster*

RASVT – Reconhecimento Automático de Sinalização Vertical de Trânsito

RdI – Região/Regiões de Interesse

RGB – Espaço cor *Red Green Blue* (Vermelho Verde Azul)

Rprc – Rácio de preenchimento do rectângulo de um *cluster*

RST – Regulamento de Sinalização de Trânsito

SMM – Sistema de Mapeamento Móvel

SVT – Sinalização Vertical de Trânsito

1. Introdução

1.1 Enquadramento

O tráfego rodoviário assume, nos dias de hoje, uma grande importância na organização da sociedade. De modo a garantir a correcta circulação dos veículos e a segurança de bens humanos e materiais, cada governo adoptou convenções e normas específicas de regulamentação rodoviária. Algumas dessas regras são comunicadas aos condutores através de sinalização vertical de trânsito (SVT). A correcta e rápida interpretação da SVT, durante a condução, tem um papel fundamental na segurança do próprio condutor, bem como na segurança dos restantes utilizadores do meio rodoviário.

Existem entidades públicas, ou privadas (nos casos das concessionárias), que detêm obrigações legais na correcta sinalização das vias rodoviárias e na manutenção do bom estado de conservação dessa mesma sinalização e que poderão ser responsabilizadas em caso de incumprimento. Nos casos particulares das concessionárias, existem inclusivamente sanções e coimas que poderão ser impostas caso as normas de sinalização que lhes são imputadas não sejam respeitadas, conforme poderá ser consultado nos contractos de concessão presentes no *website* do Instituto Nacional de Infra-Estruturas Rodoviárias (INIR) [1].

Tendo em conta estes pressupostos, existe a necessidade, tanto por parte das concessionárias, como dos institutos públicos responsáveis, de efectuar uma monitorização periódica à sinalização presente nas vias rodoviárias sob a sua alçada. Uma das possibilidades que actualmente é utilizada para realizar este tipo de monitorização baseia-se nos Sistemas de Mapeamento Móvel (SMM). O uso deste tipo de sistemas pressupõe a identificação da sinalização de forma manual por parte de um operador, uma vez que não está disponível no mercado *software* que possibilite a identificação automática da mesma. Desta forma, torna-se pertinente o desenvolvimento de uma aplicação que automatize este processo de detecção e reconhecimento da SVT.

1.2 Objectivos

O objectivo deste trabalho é o desenvolvimento de um método de Reconhecimento Automático de Sinalização Vertical de Trânsito (RASVT), nomeadamente da sinalização de perigo e alguma da sinalização de regulamentação (cedência de passagem, proibição e obrigação) que contenha coloração vermelha ou azul. O ponto de partida para o seu desenvolvimento são os dados vídeo de um Sistema de Mapeamento Móvel, obtidos em condições reais (não controladas).

Como resultado final pretende-se obter a identificação dos *frames* do vídeo em que foi detectada sinalização, assim como, a identificação da sinalização detectada.

1.3 Estrutura

O presente trabalho encontra-se estruturado em cinco capítulos fundamentais.

O primeiro capítulo consiste numa introdução onde se apresentam o enquadramento geral, que identifica a motivação da elaboração do trabalho e os principais objectivos a atingir.

Os capítulos 2 e 3 são capítulos introdutórios que enquadram o leitor às leis e regras existentes em Portugal sobre a Sinalização Vertical de Trânsito (SVT) e ao estado de arte sobre o Reconhecimento Automático de Sinalização Vertical de Trânsito. Neste último, são apresentados alguns métodos existentes, as suas fases constituintes e as abordagens mais comumente adoptadas pelos investigadores.

O capítulo 4 é o capítulo que descreve os passos que foram tomados nas fases de detecção, classificação e reconhecimento da SVT. Nele são apresentadas as explicações necessárias à compreensão do método desenvolvido e os resultados obtidos, bem como as principais dificuldades encontradas.

No quinto capítulo são apresentadas conclusões que se podem retirar com o trabalho realizado e os resultados gerais do método implementado.

Na secção de anexos, no anexo A, está presente a listagem de sinalização vertical de trânsito presente na amostra utilizada para testar o método. No anexo B está presente a documentação explicativa do código desenvolvido.

2. Sinalização Vertical de Trânsito em Portugal

A sinalização de trânsito não é mundialmente uniformizada, existindo diferentes sistemas de sinais consoante a região do Mundo sob análise. O sistema de sinalização vertical de trânsito (SVT) em Portugal evoluiu de forma lenta, com alterações esporádicas e muitas vezes impulsionadas por acordos internacionais.

Desde 28 de Setembro de 2009 [2] que Portugal adoptou as normas internacionais acordadas pela Convenção sobre a Sinalização Rodoviária de Viena de 8 de Novembro de 1968 [3]. Juntamente com Portugal outros 34 países assinaram a Convenção e um total de 62 adoptaram e ratificaram a mesma, Figura 1.

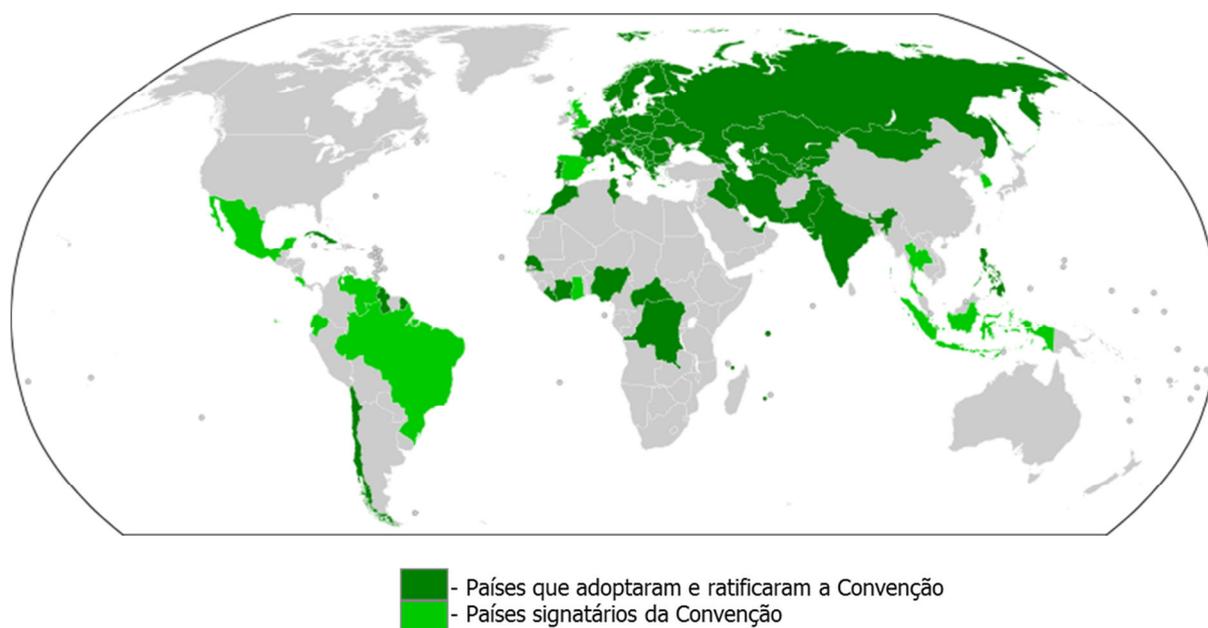


Figura 1 – Mapa esquemático dos países que assinaram e adoptaram a Convenção sobre a Sinalização Rodoviária de Viena de 8 de Novembro de 1968 [3], retirado de [4].

Nesta convenção as partes contratantes acordaram em diversas disposições gerais, que poderão ser consultadas em [2], uma vez que houve o reconhecimento que a uniformidade internacional dos sinais, dos símbolos e das marcas rodoviárias é necessária para facilitar a circulação rodoviária internacional e para melhorar a segurança nas estradas. Algumas destas disposições dizem respeito à SVT, nomeadamente na homogeneização das características da sinalização, como por exemplo os pictogramas e a localização a serem utilizados para esta.

A SVT em Portugal é regida, entre outros diplomas, pelo Regulamento de Sinalização de Trânsito (RST - Decreto Regulamentar n.º 22-A/98, de 1 de Outubro, com a redacção dada pelos Decretos Regulamentares n.º 41/2002, de 20 de agosto, e n.º 13/2003 de 21 de Junho), desenvolvido em

concordância com a Convenção anteriormente referida. Consultando o RST, particularmente nas alíneas respeitantes à SVT, verifica-se que em Portugal essa sinalização é agrupada sob cinco classes distintas e respectivas subclasses:

- ❖ Sinais de Perigo
- ❖ Sinais de Regulamentação
 - sinais de cedência de passagem
 - sinais de proibição
 - sinais de obrigação
 - sinais de prescrição específica
 - sinais de selecção de vias
 - sinais de afectação de vias
 - sinais de zona
- ❖ Sinais de Indicação
 - sinais de informação
 - sinais de pré-sinalização
 - sinais de direcção
 - sinais de confirmação
 - sinais de identificação de localidades
 - sinais complementares
 - painéis adicionais
- ❖ Sinalização de Mensagem Variável
- ❖ Sinalização Turístico-cultural

Das classes acima mencionadas apenas duas tomam especial relevância para o método a ser desenvolvido, os sinais de perigo e os sinais de regulamentação, particularmente os de cedência de passagem, proibição e obrigação e que contenham as cores vermelha ou azul.

2.1 Sinais de Perigo

“Os sinais de perigo indicam a existência ou a possibilidade de aparecimento de condições particularmente perigosas para o trânsito que imponham especial atenção e prudência ao condutor” [5].

Na sua generalidade, os sinais de perigo são representados por um triângulo equilátero orlado de vermelho, com um pictograma preto aplicado sobre fundo branco, Figura 2. Como excepção destacam-se três sinais, o de trabalhos na via que apresenta fundo amarelo e os sinais de local de passagem de nível sem guarda de uma ou duas ou mais vias, que são representados por uma cruz de fundo branco orlada a vermelho, Figura 2.



Figura 2 – Exemplos de sinalização de perigo. (a) Sinal com as características típicas de sinal de perigo. (b) Sinais com características que se diferenciam da generalidade da sinalização de perigo.

2.2 Sinais de Regulamentação

“Os sinais de regulamentação destinam-se a transmitir aos utentes obrigações, restrições ou proibições especiais e subdividem-se em:

- Sinais de cedência de passagem — informam os condutores da existência de um cruzamento, entroncamento, rotunda ou passagem estreita, onde lhes é imposto um determinado comportamento ou uma especial atenção;
- Sinais de proibição — transmitem aos utentes a interdição de determinados comportamentos;
- Sinais de obrigação — transmitem aos utentes a imposição de determinados comportamentos; (...)” [5].

Os sinais de cedência de passagem têm formas diversificadas podendo ser circulares, triangulares, quadrangulares ou hexagonais, apresentando cor vermelha, branca, azul e/ou amarela. Existe no entanto uma predominância de sinais triangulares orlados a vermelho, com pictograma preto sobre fundo branco, Figura 3.



Figura 3 – Exemplos de sinais de cedência de passagem que apresentam formas e cores distintas.

Os sinais de proibição apresentam uma forma circular, normalmente orlada a vermelho, com ou sem pictograma sobre fundo branco. Existem no entanto algumas exceções, como os sinais de proibição de estacionamento e/ou paragem, que apresentam um fundo azul, ou os sinais de fim de

proibição que não se apresentam orlados a vermelho, sendo brancos com traços diagonais pretos e pictograma cinzento (quando presente), Figura 4.

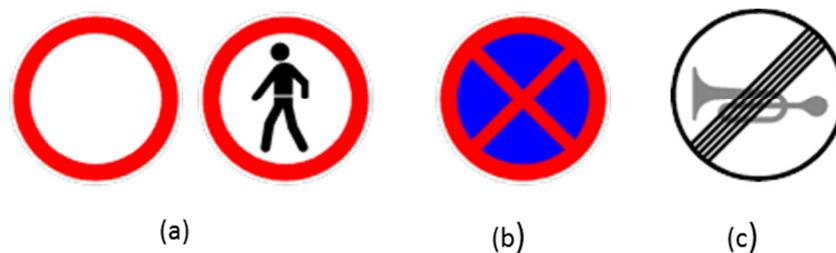


Figura 4 – Exemplos de sinais de proibição. (a) Sinais com características típicas de sinal de proibição. (b) Sinal de paragem e estacionamento proibidos. (c) Sinal de fim de proibição.

Os sinais de obrigação apresentam uma forma circular, cor azul e pictograma branco, podendo ou não ter um traço diagonal a vermelho que representa o fim de uma obrigação específica, Figura 5.

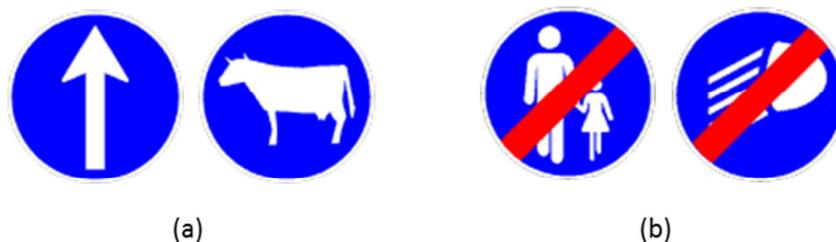


Figura 5 – Exemplos de sinalização de obrigação. (a) Sinais de obrigação. (b) Sinais de fim de obrigação.

3. Estado de arte do Reconhecimento Automático de Sinalização Vertical de Trânsito

O Reconhecimento Automático de Sinalização Vertical de Trânsito (RASVT) a partir de imagens digitais é um problema que concentra esforços por parte da comunidade científica há cerca de 30 anos, tendo o primeiro estudo sobre este tema sido efectuado no Japão no ano de 1984 [6]. Desde essa altura que diversos investigadores de diferentes áreas da ciência têm abordado este tema. Apesar de inicialmente este parecer um problema de fácil resolução, existem diferentes abordagens que mostram formas alternativas de como melhorar a robustez, ou aumentar a eficácia ou ainda como melhorar a solução final. Até ao momento, podemos afirmar que nenhuma das soluções encontradas prevalece sobre as restantes no âmbito geral, havendo soluções que são consideradas mais ou menos eficientes em determinadas situações muito específicas [7]. Apesar desta pluralidade de abordagens, de acordo com [8], podemos considerar que existem três fases distintas em qualquer método de RASVT, Figura 6.

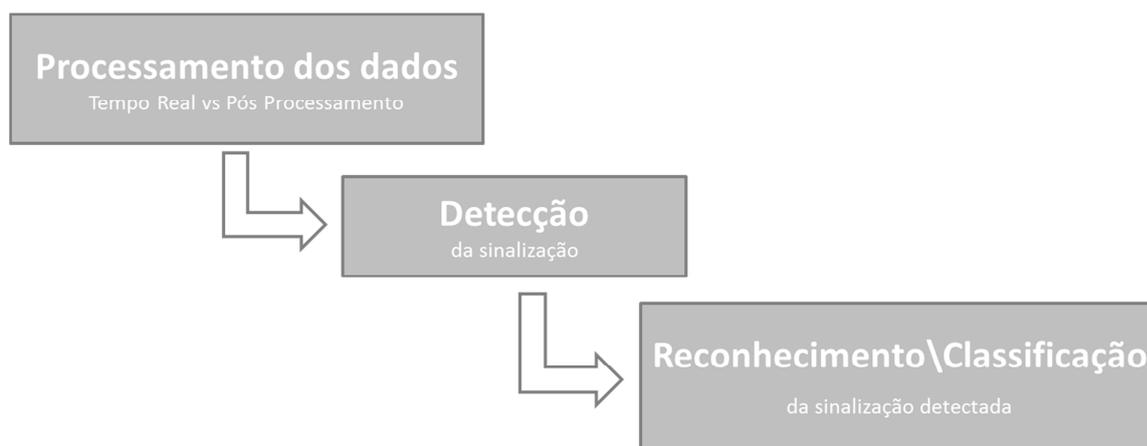


Figura 6 – Diferentes fases constituintes dos sistemas de reconhecimento automático de sinalização vertical de trânsito, de acordo com [8].

Na primeira etapa é necessário decidir se o método irá trabalhar com os dados vídeo em tempo real ou se, por outro lado, é possível realizar o processamento dos dados após a sua obtenção. A detecção em tempo real pode tornar-se complexa, logo a maioria dos métodos existentes opta por realizar um pós processamento dos dados vídeo. Este processo consiste na extracção de cada *frame* e na sua análise de forma isolada dos restantes.

A segunda etapa consiste essencialmente em processos de segmentação da imagem segundo diversos critérios. O seu principal objectivo é reduzir a área de busca dentro da imagem para zonas mais limitadas que cumpram os critérios impostos. Assim, obtêm-se Regiões de Interesse (RdI)

candidatas a conterem sinalização. Esta é considerada a parte mais importante, pois tem de ser um processo eficiente e rápido de modo a reduzir o tempo de processamento geral do método.

A terceira etapa é a etapa final, que consiste em testar cada RdI candidata de acordo com um determinado critério. Os critérios permitem inferir se a região corresponde a um sinal ou não, em caso afirmativo pode-se classificar o sinal de acordo com o critério imposto.

3.1 Métodos de Detecção da Sinalização

Fazendo referência ao capítulo 2 do presente documento, verifica-se que a Sinalização Vertical de Trânsito (SVT) tem características próprias que permitem a sua diferenciação do meio envolvente. Algumas das principais características responsáveis por essa notoriedade são a cor, forma, simbologia e texto. Destas características destacam-se duas para a detecção de SVT: a cor e a forma geométrica. Assim, os métodos de detecção de SVT podem ser agrupados da seguinte forma:

- Métodos baseados na detecção por cor. Estes segmentam a imagem de acordo com um determinado critério baseado na cor de um determinado espaço cor (os mais usados são os espaços-cor RGB, HSV e HSI).
- Métodos baseados na detecção da forma geométrica. São identificados objectos presentes na imagem que se enquadrem nas formas presentes na SVT (essencialmente círculos, triângulos equiláteros, quadrados e rectângulos).
- Métodos baseados na detecção por cor e forma geométrica. Estes consistem em efectuar a detecção numa fusão dos dois métodos anteriores, ou seja, baseiam-se na cor e forma para detectar a SVT.

Dos tipos de métodos de detecção mencionados, os que obtêm uma maior relevância são os métodos baseados na detecção da forma geométrica e os baseados na detecção por cor e forma geométrica. O método baseado na forma geométrica apresenta-se como a única solução de detecção nos casos em que os dados a trabalhar se encontram em escala de cinzentos, o que inviabiliza a detecção por cor. Os métodos de detecção por cor e forma geométrica são mais utilizados quando os dados a trabalhar contêm cor. Assim, é possível realizar-se uma detecção mais robusta e fiável por comparação com a detecção unicamente por cor ou unicamente por forma geométrica, juntando as vantagens de ambos os métodos.

3.1.1 Métodos Baseados na Detecção por Cor

Os autores que optam pela abordagem de segmentação por cor têm como objectivo encontrar RdI e descartar zonas em que haja uma pequena probabilidade de encontrar SVT. A segmentação de uma imagem baseada no atributo da cor pode ser realizada segundo diversas técnicas, como por

exemplo, as técnicas de *thresholding*, de agregação dinâmica de pixels ou de manipulação de histogramas.

Para se poder proceder à segmentação por cor há que definir o espaço cor no qual se pretende efectuar este processo. Os espaços cor mais utilizados são os espaços-cor RGB e HSV / HSI.

3.1.1.1 O espaço cor RGB

O espaço cor RGB é usado em aparelhos electrónicos e principalmente em interfaces gráficas computacionais. É um espaço cor linear aditivo, em que cada cor consiste na adição de três componentes fundamentais, vermelho, verde e azul e é representado num sistema de coordenadas cartesiano tridimensional, conforme ilustrado na Figura 7. A sua representação numérica consiste numa quantificação dos valores das três cores fundamentais. Por exemplo, o azul representado na Figura 7 contém quantificados valores de cada uma das componentes que o integram. É ainda de referir que normalmente os valores de cada uma das componentes variam entre 0 e 1 (como na Figura 7) ou entre 0 e 255, existindo no entanto algumas situações onde este intervalo é escalado para valores que sejam de maior conveniência. O espaço cor RGB é muito sensível às variações de luminosidade, uma vez que essa variação implica uma alteração pouco previsível e simultânea nas suas três componentes (vermelho, verde e azul) [9].

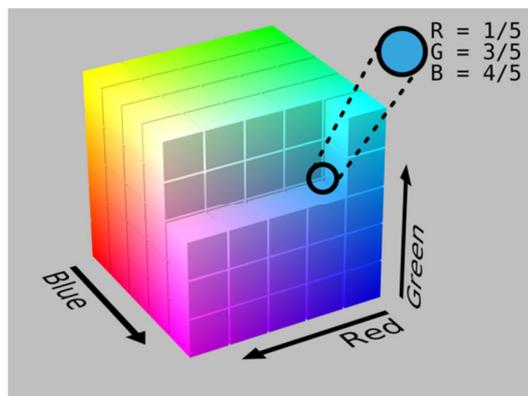


Figura 7 – Representação gráfica do espaço cor RGB, imagem retirada de [10].

Alguns autores optaram por empregar o espaço cor RGB em detrimentos de outros, alguns exemplos são apresentados de seguida.

Os autores de [11] utilizaram a segmentação por cor através de *thresholding* da imagem. O critério utilizado foi a distância (no espaço cor RGB) entre a cor presente em cada *pixel* e uma cor de referência. Um determinado *pixel* é considerado como *pixel* de interesse caso a distância entre a sua cor e a cor de referência esteja contida num determinado intervalo. Outros autores, como por exemplo [12], realizaram estudos sobre a variação da cor (RGB) sob diferentes condições de iluminação. Neste

caso, a segmentação é realizada baseando-se na diferença entre as componentes de Vermelho, Verde e Azul da imagem, pois estas diferenças são de magnitude suficiente para definir uma cor sob diferentes condições de iluminação. Em [13] os autores propõem uma análise detalhada da variação da cor dos *pixels* no espaço cor RGB e desenvolvem subespaços correspondentes à maior variação presente nas cores de SVT. Os subespaços são definidos como regiões canónicas colocadas em linhas rectas do espaço cor RGB para uma combinação de cores primárias, correspondentes às da SVT. Os autores de [14] propuseram uma segmentação da imagem utilizando o espaço cor RGB. Nesta abordagem os autores utilizaram um parâmetro para correcção do efeito da iluminação em cada imagem. Este parâmetro é calculado com base na análise da primeira linha de *pixels* de cada *frame*, normalmente correspondente ao céu. Através deste parâmetro cada *frame* é corrigido e de seguida é aplicado um *threshold* para extracção da RdI. Em [15] os autores identificaram, em cada imagem, um objecto branco. De seguida, compararam a cor do objecto e o branco teórico (no espaço cor RGB). Desta forma calcularam um valor de correcção a aplicar à imagem para, de seguida, extraírem a RdI.

3.1.1.2 O espaço cor HSV / HSI

O espaço cor HSV / HSI é um espaço cor muito utilizado no processamento de imagem. Este é obtido a partir do espaço cor RGB através de uma transformação não linear. Geometricamente este espaço cor é representado por um cilindro ou por um cone, Figura 8. Neste espaço cor existem três componentes fundamentais, a Tonalidade (*Hue*), a Saturação (*Saturation*) e a Intensidade/Valor (*Intensity/Value*). A componente de Tonalidade é normalmente expressa com valores entre 0 e 360 graus ou com valores entre 0 e 1. Através do valor desta componente é possível definir qual a cor que se pretende representar, por exemplo o vermelho corresponde ao valor 0°, o verde ao valor 120° e o azul ao valor 240°. A componente de Saturação permite definir a pureza da cor, sendo que a cor é tanto mais “pura” quanto maior for o valor associado à Saturação e tanto mais esbatida quanto menor for esse valor. Esta componente tem normalmente uma gama de variação entre 0 e 1. A componente Valor/Intensidade permite definir uma cor mais escura ou mais clara, sendo o seu valor máximo correspondente à cor branca e o seu valor mínimo associado à cor preta. Esta componente tem, por norma, um intervalo de variação igual à componente de Saturação. Existem outros espaços cor muito semelhantes ao espaço cor HSV, sendo apenas diferenciáveis em pequenos detalhes, são exemplos disso os espaço cor HSL (*Hue, Saturation, Lightness*) e HSB (*Hue, Saturation, Brightness*) [16].

Comparando com o espaço cor RGB, o espaço cor HSV representa a cor de uma maneira mais próxima da maneira que o olho humano a processa. Desta forma a identificação de uma cor sob diferentes condições de iluminação é muito mais simples do que no espaço cor RGB, uma vez que no espaço cor HSV a cor (Tonalidade (*Hue*)) se encontra numa componente independente relativamente à iluminação (Valor (*Value*)) [16], [17].

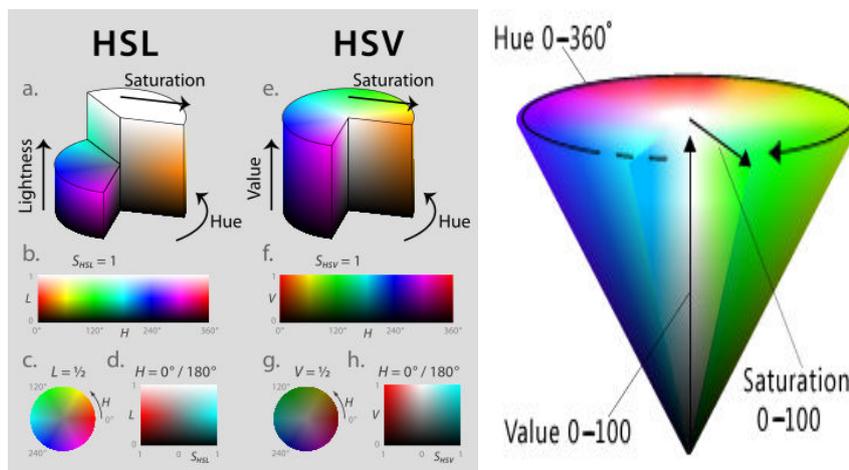


Figura 8 – Representação gráfica dos espaços cor HSL e HSV (cilindro e cone), imagem retirada de [18] e [19].

Alguns exemplos de métodos que empregam este espaço cor são de seguida descritos.

Os autores de [20] converteram a imagem do espaço cor RGB para o espaço cor HSI. De seguida os valores da Tonalidade (H) e de Saturação (S) foram recalculados de maneira a enfatizar a cor vermelha. Posteriormente foram multiplicadas as componentes H e S, o que serve para estas se corrigirem mutuamente, ou seja, caso uma componente contenha valores erróneos é assumido que a outra componente tem os valores dentro da gama correcta. Este processo permite extrair a RdI. Em [21] os autores utilizaram o espaço cor HSI para comparar a cor de cada *pixel* da imagem com uma cor de referência previamente registada. A RdI consiste nos *pixels* que contenham cores suficientemente próximas relativamente às cores de referência. Os autores de [22] optaram por fazer uma transformação própria do espaço cor RGB para o espaço cor HSI. De seguida, realizaram *thresholds* na componente H, para selecção da cor de interesse, e nas componentes S e I, para selecção da gama de variação da luminosidade de interesse, obtendo a RdI.

3.1.2 Métodos Baseados na Detecção de Forma Geométrica

As abordagens tomadas pela maioria dos autores empregam recorrentemente técnicas similares, sendo apenas diferenciáveis por pequenos detalhes. Algumas das técnicas que mais se utilizam são as transformadas de Hough (com pequenas variações entre as diferentes abordagens), os detectores de limites/fronteiras seguidos de *template matching* e as transformadas de simetria radial.

As transformadas de Hough são geralmente utilizadas para encontrar uma forma geométrica arbitrária numa imagem e é uma técnica usada há alguns anos na área da visão computacional [7]. Esta técnica consiste em aplicar numa imagem binária uma transformação, tal que todos os pontos pertencentes a uma mesma linha sejam mapeados num único ponto de um novo espaço de

parametrização da linha em questão. Devido a ter sido uma das primeiras técnicas a ser desenvolvida existem, actualmente, diversas modificações e variantes da mesma.

Em [23], o autor propôs um método geral para detectar polígonos regulares aplicando-o à detecção de SVT. Este método baseia-se na transformada de simetria radial, que revela uma metodologia muito semelhante à presente nas transformadas de Hough. Primeiramente foi construída a imagem gradiente e, de seguida, foram identificados nesta imagem os *pixels* com intensidade mais elevada (correspondentes aos limites dos objectos). Cada um dos *pixels* seleccionados foi então considerado como tendo uma zona de influência, esta define a possível localização do centro geométrico de uma figura geométrica regular. Com base nesta localização foi então possível definir que figura geométrica corresponde à do objecto em análise.

Uma abordagem diferente foi sugerida por [24] em que, primeiramente foi encontrada uma Rdl e, de seguida, foi aplicada a detecção de vértices de Harris [25] à referida Rdl. Na imagem binária resultante do processo anterior foi efectuada uma busca em seis regiões pré-determinadas, em que a existência de um vértice em cada região tem um peso associado (Figura 9). De seguida, a forma do objecto é classificada consoante a combinação da existência ou não de vértices nessas seis regiões.

Os autores de [26] e [27] propuseram um método em que foi usada informação sobre a SVT conhecida *a priori*, como por exemplo o posicionamento esperado da SVT nas imagens. Para além desta informação os autores usaram um detector de limites de *Canny* [28] nas Rdl e fizeram uma análise geométrica a cada *cluster* de *pixels* classificado como limite pelo detector. No caso das formas triangulares, estas foram detectadas classificando cada *cluster* de *pixels* limite como tendo uma orientação horizontal, vertical ou oblíqua, na imagem. Para as formas circulares foi utilizada a correlação entre os *clusters* de *pixels* limite e diversas máscaras¹ circulares pré existentes.

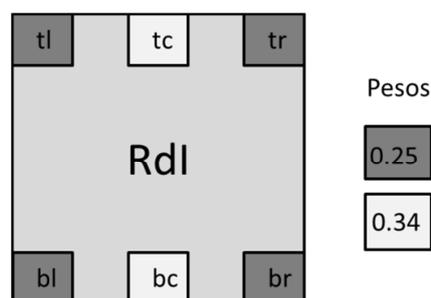


Figura 9 – Distribuição e pesos de cada uma das seis regiões pré-determinadas de busca de vértices referidas em [24].

¹ Máscaras - imagens binárias que são combinadas com a imagem original de modo a isolar zonas da imagem original.

Outros autores, como por exemplo [29] desenvolveram um método que se baseia em detectar limites/fronteiras e respectivos ângulos. O método baseia-se na geração e localização de máscaras de convulsão que usam o ângulo existente entre *pixels* de interesse. Este método é aplicado como um filtro e é usado para detectar limites/fronteiras circulares.

3.1.3 Métodos Baseados na Detecção por Cor e Forma Geométrica

As imagens que contêm informação de cor permitem a utilização conjunta de métodos de detecção de cor e de forma geométrica, desta forma é possível tirar partido dos pontos fortes de ambas as técnicas. Uma vez que são técnicas independentes, obtém-se assim uma mais-valia adicional que permite melhorar a robustez do método, mesmo sob condições que poderiam ser consideradas adversas para uma ou outra técnica isoladamente.

A Tonalidade (H) e a Saturação (S) de um espaço cor HSL melhorado, foi usado por [30], de modo a reconhecer SVT em imagens nocturnas. Foram aplicados *thresholds* dinâmicos aos histogramas das componentes H e S, respectivamente. De seguida, uma imagem binária foi calculada com base nas imagens binárias resultantes dos passos anteriores aplicados às componentes H e S. Na imagem binária obtida foi aplicado um *threshold* de maneira a alocar cada *pixel* em sete fronteiras padrão, sendo essa alocação dependente do *pixel* alvo e da sua vizinhança. Estas fronteiras padrão foram de seguida usadas para especificar os limites da forma geométrica do sinal em questão.

Outros [31], utilizaram o espaço cor HSV para segmentação da imagem, em particular as componentes Tonalidade (H) e Saturação (S). A imagem foi dividida em sub-regiões quadradas de dezasseis *pixels* e, a cada sub-região foi atribuído o valor zero ou um, consoante o número de *pixels*, pertencente a um determinado intervalo de cor, exceda ou não um valor de referência. De seguida, foi efectuada uma busca das sub-regiões às quais tenha sido atribuído o valor um. Por fim, é realizada uma detecção da forma geométrica dos limites/fronteiras resultantes dos passos anteriores.

Usando ainda o espaço cor HSV, foi implementado um método [32] em que este espaço cor é subdividido em subespaços (regiões). Neste método foram usadas as componentes Saturação (S) e Valor (V) para detectar o subespaço onde se encontrava a respectiva Tonalidade (H), tendo sido gerada uma imagem binária neste processo. De seguida, a imagem binária foi analisada por um extractor de forma geométrica, obtendo as formas geométricas presentes. As Rdl foram então comparadas com vários *templates* de diferentes formas geométricas da SVT, presentes numa base de dados, tendo no final sido seleccionado o *template* que apresentou correlação máxima para cada Rdl.

Os autores de [33] propuseram um método que constitui um avanço do descrito em [32]. Nele, foi melhorado o processo de segmentação por cor da imagem inicial. Este passou a gerar Rdl baseadas num algoritmo de crescimento de regiões, a partir de *pixels* base definidos através de *thresholding*.

Relativamente à detecção da forma geométrica presente em cada RdI, esta foi encontrada através de *template matching* das RdI com uma base de dados com imagens da SVT.

Outra abordagem foi tomada pelos autores de [34], sendo esta mais complexa do ponto de vista computacional. A imagem foi primeiramente manipulada por um filtro Laplaciano de Gauss. De seguida, foi utilizada uma rede neuronal de detecção de cor, no espaço cor RGB, para segmentar a imagem e extrair as RdI. Outra rede neuronal foi utilizada para identificar a forma geométrica presente na RdI. Em caso de ter sido detectada uma forma geométrica conhecida na RdI, esta foi apontada como contendo um sinal.

3.2 Métodos de Reconhecimento e Classificação da Sinalização

Uma vez determinada uma RdI, e de nela ter sido detectada SVT, a sinalização terá que ser identificada. A esta fase de identificação dá-se o nome de reconhecimento. O reconhecimento da SVT é normalmente realizado tendo como ponto de partida uma base de dados de imagens de toda a SVT necessária. A grande maioria dos métodos existentes, para reconhecimento de SVT, pode alicerçar-se numa de duas técnicas distintas de relacionar a sinalização presente na base de dados com a sinalização detectada: aprendizagem computacional (redes neuronais) ou *template matching*.

As primeiras técnicas mencionadas exigem métodos computacionalmente complexos e uma grande disponibilidade de dados. Estes dados são necessários para o processo de aprendizagem do classificador associado à rede neuronal. Este processo de aprendizagem desempenha um papel fundamental no correcto funcionamento do classificador. Por outro lado, quando comparadas com as técnicas de *template matching*, as redes neuronais apresentam resultados mais robustos e precisos [6]. De seguida, são aprofundados vários métodos que utilizam as duas técnicas referidas.

3.2.1 Métodos Baseados em Redes Neuronais

As redes neuronais tiveram um grande incremento tecnológico nos anos 90, altura em que o RASVT estava também em fase de grande desenvolvimento. Como tal, as redes neuronais são uma técnica muito utilizada por grande parte dos autores de métodos de RASVT. Este facto deve-se não só à eficácia do uso das mesmas na fase de identificação da SVT, mas também porque a partir de determinada altura os investigadores optaram por tentar inovar, evitando utilizar os classificadores tradicionais usados até então [35]. Por outro lado, esta abordagem veio a revelar alguns inconvenientes referidos em [8], tais como a necessidade de ter a disponibilidade de grande número de imagens de SVT, obtidas sob diferentes condições, de modo a serem fornecidos à rede neuronal para a fase de aprendizagem da mesma. Deste modo, se os dados iniciais não forem suficientes ou se não retractarem as condições de iluminação, condições atmosféricas, ou outras condicionantes existentes no terreno

aquando da obtenção dos dados a serem analisados, a precisão e robustez da fase de reconhecimento do método de RASVT poderá estar comprometida. Outro inconveniente do uso desta técnica é a elevada complexidade computacional presente na sua elaboração, o que exige conhecimentos avançados nas áreas da informática e computação. De seguida, passarão a ser enumeradas algumas técnicas já desenvolvidas.

Em [33] os autores optaram por usar uma rede neuronal que classifica os sinais de acordo com a informação que se encontra no seu interior. A classificação foi realizada fornecendo ao classificador da rede neuronal regiões de interesse quadradas, de 36 *pixels* de lado, de modo a ser reconhecido o sinal presente. Foram obtidas classificações correctas para 84 %, 88 % e 100 % da sinalização apresentada, respectivamente para sinais circulares vermelhos, triangulares vermelhos e circulares azuis.

Em [36] foi usada uma rede neuronal de propagação inversa para reconhecer sinalização de perigo. Como dados de entrada foram usadas RdI quadradas, com 10 *pixels* de lado, da parte interior de cada sinal (pictograma). Este método conseguiu reconhecer 55 % dos sinais de perigo detectados.

Outros [37], optaram por usar uma rede neuronal para detectar a posição e tamanho de sinais de limite de velocidade. O classificador usado na rede neuronal foi aplicado para o reconhecimento dos números presentes nestes sinais. Foi conseguida a detecção de 100 % da sinalização presente e o reconhecimento de 98.3 % da sinalização detectada.

Em [38] os autores utilizaram um método de elevada complexidade [7], que consiste numa rede neuronal que emprega uma variação do detector Viola Jones [39], adaptada ao RASVT. Para além desse, outros classificadores são utilizados, tendo por base a técnica de *AdaBoost* [40], que permite usar vários classificadores de fraco desempenho e combiná-los num único classificador que obtenha resultados robustos e precisos. Foi conseguido um reconhecimento global de 97 % da sinalização detectada.

3.2.2 Métodos Baseados em *Template Matching*

As técnicas de *template matching* existentes para reconhecimento/classificação de SVT baseiam-se em quantificadores de similaridade, de forma a aferir se a imagem candidata corresponde, ou não, a uma determinada imagem pré-existente. O quantificador mais utilizado é a correlação cruzada normalizada (*normalized cross-correlation*). As principais vantagens no uso deste quantificador são a sua simplicidade, robustez às variações de iluminação e facilidade de interpretação estatística [6]. Do ponto de vista prático de implementação é também um quantificador muito utilizado

pois, apesar da sua simplicidade, demonstrou ter resultados equivalentes a um correlacionador óptico² [41]. Além da correlação cruzada normalizada, outros quantificadores de semelhança podem ser usados, como a correspondência de histogramas, a correspondência de padrões ou a correspondência de diferentes tipos de transformadas (Distância, Hough, Fourier).

Os autores de [31] classificaram a SVT segundo a forma geométrica detectada. De seguida, a RdI quadrada foi redimensionada através de interpolação bilinear, de modo a ficar com um tamanho normalizado de 50 *pixels* de lado. Uma vez efectuado o redimensionamento da RdI, foi aplicada a técnica de *template matching*, através da correlação cruzada normalizada, com cada um dos sinais de uma base de dados que correspondessem à classe da RdI. Os *templates* que tivessem uma correspondência superior a um determinado valor foram seleccionados para, de seguida, do grupo previamente identificado, ser seleccionado o sinal da base de dados com maior correspondência com o sinal presente na RdI.

Em [42] a SVT, depois de detectada, foi classificada de acordo com a cor, forma e conteúdos. De seguida, foi realizado o reconhecimento da mesma por *template matching*, usando uma base de dados com *templates* da sinalização. A técnica de *template matching* utilizada é similar à técnica descrita em [43], e emprega uma codificação baseada nas cores a considerar. Neste caso, foi obtido o correcto reconhecimento de 90 % da sinalização detectada.

Os autores de [34] dividiram cada RdI em áreas quadradas e em cada uma destas áreas procederam à sua filtragem, por forma a preencher ou apagar zonas de ruído ou pequenos defeitos. De seguida, procederam a uma normalização radiométrica de cada RdI e ao redimensionamento das mesmas, com o objectivo de ter iguais dimensões às dos *templates* presentes na base de dados. Depois, foi realizado o *template matching* e obtido um correcto reconhecimento para cerca de 95 % da sinalização detectada.

Outra abordagem foi tomada em [30] onde os autores transformaram as RdI e as imagens da base de dados através do método de mapeamento logarítmico complexo [44]. De seguida, aplicaram-lhes transformadas rápidas de Fourier a duas dimensões, de modo a transformá-las para o domínio das frequências. Após este processo, foi realizada a comparação entre cada RdI e cada *template*, sendo seleccionada como correcta a correspondência que detivesse melhor valor de similitude. Todo este processo foi aplicado apenas para sinais de stop, de proibição de estacionamento e de proibição de circulação, tendo sido reconhecidos 100 %, 94 % e 95 %, respectivamente, dos sinais detectados.

Em [45] os autores utilizaram uma técnica de classificação da sinalização que usa o *template matching* baseado em transformadas de distância. Neste método, em cada *template* foram

² Correlacionador óptico: dispositivo utilizado na comparação de dois sinais que utiliza as propriedades de transformação de Fourier de uma lente.

seleccionadas RDI que diferenciassem esse *template* dos restantes. Para tal, foram maximizadas as somas das dissemelhanças entre o *template* em questão e todos os outros *templates*. No vídeo, a cada sinal detectado que esteja presente numa sequência de *frames* é atribuído um determinado peso, sendo tanto maior a sua relevância quanto maior for a sequência em que esse sinal é detectado. A ideia principal desta abordagem é a seguinte: quando um sinal é detectado pela primeira vez, as suas dimensões na imagem são relativamente pequenas, o que dificulta o processo de classificação. Por esta razão, são apenas usados na classificação sinais de maior dimensão na imagem, o que acontece no final da sequência de *frames*. Este método obteve o correcto reconhecimento e detecção de 79.4 % da SVT circular vermelha e 97.3 % da SVT circular azul.

4. Método Desenvolvido & Resultados Obtidos

O método desenvolvido, que será apresentado e explicado ao longo do presente capítulo, teve como ponto de partida os dados vídeo provenientes de um Sistema de Mapeamento Móvel (SMM) que, por sua vez, foram resultantes de um processamento de dados brutos (*raw*) desse mesmo SMM.

Na implementação de um método de reconhecimento automático de sinalização vertical de trânsito (RASVT) o processamento dos dados é feito em tempo-real ou em pós-processamento (consultar o capítulo 3). Neste caso, teve que ser adoptada a única via possível, a de implementar um método de RASVT que realize um pós-processamento dos dados.

O método apresentado pode ser separado em três fases principais, a fase de detecção da SVT, a fase de classificação da SVT detectada e a fase de reconhecimento da SVT classificada, Figura 10.

Na primeira fase, a fase de detecção da SVT, é efectuada uma extracção dos *frames* que constituem o vídeo anteriormente referido. Em cada *frame* é efectuada uma busca de regiões de interesse (RdI), tendo como principal critério as cores presentes na sinalização. A fase seguinte, de classificação, baseia-se nas RdI obtidas na fase de detecção e consiste em agrupar estas regiões em classes coerentes com a coloração e forma geométrica da SVT. Na terceira e última etapa é efectuada uma correspondência entre as RdI obtidas e os *templates* da SVT com a finalidade de se obter a identificação da SVT presente no vídeo inicial.

Os produtos resultantes da execução da aplicação desenvolvida são constituídos por um vídeo e um ficheiro de texto. O vídeo é semelhante ao vídeo inicial, mas com a indicação gráfica do posicionamento da sinalização que foi detectada, bem como com a respectiva identificação realizada. O ficheiro de texto é constituído por uma listagem de todos os *frames* do vídeo onde foi detectada sinalização, da identificação da respectiva sinalização e das coordenadas imagem do centro geométrico de cada um dos sinais detectados. A informação presente neste ficheiro permite que o mesmo seja utilizado para, numa fase posterior de desenvolvimento, ser desenvolvida uma aplicação que permita coordenar a sinalização que tenha sido reconhecida

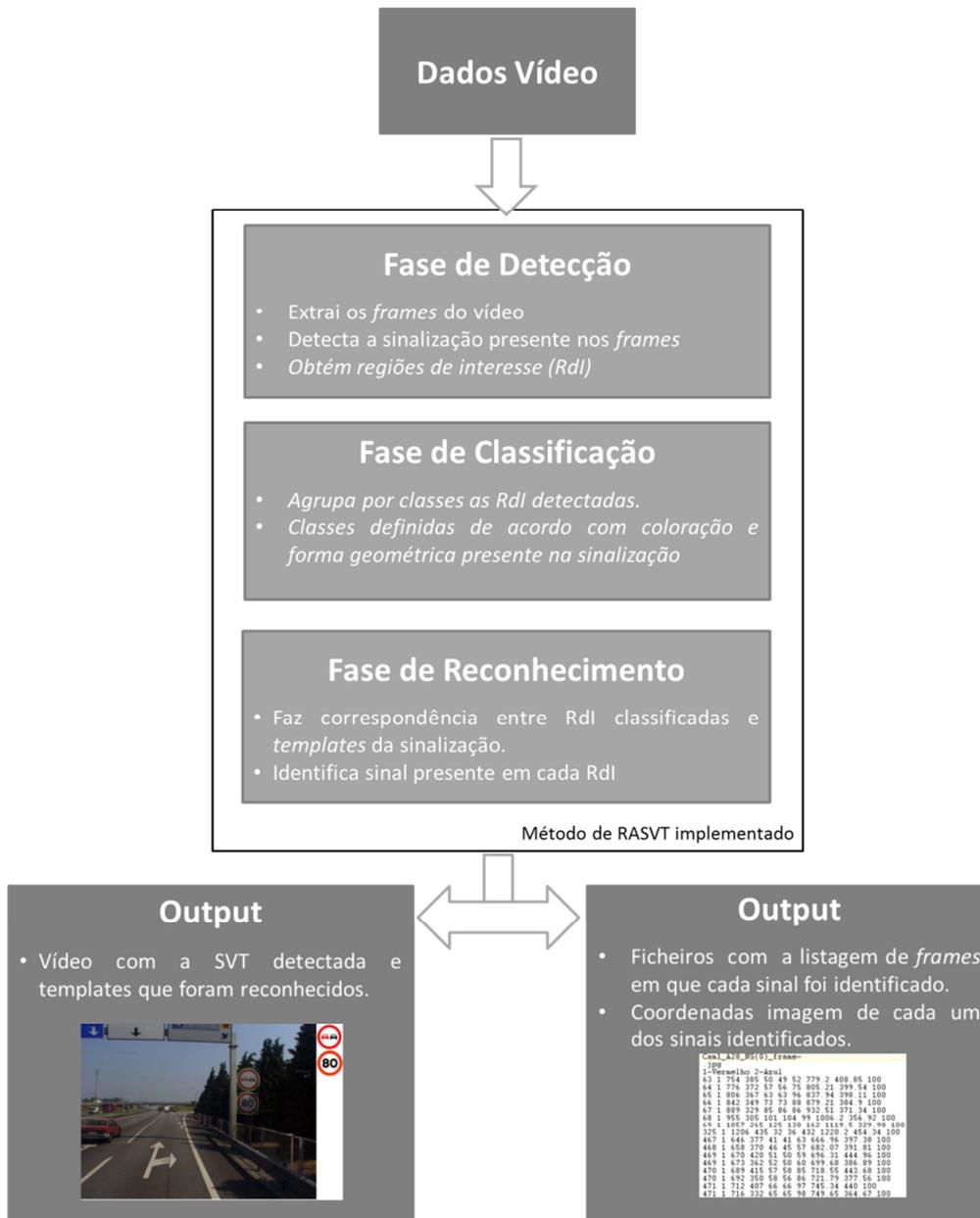


Figura 10 – Esquema explicativo das diferentes fases integrantes do método de reconhecimento automático de sinalização vertical de trânsito (RASVT) implementado.

4.1 Detecção da Sinalização

A detecção é a parte mais importante num sistema de reconhecimento automático de sinalização de trânsito. Essa importância pode ser compreendida, pois um sinal que não seja detectado é um sinal que nunca será processado, sendo portanto a sua presença ignorada. Outro motivo é esta fase ser crucial na definição do tempo de processamento geral do método, uma vez que têm que ser analisados todos os *frames* do vídeo inicial com a finalidade de identificar regiões de interesse (RdI).

Dado que a SVT que se pretende detectar corresponde a sinalização de perigo e a alguma sinalização de regulamentação, e uma vez que esta sinalização apresenta características de cor e forma geométrica bem definidas (consultar capítulo 2), a opção tomada consistiu na implementação de um método de detecção que utilizasse ambas as características anteriormente mencionadas, semelhante ao método apresentado em [31]. Este método utiliza o espaço cor HSV para segmentar a imagem original e assim extrair zonas da imagem com a coloração desejada. De seguida são efectuadas uma série de análises às zonas anteriormente assinaladas de maneira a descartar zonas que não correspondam a SVT. Na Figura 11 estão esquematizadas as etapas que definem a fase de detecção da sinalização do método implementado.

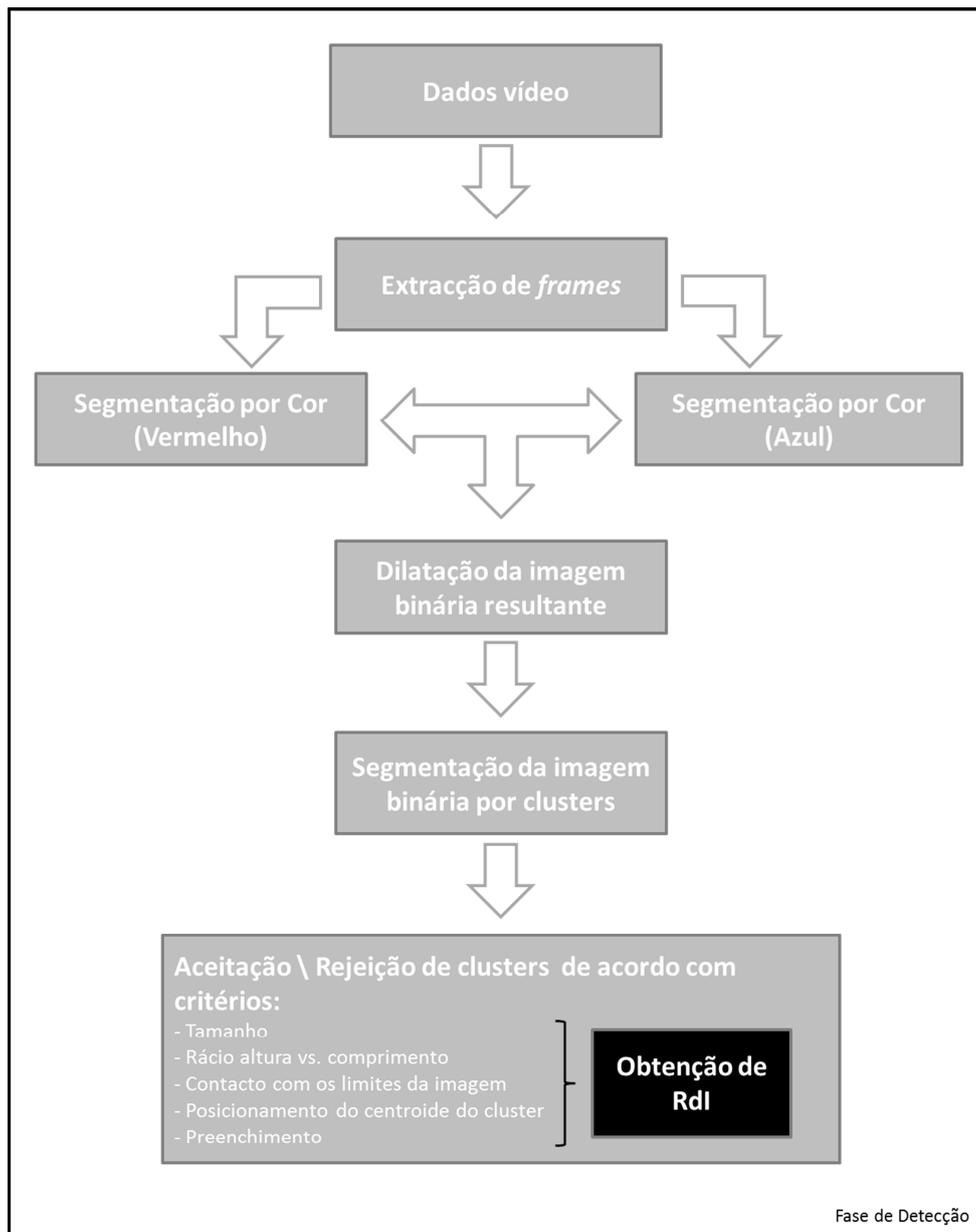


Figura 11 – Esquema dos processos que constituem a fase de detecção de sinalização do método de reconhecimento automático de sinalização implementado.

Na fase de detecção o passo inicial consiste em extrair os *frames* que constituem o vídeo obtido pelo SMM. Neste processo apenas são extraídos os *frames* que sejam únicos, uma vez que no vídeo original existe um grande número de *frames* repetidos.

O segundo passo consiste numa segmentação por cor de cada um dos *frames* anteriormente extraídos. Esta fase é crucial no processo de detecção pois, conforme será desenvolvido no subcapítulo 4.1.1, só as regiões assinaladas nesta segmentação serão alvo de uma análise mais aprofundada relativamente à existência ou não de SVT, sendo que as regiões descartadas neste processo são ignoradas pelos passos seguintes, consultar Figura 12.



Figura 12 – (a) *frame* extraído do vídeo original. (b) imagem binária resultante do processo de segmentação por cor descrito no subcapítulo 4.1.1.

Os dois passos seguintes consistem em aplicar à imagem binária, resultante da segmentação anterior, uma dilatação morfológica e uma agregação de *pixels* por critério de conectividade (criação de *clusters*).

A dilatação morfológica tem o objectivo de conectar conjuntos de *pixels* que estejam relativamente próximos uns dos outros. Desta forma, e tendo em conta a distribuição da cor existente na SVT a detectar, pretende-se que a orla vermelha ou azul dos sinais fique totalmente conectada numa única área comum, consultar Figura 13.

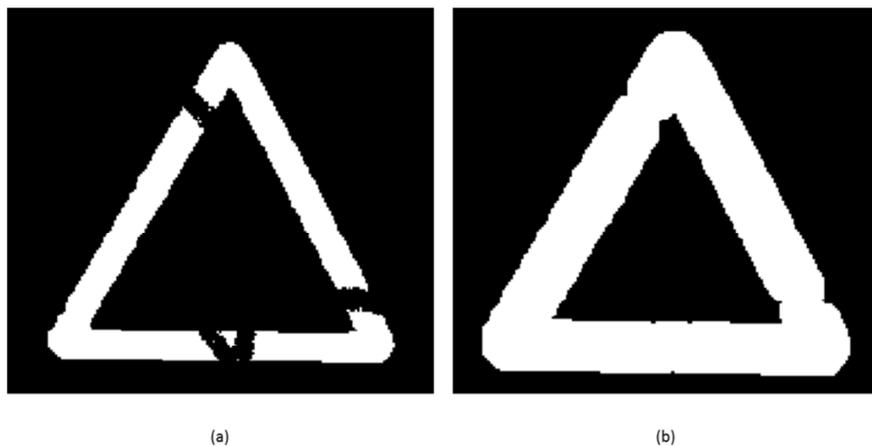


Figura 13 – (a) exemplo de imagem binária de um triângulo resultante do processo de segmentação por cor. (b) a mesma imagem após ter sido aplicado o processo de dilatação morfológica com o intuito de tornar o triângulo num *cluster* único e fechado.

A criação de *clusters* é realizada após ter sido efectuada a dilatação morfológica. Este processo consiste numa agregação de *pixels* que estejam conectados na imagem binária (Figura 14). Após esta agregação cada um destes conjuntos de *pixels* agregados (*clusters*) é catalogado para posteriormente poder ser identificado e analisado.

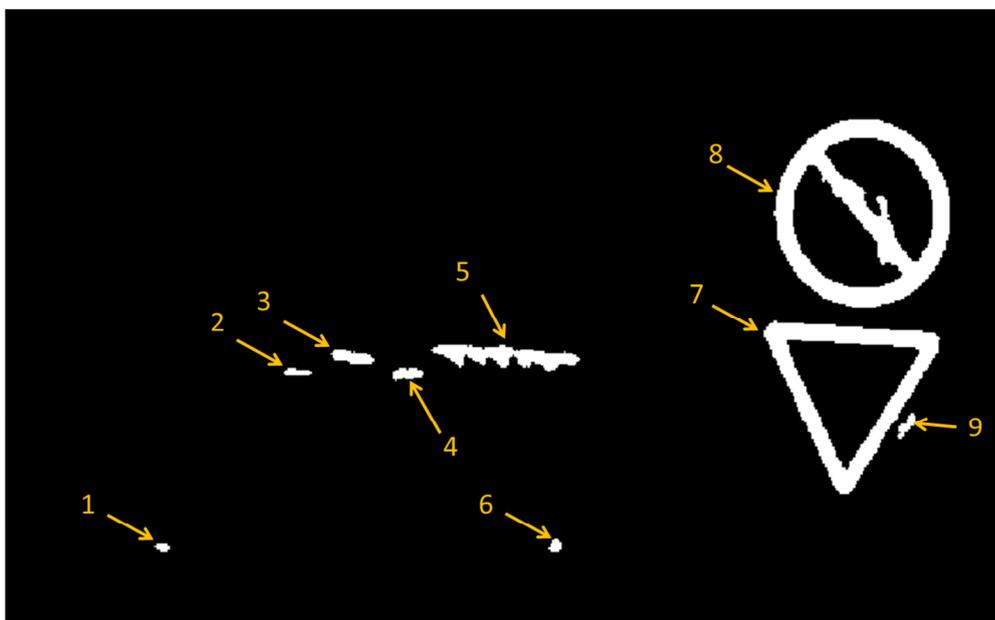


Figura 14 – Identificação e enumeração dos nove *clusters* presentes na imagem. Cada *cluster* corresponde a uma área delimitada, constituída por *pixels* brancos conectados entre si.

Uma vez identificados os diferentes *clusters* existentes, inicia-se o quinto e último passo da fase de detecção, que será aprofundado no subcapítulo 4.1.2. Este passo consiste em testar os *clusters* segundo vários parâmetros e critérios, que permitirão aceitar ou rejeitar os mesmos. No caso de um *cluster* ser aceite em todos os critérios, consideramos que estamos perante uma região de interesse (RdI) da imagem. As RdI obtidas são regiões onde, tendencialmente, existirá SVT.

Esta fase de detecção da SVT tem como principal objectivo a identificação das RdI, sendo que estas serão o ponto de partida para a fase seguinte do método, a fase de classificação.

4.1.1 Segmentação por Cor

Os *frames* extraídos do vídeo inicial estão, por pré-definição, no espaço cor RGB. Este espaço cor, tal como foi referido no subcapítulo 3.1.1.1, apesar de ser amplamente usado no mundo da computação, apresenta algumas desvantagens relativamente ao processamento digital de imagem. A principal desvantagem consiste neste espaço cor ser muito sensível às variações de luminosidade.

De maneira a ultrapassar a dificuldade levantada pela heterogeneidade de efeitos originados pela iluminação, foi realizada uma alteração de espaços cor, passando-se assim a trabalhar no espaço cor HSV. Neste espaço cor os efeitos resultantes de diferentes condições de iluminação são mais facilmente minimizados e previsíveis do que no espaço cor RGB. Tal facto deve-se a no espaço cor HSV a cor (Tonalidade (*Hue*)) se encontrar numa componente independente relativamente à iluminação (Valor (*Value*)), conforme referido no subcapítulo 3.1.1.2.

Uma vez escolhido o espaço cor HSV, foram realizados alguns testes através do uso de uma paleta de cores, Figura 15, com o intuito de aferir qual a gama de variação da componente H (Tonalidade (*Hue*)) para cada uma das cores a detectar, respectivamente a cor vermelha e azul. Assim foram encontradas gamas de variação para a componente H (Tonalidade (*Hue*)) que permitissem identificar áreas onde estavam presentes as cores pretendidas. Este processo foi realizado por limiarização (*threshold*), usando como parâmetros as gamas de variação da componente H.

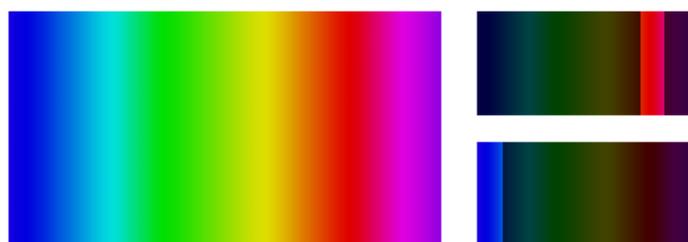


Figura 15 – Paleta de cores usada para determinar as gamas de variação da componente H (Tonalidade (*Hue*)) para as cores vermelha (à direita em cima) e azul (à direita em baixo).

Após algumas experiências iniciais, foi verificado que nos *frames* sob análise existiam diversas zonas com tons de vermelho e azul que, embora não correspondendo a sinalização, estavam a ser incluídas devido à sua cor ser próxima da cor da sinalização, Figura 16. Estas zonas foram sinalizadas como ruído. Assim, ao contrário do que seria inicialmente expectável, revelou-se necessário aplicar uma restrição nas restantes componentes da imagem, as componentes S (Saturação (*Saturation*)) e V (Valor (*Value*)), com a finalidade de melhorar as zonas detectadas e eliminar grande parte do ruído existente, Figura 17.

Analisando a Figura 16 e comparando-a com a Figura 17, verifica-se que grande parte do ruído foi removido através da restrição das componentes de Saturação (S) e Valor (V). É possível ainda verificar que as principais zonas de ruído removidas da segmentação da cor azul correspondem a zonas de sombra (valores reduzidos da componente V) e a zonas correspondentes ao céu (valores reduzidos da componente S). Relativamente à segmentação da cor vermelha foi possível remover a maioria do ruído, correspondente a zonas de sombra e de asfalto (valores reduzidos da componente V). Ainda assim, permanecem algumas zonas de ruído, correspondentes a zonas de solo exposto que apresentam colorações avermelhadas e, portanto, inviáveis de remover pela restrição das componentes de Saturação e Valor da imagem.

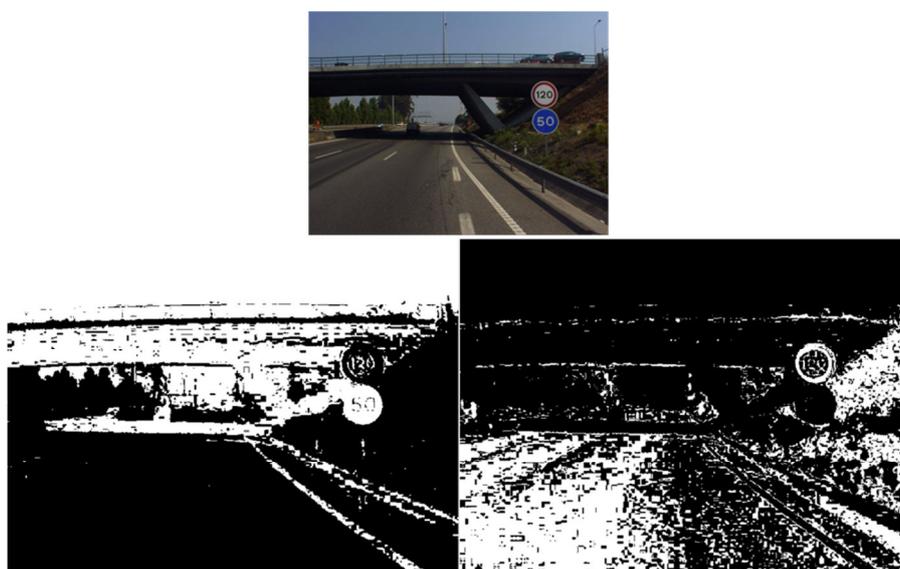


Figura 16 – Nesta figura é apresentado um *frame* (em cima ao centro) e as segmentações realizadas a esse *frame* para a cor azul (à esquerda em baixo) e para a cor vermelha (à direita em baixo) sem restrição das componentes Saturação (S) e Valor (V) da imagem original.



Figura 17 – Nesta figura é apresentado o mesmo *frame* que o apresentado na Figura 16 (em cima ao centro) e as segmentações realizadas a esse *frame* para a cor azul (à esquerda em baixo) e para a cor vermelha (à direita em baixo) com os mesmos parâmetros para a componente Tonalidade (H) que os usados na Figura 16, mas com restrição dos parâmetros de Saturação (S) e Valor (V).

4.1.2 Critérios de Aceitação / Rejeição de “Clusters”

Após um trabalho inicial de segmentação efectuado sobre os *frames* extraídos do vídeo, obtiveram-se imagens binárias que foram submetidas a um processo de criação de *clusters*. Este processo consiste na identificação de *pixels* que sejam conexos na imagem binária para, posteriormente, os agregar sob a forma de *clusters*, conforme foi anteriormente ilustrado na Figura 14.

Uma vez obtidos os diferentes *clusters* presentes na imagem, estes são de seguida analisados segundo cinco critérios que permitem diferenciar *clusters* de interesse de *clusters* considerados ruído:

- I. Dimensões do rectângulo do *cluster*³
- II. Rácio altura-comprimento do rectângulo do *cluster*
- III. Proximidade do *cluster* com os limites da imagem
- IV. Posicionamento relativo entre o centróide do *cluster* e o do rectângulo do *cluster*
- V. Rácio de preenchimento do rectângulo do *cluster* por *pixels* pertencentes ao *cluster*

De forma a facilitar a análise dos elementos presentes numa imagem, foi criado, para cada um dos *clusters* detectados, um rectângulo que os delimitasse, de forma às suas dimensões serem as mínimas possíveis – rectângulo do *cluster*. Desta forma, com o fim de simplificar a implementação

³ Rectângulo de tamanho mínimo que delimita o *cluster*, ver Figura 18.

computacional, alguns dos critérios acima mencionados são aplicados a esse rectângulo em vez de serem aplicados ao *cluster* em si. Para clarificação deste conceito, consultar a Figura 18.



Figura 18 – *Cluster* criado a partir de um *frame* (à esquerda). O mesmo *cluster* com o rectângulo delimitador de dimensões mínimas utilizado nos diferentes critérios usados para aceitar ou rejeitar *clusters* (à direita).

De seguida serão abordados em mais pormenor os cinco critérios utilizados na aceitação ou rejeição de cada um dos *clusters*.

Critério I – Dimensões do rectângulo do *cluster*

Conforme ilustrado pela Figura 19, cada sinal que se encontre numa sequência consecutiva de *frames*, apresenta uma variação nas suas dimensões, ocupando uma área tanto maior, conforme se encontra mais próximo do ponto de tomada de imagem. Assim, para uma amostra aleatória, foi elaborado um levantamento das dimensões (em *pixels*) dos sinais presentes em diferentes *frames* seguidos. Esta amostra foi seleccionada aleatoriamente, pois as dimensões dos sinais a detectar são próximas entre si, conforme poderá ser consultado em [5] e [46]. Assim, foi possível estabelecer uma gama de variação das dimensões para a sinalização presente.



Figura 19 – Sinais captados numa sequência consecutiva de *frames* com a indicação do rectângulo do *cluster* delimitador do sinal e das suas dimensões (em *pixels*) em cada fragmento de *frame*.

Após se ter estimado a gama de variação das dimensões dos sinais nos *frames* (desde o momento em que o sinal aparece até que desaparece da sequência de *frames*), foi realizada uma filtragem dos *clusters* presentes em cada uma das imagens binárias. Por conseguinte, foi possível eliminar alguns *clusters* cujas dimensões não fossem coerentes com as dimensões verificadas. Assim, os *clusters* eliminados foram considerados ruído e excluídos do processo de verificação.

Na Tabela 1 está representado um exemplo de algumas dimensões usadas na definição da gama de variação anteriormente referida. Nela estão presentes as dimensões do retângulo de *cluster* de alguns sinais, no *frame* onde este apresenta dimensões mínimas e no *frame* onde este apresenta dimensões máximas. Está também representada a diferença entre os *frames* considerados para as duas situações referidas. Os valores mínimos utilizados na gama de variação das dimensões dos *clusters* foram de 25 *pixels* e 35 *pixels* para a sinalização de cor azul e vermelha, respectivamente. Estes valores foram utilizados pois para dimensões menores eram incluídas demasiadas zonas de ruído. Enquanto os valores máximos utilizados na gama de variação das dimensões dos *clusters* foram de 250 *pixels*.

Tabela 1 – Exemplo da variação das dimensões (em *pixels*) mínimas e máximas da sinalização presente em zonas distintas do vídeo. Apenas foram consideradas dimensões maiores do que 25 *pixels* para sinalização de cor azul e 35 *pixels* para sinalização de cor vermelha. Os sinais encontram-se identificados de acordo com a nomenclatura utilizada em [5].

Sinal	Dimensões mínimas			Dimensões máximas			Diferença de frames
	<i>frame</i>	comprimento	altura	<i>frame</i>	comprimento	altura	
C13	63	50	49	69	125	130	6
C14a	467	41	41	475	143	143	8
C13	469	52	50	475	147	144	6
B9d	531	42	37	539	157	146	8
D8	56	30	29	69	129	134	13
D1c	554	29	30	566	143	159	12
C13	581	38	39	590	137	147	9
C13	602	41	42	610	135	139	8
A1a	605	58	50	610	135	125	5
	Mínimo	29	29	Máximo	157	159	

Critério II – Rácio altura-comprimento do rectângulo do cluster

Após a filtragem por dimensões, foi analisado o rácio altura-comprimento (Rac) do rectângulo delimitador de cada *cluster*, definido pela Equação 1.

$$Rac = \frac{\text{altura rectângulo delimitador cluster}}{\text{comprimento rectângulo delimitador cluster}} \quad (1)$$

Na Figura 20 e Tabela 2 podem-se visualizar alguns rácios calculados, a título de exemplo, que permitiram estimar um intervalo de aceitação para os *clusters*, ou seja, um intervalo para o rácio que inclua a sinalização esperada e que ao mesmo tempo consiga eliminar o ruído. Desta forma, foi possível eliminar *clusters* que não continham um rácio coerente com o verificado para a sinalização que se pretendia detectar.



Figura 20 – Fragmentos de *frames* nos quais estão presentes sinais, com a indicação do respectivo rácio altura-comprimento (Rac) do rectângulo do cluster para cada situação.

Neste processo verificou-se que, a variação do Rac para um mesmo sinal numa sequência de *frames* (entre o Rac do *cluster* correspondente ao sinal no primeiro *frame* e o Rac no último *frame*) é sempre inferior a 10 %. Assim, foi considerada a gama de variação para o Rac com valores compreendidos entre 0.83 e 1.2.

Tabela 2 – Exemplo de rácios altura-comprimento (Rac) de sinalização presente no vídeo. Os sinais encontram-se identificados de acordo com a nomenclatura utilizada em [5].

Sinal	Dimensões mínimas		Dimensões máximas		Variação do Rac entre <i>frames</i>
	<i>frame</i>	Rac	<i>frame</i>	Rac	
C13	63	0.98	69	1.04	0.06
C14a	467	1.00	475	1.00	0.00
C13	469	0.96	475	0.98	0.02
B9d	531	0.88	539	0.93	0.05
D8	56	0.97	69	1.04	0.07
D1c	554	1.03	566	1.11	0.08
C13	581	1.03	590	1.07	0.05
C13	602	1.02	610	1.03	0.01
A1a	605	0.86	610	0.93	0.06
	Mínimo	0.86	Mínimo	0.93	
	Máximo	1.03	Máximo	1.11	

Critério III – Proximidade do *cluster* com os limites da imagem

O terceiro critério implementado foi o da eliminação de todos os *clusters* que contactassem com os limites do *frame*. Este critério foi implementado pois, mesmo que um *cluster* represente um sinal de trânsito, se este estiver em contacto com os limites do *frame*, significa que o sinal em questão apenas foi captado parcialmente. Estas situações foram excluídas, pois em condições normais os sinais em questão já teriam sido detectados (nos *frames* anteriores) e, para as fases posteriores (classificação e reconhecimento), estas seriam situações previsivelmente problemáticas. Um exemplo poderá ser consultado na Figura 21.



Figura 21 – Exemplo de um sinal/*cluster* rejeitado devido a estar em contacto com os limites da imagem.

Critério IV – Posicionamento relativo entre o centróide do *cluster* e o centro geométrico do rectângulo do *cluster*

O centróide do *cluster*, no caso particular do *cluster* corresponder a um sinal de trânsito, será determinado única e exclusivamente pela zona colorida do sinal. Na SVT (Sinalização Vertical de Trânsito) que se pretende detectar, essa zona colorida corresponde à orla do sinal, ou ao sinal em si. Assim, o pictograma presente nos sinais, independentemente da sua cor, não afectará o posicionamento do centróide. Foi, desta forma, identificado o posicionamento esperado do centróide para a sinalização a detectar, Figura 22.

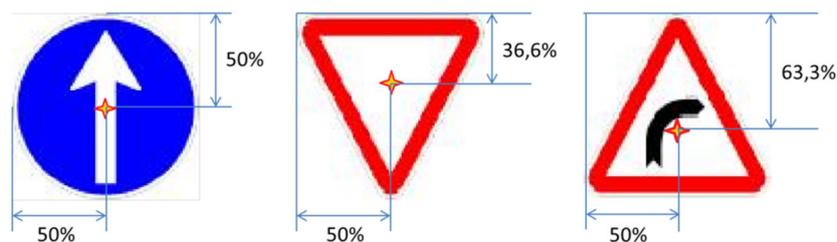


Figura 22 – Posicionamento esperado para o centróide dos sinais circulares e triangulares (normais e em posição inversa).

De seguida, foi identificada uma área em torno do centro geométrico do rectângulo do *cluster* e aferido o posicionamento do centróide do *cluster* correspondente. No caso de o centróide não ter sido encontrado no interior dessa área, o referido *cluster* foi rejeitado. Desta maneira, foi possível eliminar todos os *clusters* cujo posicionamento do seu centróide não fosse coerente com os verificados para a sinalização a detectar, conforme ilustrado na Figura 22 e Figura 23.

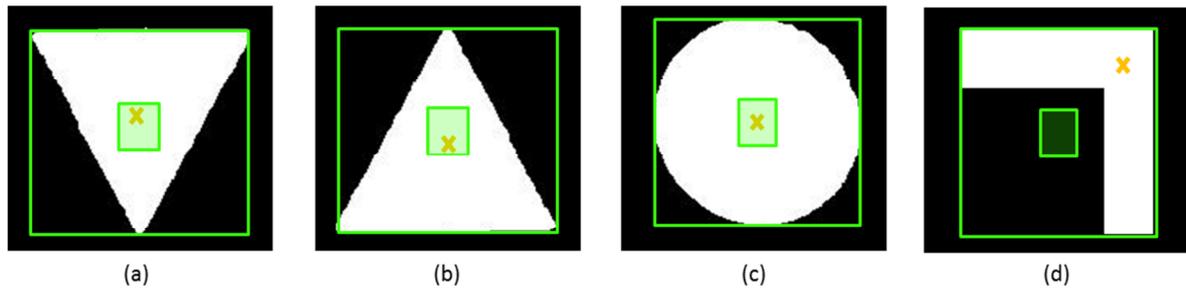


Figura 23 – Esquema exemplificativo do critério de aceitação/rejeição de um *cluster* baseado no posicionamento do seu centróide numa área circundante do centro geométrico do seu rectângulo para quatro situações. Nas situações (a), (b) e (c) o *cluster* é aceite. Na situação (d) o *cluster* é descartado.

Critério V – Rácio de preenchimento do rectângulo do cluster

O último critério de análise dos *clusters* baseia-se num rácio de preenchimento do rectângulo do *cluster* (Rprc). Para tal, é efectuado um preenchimento de todos os *clusters* que contenham “buracos” no seu interior. Assim, garante-se que todos os *clusters* que irão ser analisados são constituídos por uma forma totalmente preenchida, o que irá ser usado para calcular o rácio em questão, consultar Figura 24.



Figura 24 – *Clusters* fechados, com “buracos” no seu interior (identificados com a seta amarela) e respectivo resultado obtido depois de efectuado processo de preenchimento dos “buracos”.

Foi calculado o Rprc tendo por base a imagem binária totalmente preenchida. O Rprc consta na quantificação da proporção existente entre o número de *pixels* pertencentes ao *cluster* e o número total de *pixels* delimitados pelo rectângulo do *cluster*, Equação 2. Os rácios calculados foram, de seguida, analisados e todos os que não se apresentassem coerentes com os rácios pertencentes à sinalização, presentes na Tabela 3, foram descartados.

$$R_{prc} = \frac{\sum \text{pixels pertencentes ao cluster preenchido}}{\sum \text{pixels pertencentes ao rectângulo delimitador do cluster}} \quad (2)$$

Tabela 3 – Rácios de preenchimento do rectângulo de *cluster* (Rprc) esperados para sinalização do tipo circular, triangular e triangular invertida.

Tipo Sinal	Rprc
Circular	0.78
Triangular	0.52
Triangular invertido	0.52

Em suma, uma vez descartados todos os *clusters* que não respeitem cada um destes cinco critérios, obtêm-se os *clusters* que correspondem a Regiões de Interesse (RdI).

4.1.3 Resultados Obtidos

Para testar a fase de detecção implementada foi criada uma amostra de alguns *frames* do vídeo original. Esta amostra foi escolhida de maneira a conter não só zonas consideradas ideais (com boas condições de iluminação, posicionamento dos sinais e com poucos elementos previsivelmente problemáticos), mas também zonas sensíveis e/ou problemáticas (com elementos que poderiam levar a detecções erradas, bem como sinais que se encontravam em condições particularmente difíceis devido a sombras, posicionamento e/ou orientação).

Assim, a amostra seleccionada e com a qual foram realizados os testes, que originaram os resultados a seguir expostos, foi constituída por 987 *frames*. Nesta amostra foi identificada previamente, de forma manual, a SVT existente que, conforme se pode consultar no anexo A, consiste num total de 64 sinais. Dos 64 sinais presentes na amostra cerca de 83 % de SVT é de coloração vermelha e 17 % de SVT é de coloração azul (53 sinais de coloração vermelha e 11 de coloração azul).

Na Tabela 4 estão presentes os resultados obtidos para a fase detecção. Nesta, foram correctamente identificados 57 dos 64 sinais presentes na amostra, o que corresponde a uma correcta detecção de 89 % do total de sinalização presente. É de salientar no entanto que se considerarmos os resultados falsos positivos⁴ a percentagem de sucesso desta fase é de 32 % (o que corresponde a 68 % de ocorrências falsas positivas). Recorde-se que nesta fase estes falsos positivos não são relevantes,

⁴ Falsos positivos: RdI que foram detectadas e que não correspondem efectivamente a sinalização.

uma vez que, posteriormente, as RdI obtidas serão alvo de tratamento na fases de classificação e reconhecimento, o que permitirá descartar grande parte destas ocorrências dos resultados finais do método. Examinando os resultados falsos negativos⁵ denota-se que cerca de 11 % da SVT presente na amostra não foi detectada. Estas ocorrências, ao contrário dos falsos positivos, são de maior relevância pois a SVT não detectada não será alvo de análise posterior, influenciando assim a taxa de sucesso global do método implementado.

Tabela 4 – Resultados obtidos para a fase de detecção.

Fase de Detecção	Total RdI detectadas	Sinais existentes	RdI que correspondem a sinalização	Sinais não detectados	% sinais correctamente detectados	% RdI correctamente detectadas
Cor vermelha	82	53	47	6	88.7	57.3
Cor azul	96	11	10	1	90.9	10.4
Total	178	64	57	7	89.1	32.0

Analisando globalmente os resultados, verifica-se que o desempenho é superior para a sinalização de cor vermelha. Tal, deve-se à elevada percentagem de resultados falsos positivos correspondentes à cor azul (cerca de 90 % do total de RdI azuis detectadas não correspondem a sinalização). Desta forma, ignorando os resultados falsos positivos desta análise e analisando apenas a percentagem de sinalização correctamente detectada, observa-se que a taxa de sucesso é bastante semelhante para a sinalização de cor azul e vermelha, sendo no entanto esta percentagem ligeiramente superior para a sinalização de cor azul (cerca de 89 % para a sinalização vermelha e 91 % para a sinalização azul).

Particularizando as situações em que a sinalização não foi correctamente detectada, verifica-se que:

- O único erro existente para a sinalização de cor azul deve-se às condições de iluminação do *frame*. Conforme se pode ver na Figura 25, o sinal encontra-se numa situação de contra luz, o que faz com que o mesmo não seja abrangido pelo passo de segmentação por cor inicial.

⁵ Falsos negativos: SVT que não foi detectada.



Figura 25 – Sinalização de cor azul não detectada devido às condições de iluminação (contra luz).

- Os erros existentes para a sinalização de cor vermelha, devem-se a diversas razões, ilustradas na Figura 26 e descritas de seguida:
 - No caso do sinal “stop” presente na situação (a), este não cumpriu os critérios I e II de aceitação de *clusters* (dimensões do rectângulo do *cluster* e rácio altura-comprimento do rectângulo do *cluster*, respectivamente), o que levou à não detecção do mesmo.
 - Para a situação apresentada em (b), não foi cumprido o critério IV (posicionamento relativo entre o centróide do *cluster* e o do rectângulo do *cluster*) devido à presença de ruído na imagem, o que influenciou o posicionamento do centróide do *cluster*.
 - Nas situações apresentadas em (c), partes da sinalização não foram detectadas no passo inicial de segmentação da imagem por cor. Desta forma o *cluster* correspondente às mesmas não representou uma forma fechada, tendo assim sido excluído da análise.

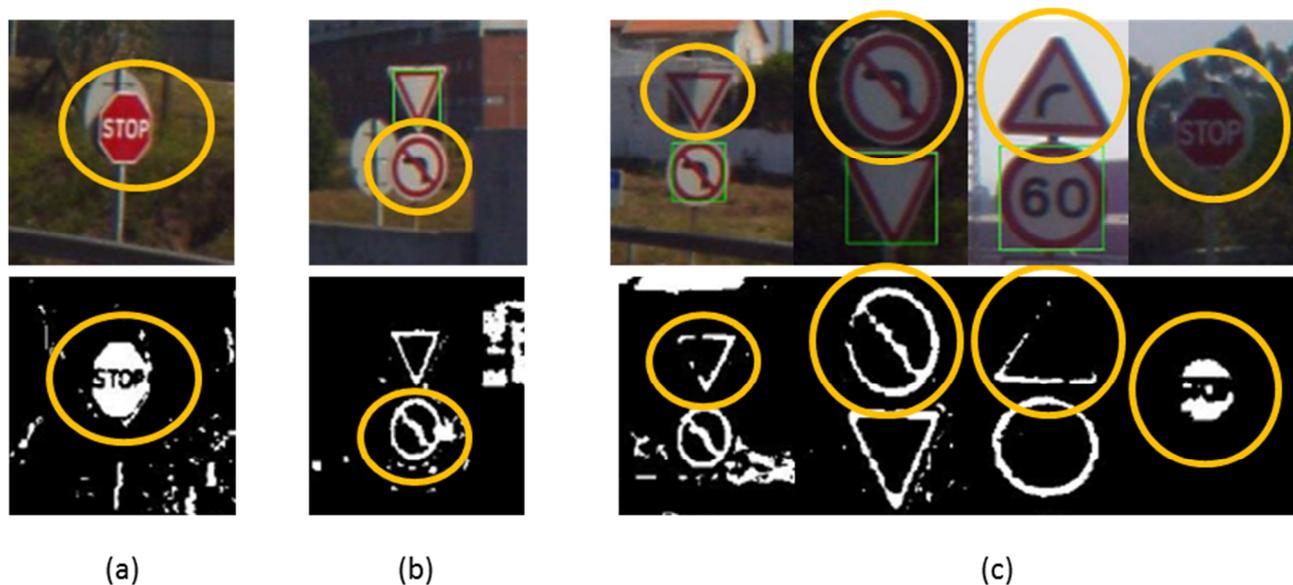


Figura 26 – Sinalização de cor vermelha não detectada. (a) Devido ao desrespeito dos critérios I e II de aceitação de *clusters*. (b) devido ao desrespeito pelo critério IV de aceitação de *clusters*. (c) devido às condições de iluminação.

Problemas Encontrados

Apesar de, numa primeira análise, a detecção de um sinal parecer uma tarefa relativamente simples de realizar num ambiente controlado, o mesmo não se poderá idealizar num ambiente real, onde existem um grande número de adversidades que dificultam esta tarefa. Assim, foram detectadas algumas situações que dificultaram a correcta detecção de alguma sinalização, ou que levaram à detecção errada de outros elementos que não são sinalização. Essas situações foram:

- Variação nas condições de iluminação
- Oclusão/obstrução da sinalização
- Presença de elementos com cor e características semelhantes à da sinalização

Variação nas condições de iluminação

Tendo em conta as características reflectivas da sinalização, facilmente se consegue depreender que as condições e tipo de iluminação são preponderantes na tonalidade da sua coloração. Assim, a coloração que um determinado sinal apresenta sofre uma grande variação consoante este sinal esteja localizado numa zona de sombra, ou sob iluminação directa, Figura 27.



Figura 27 – Exemplo da influência que diferentes condições de iluminação têm na coloração da sinalização.

Outro factor a ser tido em conta é a variação de iluminação ao longo do dia, consoante a posição relativa existente entre a sinalização e o sol, nomeadamente condições em que a sinalização se encontra em contra luz, Figura 28.



Figura 28 – Exemplo da influência que o posicionamento do sol tem relativamente à coloração da sinalização.

Estas mudanças de intensidade na iluminação levam, não só, a que as cores fiquem mais escuras ou mais claras, mas também podem alterar o tom da cor da sinalização. Esta mudança de tom dificulta o processo de detecção da sinalização, uma vez que é necessário definir uma gama de variação na busca por determinada cor. Assim, se a gama de variação considerada for demasiado abrangente, é originado ruído, o que aumenta a ocorrência de resultados falsos positivos e pode inclusivamente inviabilizar a detecção de sinalização. Se, por outro lado, a gama de variação considerada for demasiado restritiva não é detectada muita da sinalização que se encontra sob condições de iluminação menos favoráveis.

Por conseguinte, torna-se necessário encontrar um compromisso na gama de variação da tonalidade de cor que permita detectar a sinalização sob piores condições de iluminação e que, ao mesmo tempo, não inclua demasiado ruído.

Oclusão/obstrução da sinalização

Em algumas situações reais, a SVT encontra-se obstruída por diversos elementos, sejam eles de carácter temporário ou permanente. Frequentemente, nestas ocasiões, é inviabilizada a detecção da sinalização presente. Assim, são apresentadas algumas situações em que não foi possível detectar a sinalização devido à oclusão da mesma, consultar Figura 29.



Figura 29 – Situações em que a sinalização se encontra ocluída por diversos elementos presentes na via, impossibilitando a sua detecção.

Presença de elementos com cor e características semelhantes às da sinalização

Na maioria das vezes os sinais estão desobstruídos. Contudo, há situações em que a detecção se torna complexa devido ao ruído presente na sua envolvente. Existem diversas situações em que a detecção é dificultada devido à presença de elementos estranhos, elementos cuja cor e forma se assemelham a SVT. Estes elementos são, maioritariamente, veículos ou partes de veículos, painéis publicitários, terreno ou vegetação, estruturas ou edifícios. Assim, na Figura 30, estão explícitas situações onde alguns elementos presentes dificultaram o processo de detecção da sinalização e/ou levaram a detecção de resultados falsos positivos, principalmente devido à sua coloração.

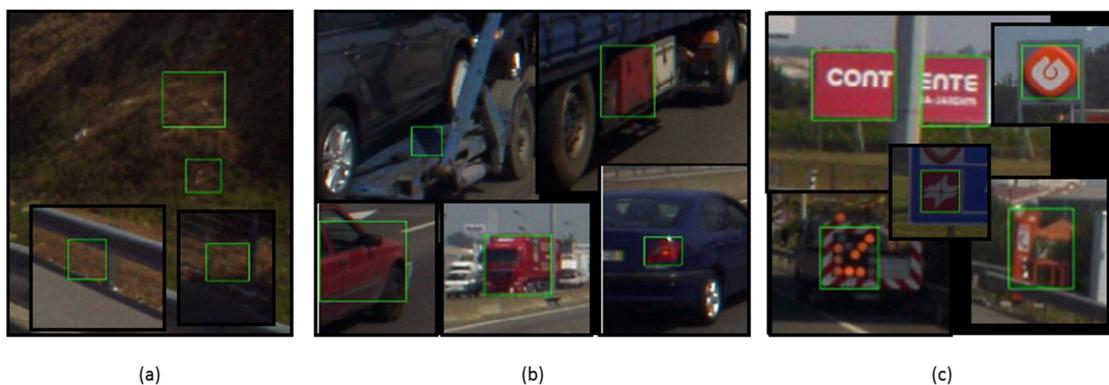


Figura 30 – Situações de falsos positivos provocados por *clusters* com características semelhantes às da sinalização presentes em: (a) terreno / vegetação, (b) veículo ou partes integrantes dos mesmos, (c) sinais publicitários ou outros sinais que não se pretendiam detectar.

4.2 Classificação da Sinalização

A fase de classificação consiste na agregação, por classes, de cada uma das regiões de interesse (RdI) identificadas na fase de detecção. Esta agregação consiste na divisão das RdI consoante a sua cor (vermelho ou azul) e a sua forma geométrica (triangular ou circular), consultar Figura 31. Tendo em conta que a segmentação por cor já tinha sido realizada na fase de detecção, restava efectuar o reconhecimento da forma geométrica presente em cada uma das RdI. Esse processo irá ser aprofundado no subcapítulo 4.2.1.

É possível agregar as RdI de acordo com as seguintes classes:

- Triângulo em posição normal Vermelho
- Triângulo em posição invertida Vermelho
- Círculo Vermelho (onde está incluído o sinal de “Stop”)
- Círculo Azul

De todas as classes, o “Triângulo em posição invertida Vermelho” é a classe que se distingue das restantes, pois contém apenas um sinal que a possa integrar, o sinal de “Cedência de passagem”. Desta forma, para esta classe em particular, a fase de classificação, é também de reconhecimento, pois é possível identificar inequivocamente o sinal em questão, apenas através da sua cor e forma geométrica.

Outra classe que é de salientar é a de “Círculo Vermelho”. Nesta classe, para além dos sinais circulares vermelhos, está também contemplado o sinal de “Stop”. Esta opção foi a tomada unicamente por motivos de facilidade de implementação, pois segundo o processo utilizado para a detecção da forma geométrica o octógono (correspondente à verdadeira forma geométrica do sinal de “stop”) é facilmente classificável como um círculo.

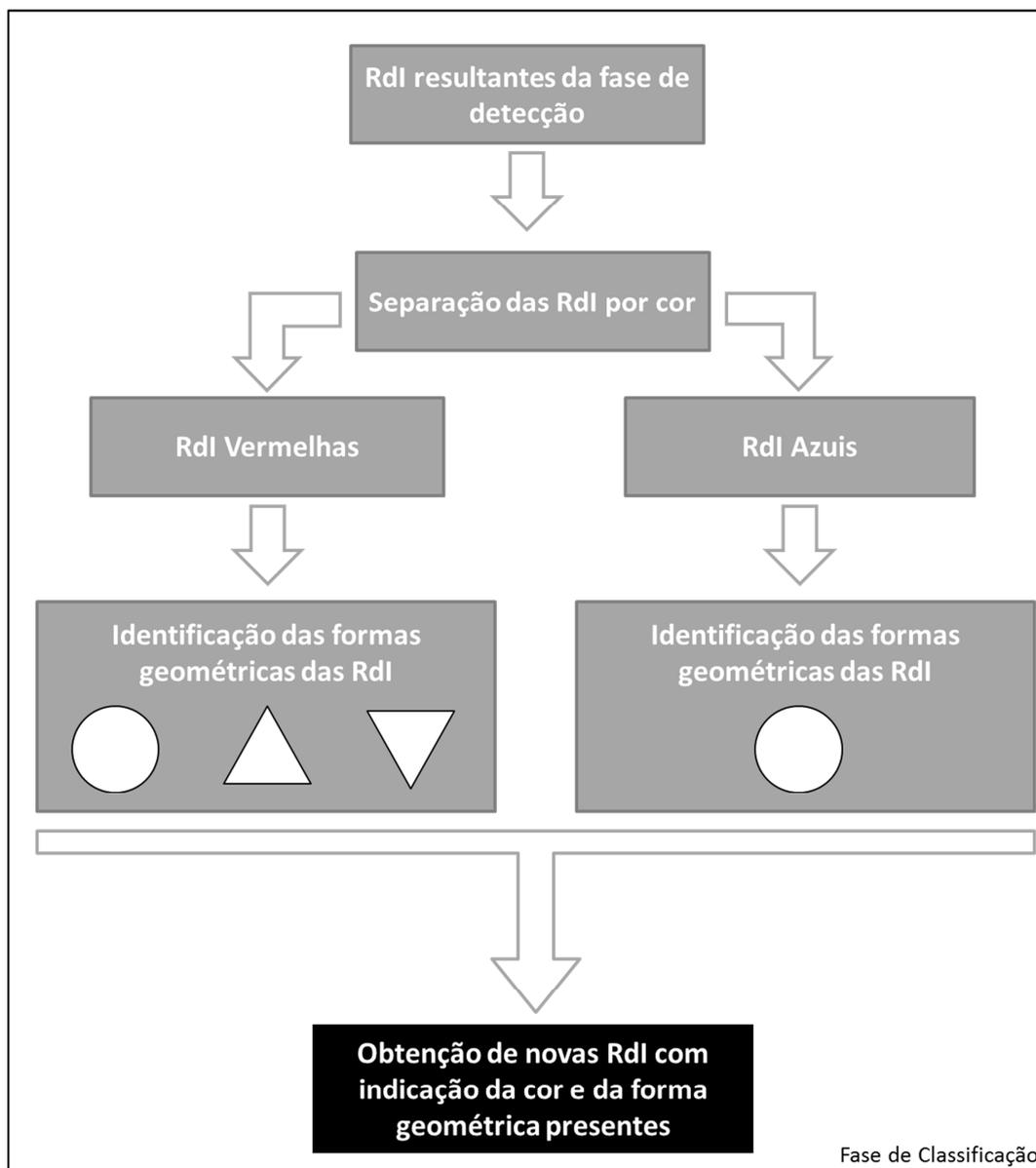


Figura 31 – Esquema dos processos que constituem a fase de classificação de sinalização do método de reconhecimento automático de sinalização implementado.

Este tipo de classificação por cor e forma geométrica foi a escolhida em detrimento de outro tipo de classificação, que à partida poderia parecer mais lógica, como a classificação presente na legislação (consultar capítulo 2). Na classificação de SVT usada em termos legais, existem sinais de características muito semelhantes em classes distintas, o que iria dificultar e tornar bastante mais complexo o processo de classificação. Alguns exemplos são apresentados na Figura 32, onde estão sinais vermelhos triangulares que, neste caso, são sinais de “Regulamentação – cedência de passagem” e “Perigo”. A única exceção em que não se verifica esta mescla de sinais, com características semelhantes distribuídos por diferentes classes, são os sinais de “Obrigações” que se distinguem dos demais devido à sua coloração azul e forma circular.



Figura 32 – Exemplos de sinalização com características de cor e forma geométrica semelhantes, mas que pertencem a diferentes classes na legislação nacional. (a) sinais pertencentes às classes de regulamentação (à esquerda) e de perigo (à direita). (b) sinais pertencentes às classes de perigo (à esquerda) e regulamentação (à direita).

Os principais objectivos desta fase de classificação consistiram em eliminar os falsos positivos e agregar correctamente os sinais detectados por classes. Esta agregação foi realizada para, na fase de reconhecimento, implementada posteriormente, ser minimizado o número de correspondências de imagens (*image matching*) a serem efectuadas. Foi, assim, possível minimizar o tempo de processamento da fase de reconhecimento, restringindo à partida o número de RdI que não correspondiam às formas geométricas de interesse. Consequentemente, foi aumentada a eficiência geral do método, não só através da diminuição do tempo de execução, como também minimizando as probabilidades da ocorrência de um reconhecimento erróneo de uma determinada RdI que não correspondesse efectivamente a SVT.

4.2.1 Identificação da Forma Geométrica

A identificação da forma geométrica realizada foi baseada nos resultados da fase anterior, a fase de detecção da SVT. Os elementos identificados como RdI foram utilizados, bem como as imagens binárias resultantes desse processo. Deste modo, cada uma das RdI foi analisada para determinar se a forma geométrica nela presente correspondia às formas geométricas de interesse. Para tal, e uma vez que se dispunha da imagem binária resultante do processo de detecção, Figura 33, esta foi tida como ponto de partida. Foram analisadas regiões pré determinadas de cada RdI, com o intuito de aferir se se estava perante uma RdI com a geometria pretendida ou não. Este processo, embora com algumas diferenças no modo de implementação, foi baseado no procedimento elaborado para detecção de forma geométrica em [24].

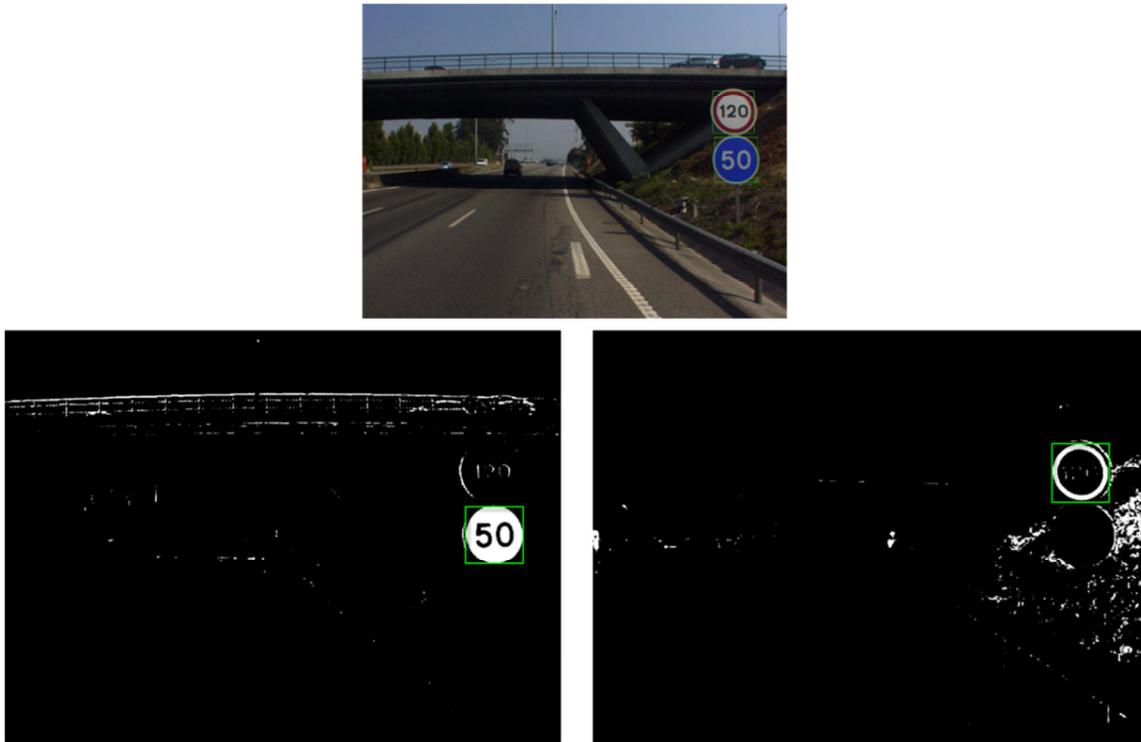


Figura 33 – Imagens binárias com as respectivas RdI, provenientes da fase de detecção, assinaladas. Imagem binária resultante da detecção de cor azul (em baixo à esquerda) e da cor vermelha (em baixo à direita).

Em primeiro lugar foi realizada a definição da localização, distribuição e dimensões das regiões da RdI a analisar. Assim, e de acordo com a Figura 34, foram encontradas oito regiões, às quais foram atribuídos diferentes pesos, que permitiam identificar as formas geométricas em questão. Esta localização, pesos e dimensões das regiões a analisar em cada RdI foram as escolhidas tendo em conta as características das formas geométricas presentes na sinalização que se pretendia detectar.

Cada RdI foi analisada (tendo por base a imagem binária) de maneira a que, caso houvesse a presença de *pixels* a branco numa região pré-determinada, o peso associado a essa região era isolado e adicionado a um parâmetro criado (ParRdI) para a RdI em questão. Assim, depois de analisadas as oito regiões da RdI, o ParRdI dessa RdI correspondia ao somatório dos pesos das regiões onde tivessem sido detectados *pixels* (a branco) pertencentes ao *cluster* da RdI (Equação 3).

$$ParRdI = \sum \text{Pesos das regiões pré determinadas contendo pixels a branco} \quad (3)$$

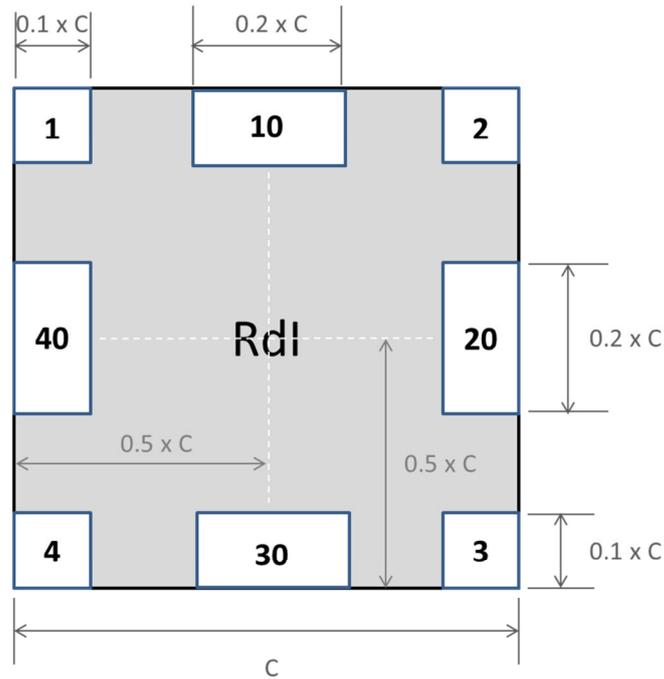


Figura 34 – Posicionamento, dimensão e peso de cada uma das oito regiões pré determinadas a analisar em cada RdI. O peso encontra-se indicado no centro de cada região.

Tornou-se assim possível identificar RdI que detivessem formas geométricas circulares ($\text{ParRdI} = 100$), quadradas ou rectangulares ($\text{ParRdI} = 110$), triangulares ($\text{ParRdI} = 47$ ou 43 , consoante a orientação do triângulo). Este procedimento é exemplificado na Figura 35.

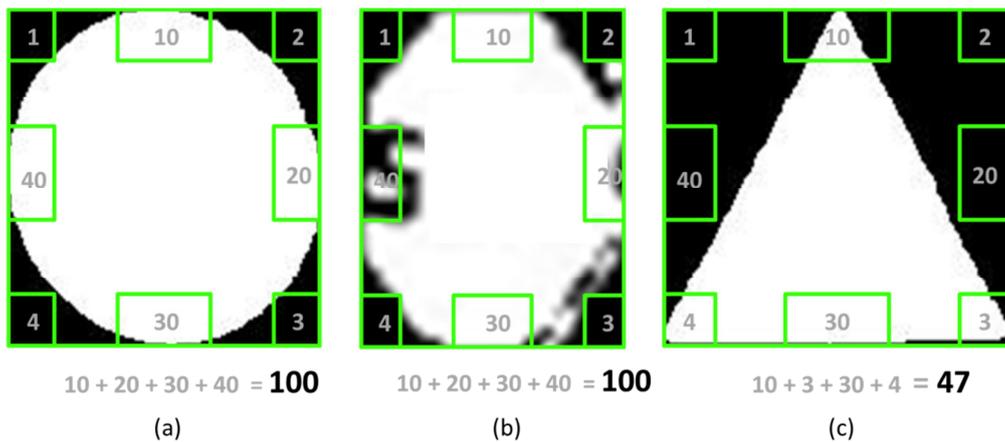


Figura 35 – Exemplo do método de identificação de forma geométrica para: (a) sinal circular; (b) sinal de “stop”; (c) sinal triangular. Com indicação do ParRdI e respectivos pesos considerados no seu cálculo.

4.2.2 Resultados Obtidos

Tendo como ponto de partida as RdI que resultaram da fase de detecção (descrita no subcapítulo 4.1.3), estas foram testadas de maneira a aferir se as formas geométricas nelas presentes seriam correspondentes às formas geométricas de interesse ou não. Nesta etapa de classificação existem dois objectivos a atingir: i) correcta classificação da sinalização detectada, e que de facto corresponde a SVT; ii) descartar o máximo número de resultados considerados falsos positivos que, conforme apresentado no subcapítulo 4.1.3, representa cerca de 68 % das RdI tidas como ponto de partida. Os resultados obtidos encontram-se explícitos na Tabela 5.

Tabela 5 – Resultados obtidos para a fase de classificação tendo por base os resultados da fase de detecção (subcapítulo 4.1.3).

Fase de Classificação		Sinais detectados	Sinais correctamente classificados	Total de RdI detectadas	Total de RdI classificadas	% sinais correctamente classificados	% sucesso da fase de classificação
Sinais Vermelhos	Circular	30	30	---	33	100.0	90.9
	Triangular normal	11	11	---	11	100.0	100.0
	Triangular invert	6	6	---	6	100.0	100.0
Totalidade Sinais Vermelhos		47	47	82	50	100.0	94.0
Sinais Azuis	Circular	10	10	96	11	100.0	90.9
Totalidade dos Sinais		57	57	178	61	100.0	93.4

Analisando estes resultados podemos constatar que a totalidade da sinalização, correctamente detectada na fase anterior, foi classificada com sucesso. Esta correcta classificação verificou-se tanto para a sinalização vermelha como para a sinalização azul. Com estes resultados é possível verificar que o primeiro objectivo desta fase foi alcançado.

Relativamente ao segundo objectivo, o de descartar os resultados considerados falsos positivos, não foi atingido na totalidade, embora o número destas ocorrências tenha sido reduzido de cerca de 68 % da totalidade das RdI para cerca de 7 % da totalidade das RdI classificadas. Assim, persistem quatro situações, três para a classe sinais vermelhos circulares e uma para a classe sinais azuis circulares, que permanecem como falsos positivos. Estas quatro situações encontram-se na Figura 36.



Figura 36 – As quatro ocorrências consideradas como falsas positivas na fase de classificação.

Analisando globalmente os resultados verifica-se que esta fase de classificação obteve uma percentagem de sucesso na ordem dos 94 % para a sinalização de cor vermelha e 91 % para a sinalização de cor azul, sendo que no total a taxa de sucesso foi de 93 %.

Problemas Encontrados

Apesar de ter sido classificada correctamente toda a sinalização, as situações apresentadas na Figura 36 constituem zonas onde se verificam resultados falsos positivos. Estas situações correspondem essencialmente a regiões que apresentam características semelhantes às da sinalização (testadas na fase de detecção, subcapítulo 4.1.3) e com forma geométrica aproximadamente circular. Analisando cada uma das situações verifica-se que:

- Existem três situações que não correspondem a nenhum tipo de sinalização, sendo duas delas (com coloração vermelha) correspondentes a zonas de terreno e a restante (com coloração azul) correspondente a parte da estrutura de um automóvel pesado.
- Existe uma situação que corresponde a sinalização, mas ao contrário do que se pretende, esta não corresponde a SVT.

Assim, estas quatro ocorrências darão entrada na fase seguinte, a fase de reconhecimento, para nessa fase serem alvo de uma análise posterior.

4.3 Reconhecimento da Sinalização

A fase de reconhecimento da sinalização consiste na identificação de cada sinal que tenha sido detectado e classificado. Esta identificação foi realizada por correspondência de imagens (*image matching*). Para tal, foi construído um repositório com *templates*⁶ da sinalização que se pretendia reconhecer, organizados e separados de acordo com as classes utilizadas na fase de classificação (ou seja por cor e forma geométrica).

Os diferentes passos que constituem a fase de reconhecimento estão esquematicamente representados na Figura 37.

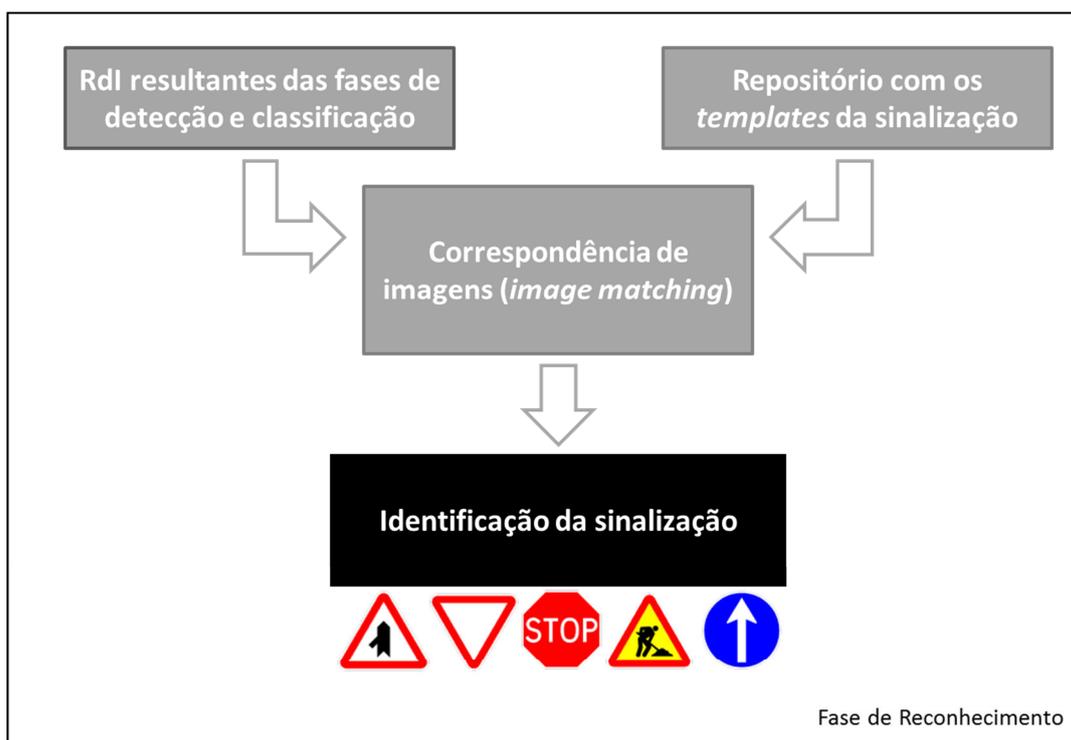


Figura 37 – Esquema dos processos que constituem a fase de reconhecimento de sinalização do método de reconhecimento automático de sinalização implementado.

Nesta fase, o primeiro passo foi obter os *templates* da sinalização, com base nos quais o reconhecimento viria a ser efectuado. Assim, os *templates* disponibilizados em [47], foram organizados num repositório segundo a cor e forma geométrica (o mesmo critério utilizado na classificação das Rdl). Após construído este repositório, foi iniciado o processo de reconhecimento, que consistiu na realização de correspondência de imagens (*image matching*) entre os *templates*

⁶ *Templates*: imagens padrão utilizadas como referência no processo de correspondência de imagens (*image matching*).

presentes no repositório e cada uma das RdI. De forma a minimizar o número de correspondências de imagens a realizar, e recordando que as RdI se encontram classificadas pelo mesmo critério usado para classificar os *templates*, cada RdI apenas efectuou correspondência com os *templates* que tivessem a mesma classificação. Ou seja, descrevendo o exemplo apresentado na Figura 38, no caso de estarmos perante uma RdI pertencente à classe “Triângulo em posição normal vermelho”, esta apenas irá fazer correspondência com os *templates* que correspondam à mesma classe.

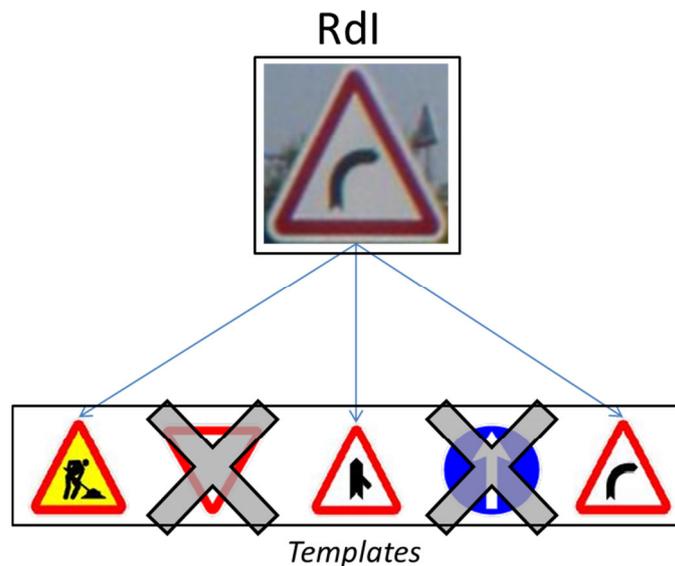


Figura 38 – A correspondência de imagens (*image matching*) é realizada somente com *templates* que tenham a mesma classe que a RdI.

Todo o processo de correspondência de imagens será aprofundado no subcapítulo 4.3.1, onde serão descritos os processos que foram ensaiados até terem sido alcançados os resultados apresentados e analisados no subcapítulo 4.3.2.

4.3.1 Correspondência de imagens (*image matching*)

O processo de correspondência de imagens consistiu na comparação de cada RdI com todos os *templates* que pertencessem à mesma classe. Esta comparação foi realizada utilizando um correlacionador que permitisse quantificar o grau de similitude entre as imagens em questão.

Tendo em consideração a revisão bibliográfica elaborada no subcapítulo 3.2.2, onde foi referido que um dos correlacionadores mais utilizados na realização de correspondência de imagens é a correlação cruzada normalizada, esta foi utilizada com o intuito de identificar a sinalização presente em cada RdI.

Identificação de SVT através de Correlação Cruzada Normalizada

De acordo com [48], a correlação cruzada normalizada (CCN) entre um *template* e uma imagem é descrita pela Equação (4).

$$CCN(u, v) = \frac{\sum_{x,y} [f(x, y) - \bar{f}_{u,v}] [t(x - u, y - v) - \bar{t}]}{\sqrt{\sum_{x,y} [f(x, y) - \bar{f}_{u,v}]^2 \sum_{x,y} [t(x - u, y - v) - \bar{t}]^2}} \quad (4)$$

Onde t representa o *template*, f a imagem, \bar{t} a média do *template* e $\bar{f}_{u,v}$ a média de $f(x, y)$ na região correspondente ao *template*, (u, v) as coordenadas correspondentes ao *template* e (x, y) as coordenadas correspondentes à imagem.

Segundo a mesma fonte, a correlação cruzada normalizada apresenta algumas vantagens relativamente a outros correlacionadores mais simples (correlação simples e correlação cruzada), tais como:

- Ser invariante a flutuações na amplitude da imagem, particularmente as causadas pela variação nas condições de iluminação.
- A gama de variação do seu valor ser independente das dimensões (u, v) da região correspondente ao *template*.
- Efectuar uma correspondência correcta nos casos em que a imagem contém uma zona com brilho elevado (*brighth spot*).

Desta forma, foi utilizada a função *normxcrr2*, presente na biblioteca de processamento de imagem do Matlab e que contém implementado um algoritmo otimizado da correlação cruzada normalizada, para efectuar a correspondência entre cada RdI e os *templates*. Para tal, os *templates* foram redimensionados para as mesmas dimensões que as apresentadas por cada RdI. Este redimensionamento foi realizado utilizando o método de interpolação bilinear e o método do vizinho mais próximo, não tendo os resultados finais sido alterados significativamente pela utilização de um e de outro método.

Contrariamente ao que seria expectável, o uso da correlação cruzada normalizada não se revelou eficaz no processo de identificação, tendo sido obtida uma reduzida taxa de sucesso, correspondente à correcta identificação de cerca de 25 % das RdI.

Assim, foi desenvolvido um método alternativo, que se baseia nos pictogramas presentes na sinalização e num simples coeficiente de correlação, para identificar quais os sinais presentes nas RdI.

Identificação de SVT através dos pictogramas e de correlação simples

Tendo em conta que a sinalização a detectar contém pictogramas únicos para cada sinal, foi desenvolvida uma metodologia de identificação da sinalização baseada nesses pictogramas. Após uma análise à sinalização que se pretendia reconhecer, foi identificado que os pictogramas existentes contém cor branca ou cor preta. Existindo uma prevalência de pictogramas de cor branca para a sinalização da classe “Círculo Azul” e de pictogramas de cor preta para as restantes classes. Assim, foram desenvolvidos processos distintos, um para a identificação dos pictogramas de cor branca e outro para os pictogramas de cor preta, ver Figura 39.

Na Figura 39 estão ilustrados os passos que foram implementados e que permitem a identificação de cada sinal, baseando-se nos pictogramas existentes.

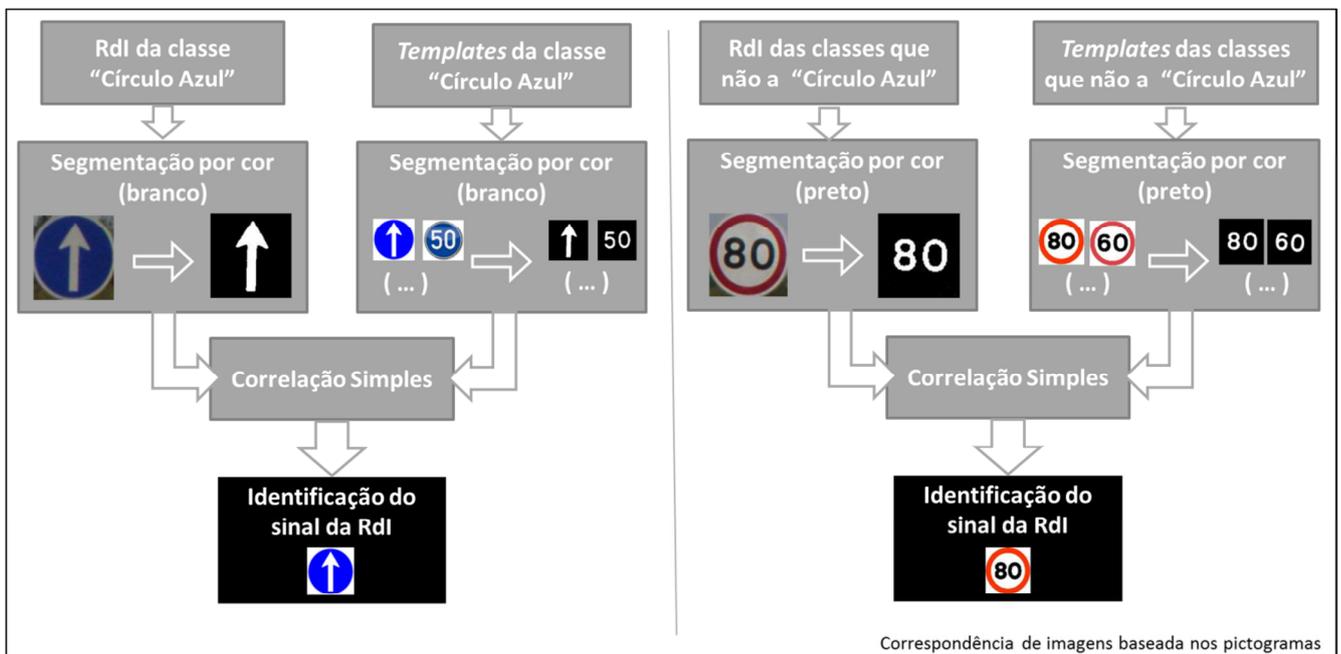


Figura 39 - Esquema dos processos que foram implementados por forma a realizar a identificação da sinalização presente nas RdI, para RdI pertencentes à classe “Círculo Azul” (à esquerda) e para RdI pertencentes às restantes classes (à direita).

Neste processo, o primeiro passo consistiu em identificar a classe da RdI sob análise. De seguida, foi realizada uma segmentação da RdI por cor. No caso de a RdI pertencer à classe “Círculo Azul” era efectuada uma busca pela cor branca, em caso contrário era efectuada uma busca pela cor preta. Esta segmentação foi realizada de acordo com as condições presentes nas Equações 5 e 6, para as cores branca e preta, respectivamente.

$$BRANCO: \quad t(x, y, 2) < 0.35 \cap t(x, y, 3) > 0.20 \quad (5)$$

$$PRETO: \quad t(x, y, 3) < \frac{t(:, :, 3)_{\max} - t(:, :, 3)_{\min}}{2.12} \quad (6)$$

Nas Equações 5 e 6, t representa a imagem no espaço cor HSV, $t(x, y, 2)$ cada *pixel* da componente Saturação (S) da imagem, $t(x, y, 3)$ cada *pixel* da componente Valor (V) da imagem, $t(:, :, 3)_{\max}$ o valor máximo presente na componente Valor (V) da imagem e $t(:, :, 3)_{\min}$ o valor mínimo presente na componente Valor (V) da imagem.

Após a segmentação da RdI ter sido efectuada, foi realizado o mesmo processo para os *templates* que tivessem a mesma classificação, sendo obtidas imagens binárias tanto para os *templates* como para a RdI. De seguida, as imagens correspondentes aos *templates* foram redimensionadas de maneira a todas as imagens binárias obtidas (da RdI e dos *templates*) ficarem com as mesmas dimensões, tendo sido efectuada o redimensionamento pelo método de interpolação bilinear. O passo seguinte consistiu em comparar as imagens de cada um dos *templates* com a imagem da RdI, por intermédio da correlação simples (CS), definida pela Equação 7.

$$CS = \frac{\sum_{x,y} [f(x, y) - \bar{f}][t(x, y) - \bar{t}]}{\sqrt{(\sum_{x,y} [f(x, y) - \bar{f}]^2) (\sum_{x,y} [t(x, y) - \bar{t}]^2)}} \quad (7)$$

Na Equação 7, f é a imagem correspondente à RdI, t a imagem correspondente ao *template*, \bar{f} a média da imagem correspondente à RdI e \bar{t} a média da imagem correspondente ao *template*.

Após terem sido correlacionados todos os *templates* com a RdI, foi considerado que o sinal presente na RdI era o correspondente ao *template* que obtivesse maior valor de correlação. De forma a descartar identificações erróneas, apenas foram consideradas correspondências com valor de correlação superior a um determinado limite. Neste caso, apenas foram consideradas como válidas correspondências que tivessem um valor de correlação superior a 3 %. Este limite foi encontrado empiricamente por método de tentativa e erro.

Após ter sido realizado o processo de identificação para a totalidade das RdI presentes nos diferentes *frames*, verificou-se que nem todas as RdI tinham sido identificadas, ou seja, a melhor correspondência efectuada, para essas RdI, tinha valor de correlação inferior ao limite considerado.

4.3.2 Resultados Obtidos

O processo de reconhecimento foi realizado para todas as RdI resultantes das fases de detecção e classificação. Através da sua aplicação pretendem-se alcançar dois objectivos: i) realizar a correcta identificação das RdI que correspondam a SVT; ii) descartar todas as RdI que não correspondam a SVT e que são, portanto, resultados falsos positivos.

Desta forma, após a aplicação deste processo à amostra testada, foram obtidos os resultados que se encontram na Tabela 6.

Tabela 6 – Resultados obtidos para a fase de reconhecimento, tendo por base os resultados da fase de detecção (subcapítulo 4.1.3) e da fase de classificação (subcapítulo 4.2.2).

Fase de Reconhecimento		RdI que contêm sinalização	RdI correctamente reconhecidas	RdI correctamente reconhecidas (%)	RdI Falsas Positivas da fase anterior	RdI Falsas Positivas que foram descartadas	RdI Falsas Positivas correctamente descartadas (%)
Sinais Vermelhos	Circular	30	27	90.0	3	3	100.0
	Triangular normal	11	9	81.8	0	0	----
	Triangular invertido	6	6	100.0	0	0	----
Totalidade Sinais Vermelhos		47	42	89.4	3	3	100.0
Sinais Azuis	Circular	10	10	100.0	1	1	100.0
Totalidade dos Sinais		57	52	91.2	4	4	100.0

Analisando estes resultados verifica-se que, relativamente às RdI falsas positivas, estas conseguiram ser correctamente eliminadas na sua totalidade, não existindo portanto, no somatório das três fases do método desenvolvido (detecção, classificação e reconhecimento), resultados falsos positivos. Assim, podemos afirmar que o segundo objectivo foi alcançado.

Focando o primeiro objectivo, verifica-se que foram correctamente reconhecidas cerca de 91 % das RdI. Sendo que os 9 % de RdI não reconhecidas, pertencem a sinalização das classes “Círculo Vermelho” e “Triângulo em posição normal Vermelho”, tendo, esta última classe sido a que obteve a pior taxa de reconhecimento, com cerca de 81 % das RdI correctamente identificadas. As RdI pertencentes às restantes classes obtiveram uma taxa de reconhecimento de 100 %, ou seja, foram correctamente identificadas todas as RdI correspondentes a sinalização.

Examinando um pouco mais aprofundadamente as RdI que correspondem a sinalização que não foi reconhecida, conforme apresentado na Figura 40, verifica-se que:

- O sinal presente na situação (a), é erradamente reconhecido como sendo um sinal de limite de velocidade de 80 km/h.
- O sinal presente na situação (b), é erradamente reconhecido como sendo um sinal de limite de velocidade de 60 km/h.
- Os sinais presentes nas situações (c) e (d), não são reconhecidos, ou seja a correlação máxima encontrada com um *template* é inferior ao valor mínimo admissível.

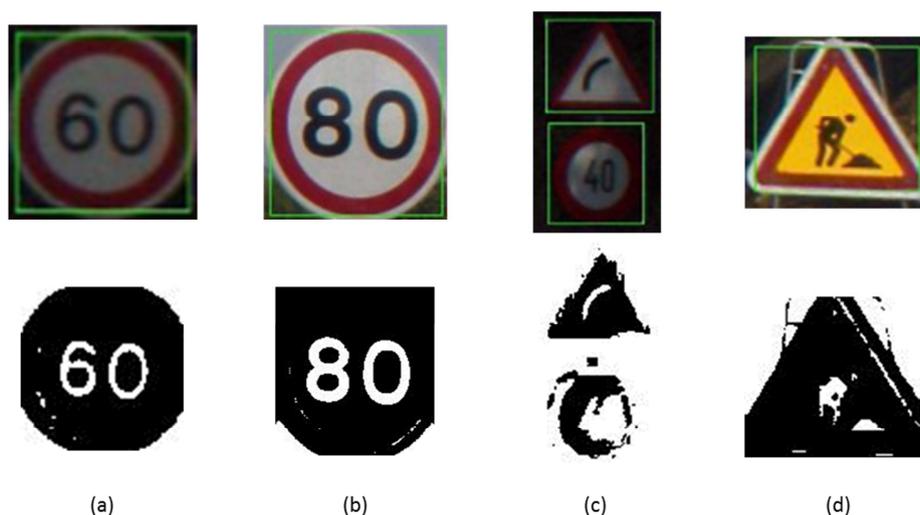


Figura 40 – Erros encontrados na fase de reconhecimento. Em cima encontram-se as RdI e em baixo as respectivas imagens binárias utilizadas na correspondência com os *templates*. As situações (a) e (b) são RdI que não foram correctamente identificadas, enquanto (c) e (d) são RdI que não foram identificadas.

Principais Dificuldades

As cinco RdI que foram erradamente identificadas ou não identificadas, apresentadas na Figura 40, correspondem à principal dificuldade encontrada. As situações (a) e (b) correspondem a uma troca de identificações entre o sinal de limites de velocidade de 60 e 80 km/h. Esta troca deve-se à existência de um elevado grau de similitude entre os pictogramas, levando a que os coeficientes de correlação entre os *templates* dos sinais de limite de velocidade de 60 e 80 km/h sejam muito próximos. As situações (c) e (d) correspondem a situações em que as RdI não foram identificadas. Esta não identificação deve-se às condições de iluminação presentes. Na situação (c) as RdI encontravam-se demasiado escuras (em zona de sombra), tendo o processo de segmentação da cor preta incluído zonas que não a do pictograma, o que influenciou o processo de correspondência com os *templates* e

levou esta a apresentar valor de correlação abaixo do limite de 3 %. Na situação (d), apesar de a RdI estar bem iluminada, a elevada reflectância da cor amarela prejudica a percepção da cor preta, o que leva a que o pictograma não seja detectado na sua totalidade e, por sua vez, o valor da correspondência fica abaixo do limite considerado.

5. Conclusão

Numa primeira abordagem à tarefa de reconhecimento automático de SVT, é usual a ideia de esta ser uma tarefa de fácil execução. Contudo, esta tarefa revela-se mais complexa do que se julgaria à partida, uma vez que as dificuldades encontradas são em grande número e algumas têm uma resolução complexa. Uma das consequências da complexidade desta tarefa é o facto de ainda não existir na comunidade científica um método que reúna consenso [7].

Apesar das dificuldades inerentes à tarefa, foi possível desenvolver um método para o reconhecimento automático da SVT. Este método foi implementado em três fases distintas, de detecção, de classificação e de reconhecimento.

A fase de detecção, consiste em detectar regiões de interesse (RdI) que possam conter sinalização. Esta detecção é feita utilizando um processo de segmentação por cor (vermelha e azul) e utilizando cinco critérios de selecção (dimensões da RdI, rácio altura-comprimento da RdI, proximidade da RdI aos limites da imagem, posicionamento relativo entre o centróide da RdI e centróide da área segmentada e rácio de preenchimento da RdI).

A fase de classificação, que tem por base as RdI obtidas na fase anterior e consiste no reconhecimento da forma geométrica de cada região, bem como na agregação por classes de cada uma dessas RdI analisadas. As classes utilizadas baseiam-se na cor e forma geométrica.

A fase de reconhecimento, consiste na identificação da sinalização que tenha sido detectada e classificada e baseia-se na correspondência dos pictogramas presentes nos sinais. Através de um processo de segmentação por cor é extraído o pictograma para de seguida ser feito o seu reconhecimento, através da correspondência entre o pictograma extraído e pictogramas *template*. Esta correspondência foi realizada, numa primeira abordagem, usando a correlação cruzada normalizada que, por não se revelar particularmente eficaz, foi excluída em detrimento da correlação simples, que apresentou resultados mais consistentes.

Embora não se tenha alcançado um total sucesso na detecção, classificação e reconhecimento dos sinais da amostra considerada, o valor obtido é considerado como satisfatório tendo em conta a simplicidade do método implementado.

A fase de detecção apresentou uma taxa de sucesso de 32 %, a fase de classificação teve uma taxa de sucesso de 93 % e a fase de reconhecimento teve uma taxa de sucesso de 91 %. A taxa de sucesso global obtido pelo método RASVT implementado é de 81 %, ou seja, da sinalização presente na amostra analisada, 81 % foi correctamente detectada, classificada e reconhecida.

Relativamente a resultados falsos positivos a sua existência não se verifica, ou seja não existem zonas que tenham sido assinaladas que não correspondam a sinalização.

Podemos ainda concluir que os factores que maior influência têm sobre os resultados finais obtidos são:

- Condições de iluminação da SVT (maioritariamente situações de sombra e contra luz).
- Oclusão/obstrução da SVT por elementos presentes na via (postes de iluminação, veículos, vegetação, etc).
- Presença de elementos externos com cor e forma geométrica semelhante à da sinalização (principalmente painéis publicitários, veículos ou partes integrantes destes, zonas de terreno ou de vegetação).

A nível de desempenho computacional, a fase de detecção foi a fase preponderante, tendo sido a que mais tempo necessitou para ser realizada, demorando, para a amostra sob análise, uma média de 5 segundos por cada *frame* analisado. As restantes fases, por utilizarem apenas as RDI provenientes da fase de detecção, e não a totalidade dos *frames*, revelaram-se mais rápidas, sendo a média por *frame* de apenas algumas fracções de segundo. Quando comparados os tempos de execução das três fases, pode-se afirmar que os tempos de execução das fases de classificação e reconhecimento são desprezáveis.

Tal como a maioria dos métodos desenvolvidos até aos dias de hoje, este método não pode ser generalizado, pois foi desenvolvido com base num certo tipo de amostra com características específicas, como por exemplo, as condições climáticas em que os dados foram obtidos, as condições de iluminação, o tipo de via e o tipo de sinalização presente.

6. Referências

- [1] “Instituto Nacional de Infra-Estruturas Rodoviárias,” [Online]. Available: <http://www.inir.pt/portal/RedeRodovi%C3%A1ria/Concession%C3%A1rias/tabid/71/language/pt-PT/Default.aspx>. [Acedido em 19 04 2013].
- [2] D. d. República, “Resolução da Assembleia da República nº 94-A/2009,” 2009. [Online]. Available: <http://dre.pt/pdf1sdip/2009/09/18801/0000200002.pdf>. [Acedido em 23 Janeiro 2013].
- [3] “Organização das Nações Unidas , ONU,” [Online]. Available: http://treaties.un.org/pages/ViewDetailsIII.aspx?&src=TREATY&mtdsg_no=XI~B~20&chapter=11&Temp=mtdsg3&lang=en. [Acedido em 23 Abril 2013].
- [4] “Imagem online,” [Online]. Available: http://en.wikipedia.org/wiki/Vienna_Convention_on_Road_Signs_and_Signals. [Acedido em 23 Abril 2013].
- [5] D. d. República, “Regulamentação de Sinalização de Trânsito - Decreto Regulamentar n.º 22-A/98,” 1 Outubro 1998. [Online]. Available: http://www2.cm-evora.pt/Transito/01regulamento%20de%20sinalizacao%20DL%2022_A_98.pdf. [Acedido em 22 Janeiro 2013].
- [6] S. Escalera, X. Baró, O. Pujol, J. Vitrià e P. Radeva, “Background on Traffic Sign Detection,” em *Springer, Traffic-Sign Recognition Systems, SpringerBriefs in Computer Science 2011*, pp 5-13, 2011.
- [7] K. Brkic, “An overview of traffic sign detection methods,” 2010. [Online]. Available: http://www.fer.unizg.hr/_download/repository/BrkicQualifyingExam.pdf. [Acedido em 15 Abril 2013].
- [8] K. Chourasia e J. Chourasia, “A Review and Comparative Analysis of Recent Advancements in Traffic Sign Detection and Recognition Techniques,” 2011. [Online]. Available: [http://www.smslucknow.com/test/fckeditor/file/all%20pdf/Archives/Vol%202\(1\)/Vol%202\(1\)-2011-P27-P34.pdf](http://www.smslucknow.com/test/fckeditor/file/all%20pdf/Archives/Vol%202(1)/Vol%202(1)-2011-P27-P34.pdf). [Acedido em 24 Janeiro 2013].
- [9] H. Shojania, “Real-time traffic sign detection,” 2000. [Online]. Available: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.90.3928&rep=rep1&type=pdf>. [Acedido em 2 Fevereiro 2013].

- [10] “Imagem online,” [Online]. Available: http://upload.wikimedia.org/wikipedia/commons/8/83/RGB_Cube_Show_lowgamma_cutout_b.png. [Acedido em 6 Abril 2013].
- [11] D. Ghica, S. Lu e X. Yuan, “Recognition of traffic signs by artificial neural network,” em *IEEE International Conference on Neural Networks, Nov/Dec 1995 (Volume: 3 Pages: 1444 -1449)*, Perth, W.A., 1995.
- [12] M. Bénallal e J. Meunier, “Real-time colour segmentation on road signs,” em *Canadian Conf. on Electrical and Computer Engineering (IEEE CCECE), vol.3, pp.1823-1826*, Montréal, Canada, 2003.
- [13] M. M. Zadeh, T. Kasvand e C. Y. Suen, “Localization and recognition of traffic road signs for automated vehicle control systems,” em *Proc. SPIE 3207, Intelligent Transportation Systems, 272 (January 27, 1998); doi:10.1117/12.300865*;
- [14] D. Kang, N. Griswold e N. Kehtarnavaz, “An invariant traffic sign recognition system based on sequential color processing and geometrical transformation,” em *IEEE, Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation, pp.88-93, 21-24 Apr 1994*, Dallas, Texas, 1994.
- [15] A. Broggi, P. Cerri, P. Medici, P. Porta e G. Ghisio, “Real time road signs recognition,” em *Intelligent Vehycles Symposium, IEEE, pp.981-986*, Istanbul, 2007.
- [16] H. Cheng, H. Jiang, Y. Sun e J. Wang, *Color Image Segmentation: Advances & Prospects*, ScienceDirect, Pattern Recognition, Volume 34, Issue 12, December 2001, Pages 2259–2281.
- [17] H. Wang e D. Suter, “Color Image Segmentation Using Global Information and Local Homogeneity,” em *VIIth Conference on digital image computing: techniques and applications, (pages:89-98)*, 2003.
- [18] “Imagem online,” [Online]. Available: http://upload.wikimedia.org/wikipedia/commons/a/a0/Hsl-hsv_models.svg. [Acedido em 6 Abril 2013].
- [19] “Microsoft,” [Online]. Available: <http://msdn.microsoft.com/en-us/library/windows/desktop/aa511283.aspx>. [Acedido em 30 Abril 2013].
- [20] A. Escalera, J. Armingol e M. Mata, “Traffic sign recognition and analysis for inteligent vehicles,” em *Science Direct, Image and Vision Computing, Volume 21, Issue 3, 1 March 2003, Pages 247–258*.

- [21] Y. Fang, S. Chen e C. Fuh, "Road sign detection and tracking," em *IEEE Transactions on Vehicular Technology*, (Volume: 52 , Issue: 5, Pages: 1329-1341), September 2003.
- [22] P. Paclík, J. Novovicova, P. Pudil e P. Somol, "Road sign classification using laplace kernel classifier," em *Pattern Recognition Letters* (Volume:21 Number:13-14 Pages: 1165-1173), 2000.
- [23] G. Loy e N. Barnes, "Fast shape-based road sign detection for a driver assistance systems," em *IEEE/RSJ Internationall Conference on Inteligent Robots and Systems, 2004* (Volume:1 Pages: 70-75).
- [24] C. Paulo e P. Correia, "Automatic detection and classification of traffic signs," em *Eighth International Workshop on Image Analysis for Multimedia Interactive Services, 2007. WIAMIS '07. (Page: 11)*, 2007.
- [25] C. Harris e M. Stephens, "A combined corner and edge detector," em *Proceedings of Fourth Alvey Vision Conference, 1988* (Pages: 147-151).
- [26] P. Parodi e P. G., "A feature-based recognition scheme for traffic scenes," em *Proceedings of the Intelligent Vehicles '95 Symposium, 1995* (Pages:229-234), 1995.
- [27] G. Piccioli, E. Micheli, P. Parodi e M. Campani, "Robust road sign detection and recognition from image sequences," em *Proceedings of the Intelligent Vehicles '94 Symposium, 1994* (Pages: 278-283).
- [28] J. Canny, "A Computacional Approach to Edge Detection," em *IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 8 Issue 6, June 1986* (Pages: 679-698) .
- [29] H. Sandoval, T. Hattor, S. Kitagawa e Y. Chigusa, "Angle-dependent edge detection for traffic signs recognition," em *Proceedings of the IEEE Intelligent Vehicles Symposium, 2000. IV 2000* (Pages:308-313).
- [30] T. Hibi e M. Kamata, "Extraction and recognition of road sign region from natural color image, by using novel HSL and coordinates transformation," em *IEEJ Transactions on Industry Applications; VOL.115; NO.12; PAGE.1484-1490;*, 1995.
- [31] G. Piccioli, M. Micheli, P. Parodi e M. Campani, "Robust method for sign detection and recognition," em *Image and Vision Computing, Volume 14, Issue 3, April 1996, Pages 209–223*.
- [32] S. Vitabile e F. Sorbello, "Pictogram road signs detection and understanding in outdoor scenes," em *SPIE Volume:3364, Pages:359-370, Enhanced and Synthetic Vision 1998*.
- [33] S. Vitabile, G. Pollacia, G. Pilato e F. Sorbello, "Road Sign Recognition using a dynamic pixel agregation technique in the HSV color space," em *11th International Conference on Image*

Analysis and Processing, 2001 (Pages: 572-577).

- [34] H. Ohara, I. Nishikawa, S. Miki e N. Yabuki, "Detection and recognition of road signs using simple layered neural network," em *IEEE, 9th International Conference Neural Information Processing, 2002 (Volume:2 Pages:626-630).*
- [35] M. Shirvaikar, "Automatic detection and interpretation of road sign," em *36th Southeastern Symposium on System Theory, 2004 (Pages:413 - 416).*
- [36] D. Kellmeyer e H. Zwahlen, "Detection of highway warning signs in natural video images using color image processing and neural networks," em *IEEE International Conference on Neural Networks, 1994. IEEE World Congress on Computational Intelligence., 1994 (Volume:7 Pages: 4226-4231).*
- [37] M. Nakamura, S. Kodama, T. Jimbo e M. Umeno, "Searching and recognition of road signpost using ring detection network," em *IEEE International Conference on Systems, Man, and Cybernetics, 1999 (Volume:2 Pages:190-195).*
- [38] R. Timofte, K. Zimmermann e L. Van Gool, "Multi-view traffic sign detection, recognition and 3d localization," em *Workshop on Applications of Computer Vision (WACV), 2009 (Pages:1-8).*
- [39] P. Viola e M. Jones, "Robust real-time object detection," em *International Journal of Computer Vision, 2001.*
- [40] Y. Freund e R. E. Schapire, "A Short Introduction to Boosting," em *Journal of Japanese Society for Artificial Intelligence, September, 1999. (Volume:14 Issue:5 Pages:771-780).*
- [41] L. Guibert, Y. Petillot e J. Tochnaye, "Real-time demonstration on an on-board nonlinear joint transformation correlator system," em *SPIE, Opt. Eng. 36(3), 820-824 (Mar 01, 1997);.*
- [42] Y. Lauzière, D. Gingras e F. Ferrie, "A model-based road sign identification system," em *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001. (Volume:1 Pages:1163-1170).*
- [43] B. Lauziere e G. Ferrie, "Bit Encoded Labelling and Quasi-Invariant Spatial Pattern Matching," em *Proceedings of 4th Asian Conference on Computer Vision, 2000 (Pages: 759-764).*
- [44] K. Sugahara, "Complex-log mapping approach to rotation and enlargement or reduction of digital images and its performance evaluation," em *IEEE 22nd Intern. Conf. Industrial Electronics, Control and Instrumentation, 1996 (Volume:3 Pages: 1655-1660).*
- [45] A. Ruta, Y. Li e X. Liu, "Real-time traffic sign recognition from video by class-specific discriminative features," em *Journal Pattern Recognition, Volume 43 Issue 1, January, 2010,*

Pages 416-430, 2010.

- [46] I. N. I.-E. R. (INIR), “Disposições Normativas Sinalização - pp.22-27,” [Online]. Available: <http://www.inir.pt/portal/LinkClick.aspx?fileticket=bHpImBCWewg%3D&tabid=116&mid=487&language=pt-PT>. [Acedido em 25 Junho 2013].
- [47] E. d. Portugal, “Normas de Sinalização - Estradas de Portugal,” [Online]. Available: <http://www.estradasdeportugal.pt/index.php/pt/areas-de-atuacao/seguranca-rodoviaria>. [Acedido em 12 Abril 2013].
- [48] J. Lewis, “Fast Normalized Cross-Correlation,” *Vision Interface, Canadian Image Processing and Pattern Recognition Society, Quebec, Canada, May 15-19, 1995 (Pages: 120-123)*.

7. Anexos

Anexo A - Sinalização presente na amostra utilizada para testar o método implementado

Sinal	Frames		Cor	Forma
	De	Até		
C13_120	31	89	1	100
D8	31	89	2	100
C14a	445	477	1	100
C13_80	445	477	1	100
C14a	477	500	1	100
C13_60	477	500	1	100
B9d	521	540	1	47
D1c	540	575	2	100
B2	540	575	1	100
C13_80	575	593	1	100
A1a	593	612	1	47
C13_60	593	612	1	100
C14a	612	637	1	100
C13_40	612	637	1	100
B9d	637	668	1	47
D1c	637	668	2	100
C11b	668	696	1	100
B1	668	696	1	43
C13_120	739	763	1	100
D8	739	763	2	100
B9d	899	914	1	47
B1	937	947	1	43
C11b	937	947	1	100
D1c	1148	1171	2	100
B1	1169	1176	1	43
C11b	1169	1176	1	100
C14a	2087	2103	1	100
C13_80	2087	2103	1	100
D1c	2231	2268	2	100
C13_120	2330	2358	1	100
D8	2330	2358	2	100
C14a	3219	3271	1	100
C13_60	3219	3271	1	100
C13_40	3219	3271	1	100
B9d	3805	3816	1	47
D1c	3816	3822	2	100
C11b	3816	3847	1	100

B1	3816	3847	1	43
A1a	3847	3859	1	47
C13_60	3847	3859	1	100
B9d	3878	3899	1	47
D1c	3899	3921	2	100
C13_120	3980	4006	1	100
D8	3980	4006	2	100
A1a	4866	4893	1	47
C13_60	4868	4893	1	100
C14a	4893	4918	1	100
C13_40	4893	4918	1	100
B9d	4918	4935	1	47
D1c	4935	4949	2	100
C11b	4949	4962	1	100
B1	4949	4962	1	43
A1a	6760	6776	1	47
C13_40	6760	6776	1	100
A23	6832	6843	1	47
C13_60	7897	7903	1	100
A1a	7897	7903	1	47
C11b	7981	7998	1	100
B1	7981	7998	1	43
C14a	10235	10257	1	100
C14a	10235	10257	1	100
B2	10361	10370	1	100
C11b	10361	10370	1	100
B1	10361	10370	1	43

A coluna “Sinal” corresponde à identificação de cada sinal de trânsito, de acordo com a nomenclatura utilizada em [5]. Na coluna “Frames” estão identificados o primeiro e o último *frames* da sequência em que cada sinal aparece. Na coluna “Cor” o algarismo 1 corresponde à cor vermelha e o 2 à cor azul. Na coluna “Forma” os algarismos 100, 47 e 43 correspondem a formas circular, triangular (triângulo na posição regular) e triangular (triângulo na posição invertida), respectivamente.

Anexo B – Documentação do código implementado para reconhecimento automático de sinalização vertical de trânsito num vídeo.

O método reconhecimento automático de sinalização vertical de trânsito desenvolvido foi implementado em Matlab, versão R2012b (8.0.0.783) de 64-bit, e emprega as bibliotecas de processamento de imagem e de sistemas de visão computacional. Todo este desenvolvimento foi efectuado no sistema operativo Windows 7 versão Home Premium de 64-bit.

O desenvolvimento de uma aplicação não foi o objectivo primário, tendo no entanto sido desenvolvidas funções que, uma vez integradas num *interface* gráfico poderão ser articuladas de maneira a ser criada uma aplicação independente. Na Figura 41, encontra-se um esquema com os dados de entrada e de saída de cada uma das funções implementadas. De seguida será apresentada uma descrição dessas funções:

- **ExtraiFrames** – Extrai os *frames* constituintes de um vídeo.

- **Recon Forma** – Generaliza a detecção e classificação para todos os *frames* presentes na directoria seleccionada.
 - MascaraCorHSVeForma_Vermelho – Função auxiliar da função “ReconForma” que efectua o processo de detecção e classificação (descritos nos capítulos 4.1 e 4.2 do presente documento) para a sinalização de cor vermelha.
 - MascaraCorHSVeForma_Azul – Função auxiliar da função “ReconForma” que efectua o processo de detecção e classificação (descritos nos capítulos 4.1 e 4.2 do presente documento) para a sinalização de cor azul.

- **Corresponde Sinais** – Generaliza o reconhecimento para todos os *frames* onde foi detectada e classificada a sinalização
 - Corr_PictogramaBW – Função auxiliar da função “Corresponde_Sinais” que faz a correspondência entre duas imagens binárias correspondentes aos pictogramas das RdI e dos *templates* (conforme descrito no capítulo 4.3.1).

- **MontaVideo** – Monta as imagens presentes numa directoria num vídeo.

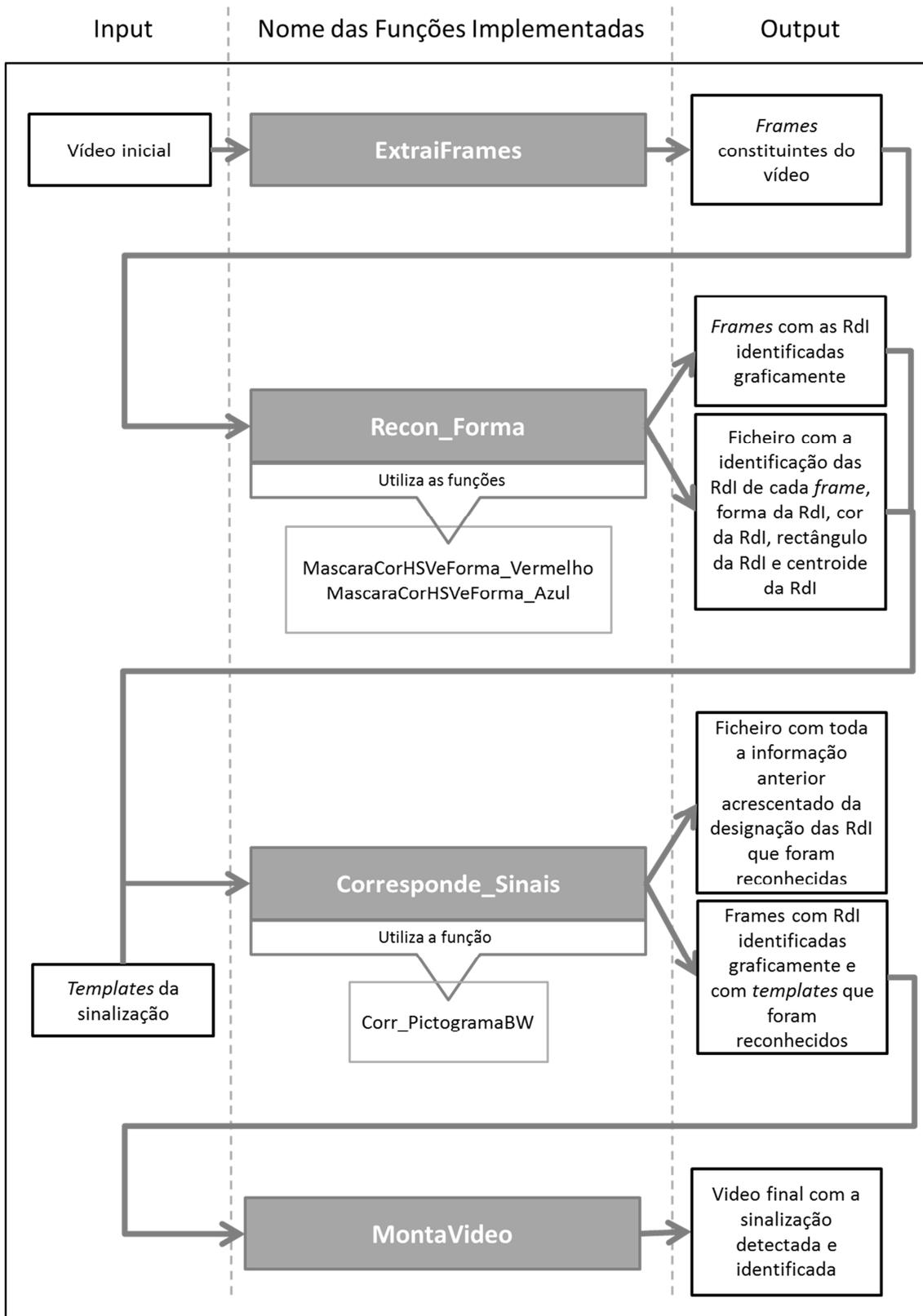


Figura 41 – Esquema ilustrativo do *input* e *output* das funções implementadas de modo a se proceder ao reconhecimento automático de sinalização vertical de trânsito.