



UNIVERSIDADE DA BEIRA INTERIOR  
Engenharia

# Classificação e reconhecimento de frutos por processamento digital de imagem para previsão da produção

Tiago Filipe da Mota Pereira

Dissertação para obtenção do Grau de Mestre em  
**Engenharia Eletrotécnica e de Computadores**  
(2º ciclo de estudos)

Orientador: Prof. Doutor Pedro Dinis Gaspar  
Departamento de Engenharia Eletromecânica  
Universidade da Beira Interior  
Covilhã, Portugal

Covilhã e UBI, outubro de 2017



## **Agradecimentos**

O espaço limitado desta seção, seguramente, não me permite agradecer devidamente, a todas as pessoas, que ao longo do meu Mestrado em Engenharia Eletrotécnica e de Computadores me ajudaram, direta ou indiretamente, a cumprir os meus objetivos e a realizar mais esta etapa da minha formação académica. Desta forma, deixo um profundo e total agradecimento aqueles sem os quais não seria possível a realização desta dissertação.

Ao Professor Doutor Pedro Dinis Gaspar, pela paciência, pela sua orientação, total apoio, disponibilidade, pelo saber que transmitiu, pelas opiniões e críticas, total colaboração no solucionar de dúvidas e problemas que foram surgindo ao longo da realização deste trabalho.

Por último, mas não menos importante, agradeço a minha família, em especial aos meus pais, pelo apoio incondicional, a presença constante e por quererem tanto quanto eu as metas que atingi ao longo da minha formação académica.

Obrigado!



## Resumo

O conceito de Agricultura de Precisão está normalmente associado à utilização de equipamento de alta tecnologia (seja hardware, no sentido genérico do termo, ou software) para avaliar, ou monitorizar, as condições numa determinada parcela de terreno, aplicando depois os diversos fatores de produção (sementes, fertilizantes, fitofármacos, reguladores de crescimento, água, etc.), em conformidade.

O tema proposto na presente dissertação tem como objetivo criar um algoritmo de análise e processamento de imagem que caracterize uma árvore, em particular pessegueiro, distinga os frutos e calcule as suas dimensões, e/ou volume, e/ou peso. A distinção dos frutos em árvore, ou seja, em ambiente natural, é complexa e requer algoritmos de segmentação igualmente complexos. No presente estudo, a segmentação de imagem é aplicada de forma a extrair as características de cor e forma usadas para identificação dos frutos. Para melhorar este reconhecimento, um método de classificação com recurso a uma máquina de suporte vetorial é usado, atingindo uma taxa de reconhecimento a rondar 70%. A estimativa da previsão de produção é feita com recurso aos valores do volume calculados para os frutos reconhecidos, obtendo um valor estimado para a previsão de produção de 29,3 toneladas por hectare.

## **Palavras-chave**

Agricultura de Precisão

Segmentação

Espaços de cor

Deteção de arestas

Máquina de suporte vetorial

Previsão de Produção

Pessegueiro

## **Abstract**

The concept of Precision Agriculture is usually associated with the usage of high-end technology equipments (being hardware or software), to first evaluate or control the conditions of a determined portion of land, applying afterwards various factors of production accordingly, like seeds, fertilizers, phytopharmaceuticals, growing regulators, water, etc.

The proposed theme of this dissertation aims to create an algorithm capable of analyse and process images to characterize trees, particularly peach trees, distinguish fruits and calculate it dimensions, like volume and weight. The recognition of peaches on their natural conditions, on trees, is complex and requires segmentation algorithms, also complex. The proposed algorithm applies image segmentation for extraction of characteristics such as colour and shape. These characteristics are then used to train a classification method through a support vector machine to improve the recognition rate of fruits, accomplishing results around 70%. The production prediction is obtained with the volume values already calculated for the recognised peaches, providing a prediction of 29.3 tons per hectare.

## **Keywords**

Precision agriculture

Segmentation

Colour spaces

Shape recognition

Support vector machine

Production prediction

Peach trees



---

# Índice

Agradecimentos .....	i
Resumo .....	iii
Abstract .....	v
Índice .....	vii
Lista de Figuras .....	ix
Lista de Tabelas .....	xiii
Nomenclatura .....	xv
<b>1. Introdução .....</b>	<b>1</b>
1.1. Enquadramento.....	1
1.2. O problema em estudo e a sua relevância .....	4
1.3. Objetivos e contribuição da dissertação.....	6
1.4. Visão geral e organização da dissertação.....	7
<b>2. Estado da Arte .....</b>	<b>9</b>
2.1. Introdução.....	9
2.2. Cor .....	10
2.2.1. Representação espectral da luz .....	10
2.2.2. Visualização de cores .....	11
2.2.3. Perspetiva de “visualização” de uma câmara digital.....	12
2.3. Espaços de Cor .....	13
2.3.1. Espaço de cores RGB .....	13
2.3.2. Outros espaços de cores .....	15
2.4. Textura .....	17
2.4.1. Métodos estruturais.....	18
2.4.2. Métodos espaciais.....	19
2.4.3. Métodos baseados em modelos .....	19
2.4.4. Métodos baseados em transformações.....	20

2.5.	Reconhecimento de contornos.....	20
2.5.1.	Técnicas de reconhecimento de contornos .....	21
2.6.	Machine Learning .....	23
2.6.1.	Algoritmo K-means Cluster .....	23
2.6.2.	K-nearest neighbors .....	24
2.6.3.	Máquinas de suporte vetorial (Support Vector Machine - SVM) .....	25
2.7.	Estudos científicos na temática.....	25
2.7.1.	Outros estudos científicos na temática da agricultura de precisão	29
<b>3.</b>	<b>Algoritmo de Reconhecimento e Previsão de Carga de Frutos.....</b>	<b>33</b>
3.1.	Introdução .....	33
3.2.	Algoritmo de reconhecimento de fruta.....	33
3.2.1.	Limite (threshold) de Cor .....	33
3.2.2.	Segmentação e análise das imagens .....	35
3.2.3.	Deteção de arestas.....	36
3.2.4.	Probabilidade de um objeto como redondo .....	37
3.2.5.	Nota Conclusiva.....	39
3.3.	Aprendizagem máquina (Machine Learning) - Treino da SVM .....	40
3.3.1.	Classificação de exemplos de treino.....	41
3.3.2.	Modelo SVM e a sua aplicação .....	42
3.3.3.	Nota Conclusiva.....	44
<b>4.</b>	<b>Análise e Discussão de Resultados.....</b>	<b>45</b>
4.1.	Introdução .....	45
4.2.	Testes e Resultados .....	45
4.2.1.	Apresentação dos resultados.....	48
4.3.	Previsão de produção .....	61
4.3.1.	Previsão de Produção - Algoritmo.....	61
4.3.2.	Previsão de Produção - Teórica.....	64
4.4.	Discussão dos resultados .....	65
4.5.	Nota Conclusiva .....	66
<b>5.</b>	<b>Conclusões .....</b>	<b>69</b>
5.1.	Conclusões .....	69
5.2.	Sugestões de trabalhos futuros .....	70
	<b>Referências Bibliográficas .....</b>	<b>73</b>

## Lista de Figuras

Figura 1 - Esquema geral de um sistema de Agricultura de Precisão (Coelho & Silva, 2009). .....	3
Figura 2 - Espectro de cores visíveis como função de comprimento de onda em nanómetros (Corke, 2011) .....	10
Figura 3 - Eletromicrografia colorida de bastonetes (branco) e células em cone (amarelo) (Corke, 2011). .....	12
Figura 4 - Filtro Bayer. Os blocos cinzentos representam o vetor de photosites de silicone sensíveis à luz sobre o qual está um vetor de filtros vermelhos, verdes ou azuis (Corke, 2011). .....	13
Figura 5 - Locus espectral no plano de cromaticidade r-g (Corke, 2011). .....	14
Figura 6 - Exemplo de reconhecimento baseado na SVM (Wang et al., 2012). .....	27
Figura 7 - Diferentes tipos de frutos e respectiva amostra gráfica (Zhang & Wu, 2012). .....	31
Figura 8 - Limite de cor no plano HSV.....	34
Figura 9 - Exemplo de imagem após segmentação. ....	35
Figura 10 - Resultados obtidos como operador de Sobel. ....	37
Figura 11 - Exemplo de contornos e respetiva métrica.....	38
Figura 12 - Fluxograma da função SVM DATA .....	40
Figura 13 - Diagrama de blocos do algoritmo final. ....	44
Figura 14 - Variação dos resultados de segmentação para as imagens da Figura 14: a) Segmentação da Figura 14 a), b) Segmentação da Figura 14 b), c) Segmentação da Figura 14 c).....	47

Figura 15 - Variação de luminosidade nos diferentes dias: a) dia #1, b) dia #2, c) correção da luminosidade do dia dois. ....	47
Figura 16 - Exemplo das diferentes dimensões dos pêsegos: a) Cultivar 1, b) Cultivar 2. ....	48
Figura 17 - Gráficos de dispersão para os melhores ensaios para ambos os casos: a) Gráfico de dispersão do conjunto de treino equivalente a 80% da amostra, b) Gráfico de dispersão do conjunto de teste equivalente a 20% da amostra, c) Gráfico de dispersão do conjunto de treino equivalente a 90% da amostra, d) Gráfico de dispersão do conjunto de teste equivalente a 10% da amostra.....	51
Figura 18 - Representação através de gráfico de barras dos resultados obtidos na Tabela 7: a) Gráfico dos resultados obtidos para os 10 ensaios com a configuração de conjuntos 80%-20%, b) Gráfico dos resultados obtidos para os 10 ensaios com a configuração de conjuntos 90%-10%. ....	52
Figura 19 - Gráficos de dispersão para os melhores ensaios para ambos os casos: a) Gráfico de dispersão do conjunto de treino equivalente a 80% da amostra, b) Gráfico de dispersão do conjunto de teste equivalente a 20% da amostra, c) Gráfico de dispersão do conjunto de treino equivalente a 90% da amostra, d) Gráfico de dispersão do conjunto de teste equivalente a 10% da amostra.....	55
Figura 20 - Representação através de gráfico de barras dos resultados obtidos na Tabela 8: a) Gráfico dos resultados obtidos para os 10 ensaios com a configuração de conjuntos 80%-20%, b) Gráfico dos resultados obtidos para os 10 ensaios com a configuração de conjuntos 90%-10%. ....	56
Figura 21 - Gráficos de dispersão para os melhores ensaios para ambos os casos: a) Gráfico de dispersão do conjunto de treino equivalente a 80% da amostra, b) Gráfico de dispersão do conjunto de teste equivalente a 20% da amostra, c) Gráfico de dispersão do conjunto de treino equivalente a 90% da amostra, d) Gráfico de dispersão do conjunto de teste equivalente a 10% da amostra.....	59
Figura 22 - Representação através de gráfico de barras dos resultados obtidos na Tabela 9: a) Gráfico dos resultados obtidos para os 10 ensaios com a configuração de conjuntos 80%-20%, b) Gráfico dos resultados obtidos para os 10 ensaios com a configuração de conjuntos 90%-10%. ....	60
Figura 23 - Contagem manual de pêsegos. ....	64

Figura 24 - Graficos de dispersão do ensaio de validação das características de treino da SVM: a) conjunto de treino, b) conjunto de teste..... 67

Figura 25 -Exemplo de figura usada para extração de informação do ensaio de validação. .... 68



## Lista de Tabelas

Tabela 1 - Valor do comprimento de onda das cores primárias da CIE 1976 (Commission internationale de L'Éclairage) (Corke, 2011). .....	13
Tabela 2 - Taxa média de reconhecimento da máquina de suporte vetorial para diferentes funções Kernel (Wang et al., 2012). .....	28
Tabela 3 - Taxa média de reconhecimento [%] da máquina de suporte vetorial para diferentes funções Kernel (Zhang & Wu, 2012). .....	30
Tabela 4 - Resultados referentes à Figura 9. ....	39
Tabela 5 - Classificação dos exemplos de treino. ....	41
Tabela 6 - Normalização dos dados. ....	43
Tabela 7 - Resultados obtidos para os testes do primeiro caso: conjunto de teste 20% e 10% respectivamente lado esquerdo e lado Direito. ....	49
Tabela 8 - Resultados obtidos para os testes do segundo caso: conjunto de teste 20% e 10% respectivamente lado esquerdo e lado Direito. ....	53
Tabela 9 - Resultados obtidos para os testes do terceiro caso: conjunto de teste 20% e 10% respectivamente lado esquerdo e lado Direito. ....	57





## Nomenclatura

### Geral:

$A$	Área, [m <sup>2</sup> ];
$c$	Velocidade da luz, [m·s];
$h$	Constante de Planck, [J·s];
$k$	Constante de Boltzman, [J/K];
$P$	Peso, [g];
$R$	Raio, [cm];
$T$	Temperatura absoluta da fonte, [K];
$V$	Volume, [cm <sup>3</sup> ];

### Simbologia grega:

$\lambda$	Comprimento de onda, [nm]
$\rho$	Densidade, [g·cm <sup>-3</sup> ];

### Acrónimos:

CIE	<i>Comission Internationale de L'Eclairage;</i>
GPS	<i>Global Positioning System;</i>
KNN	<i>K-Narest Neighbors;</i>
RISAR	<i>Rotation Invariant SAR;</i>
SAR	<i>Simultaneous Autoregressive Model;</i>
SIG	<i>Sistemas de Informação Global;</i>
SVM	<i>Support Vector Machine;</i>
UBI	Universidade da Beira Interior;
UAV	Unmanned Aerial Vehicle;



## 1. Introdução

Este capítulo introdutório inicia-se com o enquadramento do tema, onde é descrito o conceito de agricultura de precisão e os seus objetivos perante os agricultores. De seguida, é exposto o problema em estudo e a sua relevância, abordando o tema da previsão de produção e os problemas associados a uma previsão de produção assertiva. Posteriormente, são referidos os objetivos e a contribuição da dissertação, onde o principal objetivo consiste no desenvolvimento de um sistema de análise e processamento de imagem que disponibilize aos fruticultores uma previsão de produção assertiva. Por fim, é dada uma visão geral e de organização da dissertação por capítulos.

### 1.1. Enquadramento

O conceito de Agricultura de Precisão está normalmente associado à utilização de equipamento de alta tecnologia (seja hardware, no sentido genérico do termo, ou software) para avaliar, ou monitorizar, as condições numa determinada parcela de terreno, aplicando depois os diversos fatores de produção (sementes, fertilizantes, fitofármacos, reguladores de crescimento, água, etc.) em conformidade. Tanto a monitorização como a aplicação diferenciada, ou à medida, exigem a utilização de tecnologias recentes, como os sistemas de posicionamento a partir de satélites (GPS - *Global Positioning System*), os Sistemas de Informação Geográfica (SIG) ou os sensores eletrónicos, associados quer a reguladores automáticos de débito nas máquinas de distribuição quer a medidores de fluxo nas máquinas de colheita (Coelho & Silva, 2009).

A Agricultura de Precisão aparece, geralmente, associada a dois objetivos genéricos: o aumento do rendimento dos agricultores; e, a redução do impacto ambiental resultante da atividade agrícola. O primeiro destes objetivos pode, por sua vez, ser alcançado por duas vias distintas, mas complementares: a redução dos custos de produção; e, o aumento da produtividade (e, por vezes, também da qualidade) das culturas. O

cumprimento do segundo destes objetivos está relacionado com o rigor do controlo da aplicação dos fatores de produção (sobretudo, produtos químicos, atendendo às externalidades ambientais negativas que lhes estão normalmente associadas), que deverá ser feita, tanto quanto possível, na justa medida das necessidades das plantas (Coelho & Silva, 2009).

De facto, sabendo, por exemplo, que as necessidades de azoto em duas áreas distintas de uma mesma parcela de terreno não são iguais, por hipótese, em função dos resultados da análise de terras para as duas situações ter-se-á, do ponto de vista estritamente técnico e teórico, vantagem em fazer variar a quantidade de adubo azotado em conformidade. Deste modo, seria natural a tentação de aplicar mais adubo na área em que as necessidades ou o potencial produtivo são maiores, e a reduzir a quantidade na área em que o potencial produtivo ou as necessidades são menores, em vez de, como usualmente sucede, aplicar um valor médio e igual em toda a parcela. Com esta forma de atuar, não só se aumentaria a produção, aplicando mais adubo onde este é necessário, como também se reduziria os custos e o impacto ambiental da atividade (Coelho & Silva, 2009).

Na prática, todas estas proposições seriam inquestionáveis, não fora o caso das tecnologias associadas à Agricultura de Precisão serem, quase sempre, complexas e caras. É exatamente por esta razão que o conceito não se encontra hoje em dia mais divulgado, nomeadamente no nosso país. Em primeiro lugar, só fará sentido recorrer à Agricultura de Precisão se os benefícios económicos daí decorrentes forem superiores ao investimento necessário à sua adoção. Ora, infelizmente, são poucas as explorações, sobretudo em Portugal, com dimensão suficiente para, só por si, justificar ou viabilizar tais investimentos. Note-se que os investimentos referidos não passam apenas pela aquisição de determinados equipamentos (GPS, sensores, etc.), mas também pelo levantamento da situação de base e pela construção de um sistema de informação geograficamente referenciada (Coelho & Silva, 2009). Em segundo lugar, existe ainda um longo caminho a percorrer pelas tecnologias de informação associadas, especialmente no que se refere à sua facilidade de uso e de integração no negócio. Isto é, ainda é necessário um esforço razoável para efetuar a recolha e processamento da informação necessária à prática da agricultura de precisão, esforço este com custos muitas vezes inaceitáveis para o agricultor, nomeadamente quanto ao dispêndio do seu tempo/atenção. Em terceiro lugar, na hipótese de que os investimentos sejam

rentáveis, é necessário que existam pessoas (agricultores e/ou técnicos) com conhecimentos suficientes para ajustar, desenvolver e utilizar estas tecnologias (Coelho & Silva, 2009). Encontram-se inúmeras definições para um conceito tão genérico como o de Agricultura de Precisão, mas pela sua abrangência e simplicidade, é proposto o seguinte:

*“A Agricultura de Precisão envolve a aplicação diferenciada e à medida dos fatores de produção, tendo em conta a variação espacial e temporal do potencial produtivo do meio e das necessidades específicas das culturas, de forma a aumentar a sua eficiência de utilização e, assim, melhorar o rendimento económico e reduzir o impacte ambiental da atividade agrícola” (Coelho & Silva, 2009).*

Atualmente, os exemplos mais comuns de Agricultura de Precisão estão relacionados com a aplicação diferenciada no espaço, de sementes, fertilizantes, fitofármacos e água de rega, o que se justifica, sobretudo, pelo elevado peso económico que estes fatores normalmente representam nos custos totais das culturas, pela facilidade de relacionar o seu nível de utilização com a produtividade alcançada pelas culturas e pelo, maior ou menor, impacto ambiental que podem ter. Na Figura 1 é apresentado um esquema geral de um sistema de Agricultura de Precisão.

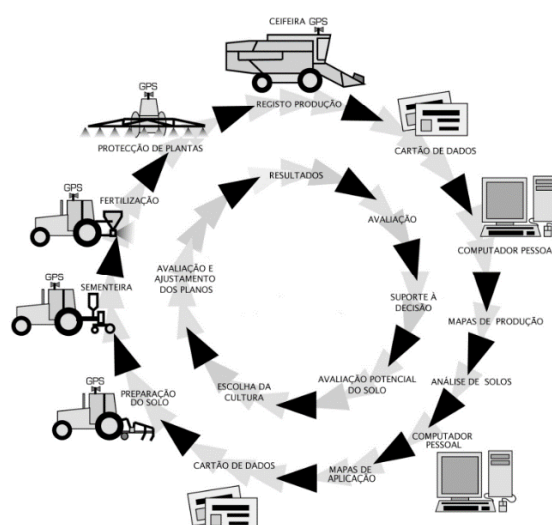


Figura 1 - Esquema geral de um sistema de Agricultura de Precisão (Coelho & Silva, 2009).

As aplicações diferenciadas de fertilizantes, por exemplo, podem não só contribuir para aumentar consideravelmente o rendimento económico das culturas, como ajudam a reduzir o arrastamento de nutrientes e a consequente contaminação das águas residuais e subterrâneas. Hoje em dia, é relativamente fácil (e barato) analisar o teor dos macronutrientes no solo, o que permite mapear a fertilidade de pequenas, médias ou grandes parcelas. Além disso, já existe tecnologia disponível que permite efetuar aplicações diferenciadas no espaço. No entanto, a inércia na adoção da Agricultura de Precisão persiste, podendo ser essencialmente explicada por três ordens de razões: (1) o baixo *know-how* específico, nestas matérias, de agricultores, técnicos e empresas ligadas ao sector; (2) o relativamente elevado custo inicial da mudança, associado aos equipamentos (hardware e software) necessários a este tipo de agricultura; (3) a relativamente modesta escala de operação da generalidade das explorações agrícolas europeias e, sobretudo, portuguesas (Coelho & Silva, 2009).

No caso dos outros fatores de produção, a importância relativa dos resultados económicos e ambientais pode ser distinta, mas existe igualmente tecnologia disponível para efetuar aplicações de acordo com necessidades espacialmente distintas e definidas. Seja em que caso for, se a variabilidade espacial existir, se for possível medi-la e determinar como afeta a produtividade e a qualidade das culturas, e se existir tecnologia para aplicar os fatores de produção de forma diferenciada, serão apenas as variáveis de natureza económica a ditar se é vantajoso adotar sistemas de Agricultura de Precisão. Neste particular, pesa muito a situação específica de cada exploração, nomeadamente no que concerne à intensidade da variação das condições do meio no seu interior e à sua dimensão física e/ou económica (Coelho & Silva, 2009).

### **1.2. O problema em estudo e a sua relevância**

Os fruticultores realizam previsões da produção para gerir todos os aspetos relacionados com a colheita e a comercialização, entre os quais os contratos de fornecimento. Usualmente é realizada uma seleção aleatória de um n.º de árvores para contagem visual

dos frutos, podendo ser apoiada por medições de calibre<sup>1</sup>. A precisão da estimativa da produção aumenta com o tamanho da amostra, mas requer um maior n.º de pessoas e tempo, o que se traduz em maiores custos. Após a celebração do contrato de fornecimento, no caso de o produtor não conseguir fornecer a quantidade contratada, poderá ter que ressarcir o distribuidor (custo adicional), contribuindo para uma deterioração das relações Produtor/Comprador. Uma previsão da produção mais assertiva permite colmatar estas situações, desenvolver sistemas de apoio, como por exemplo de seguros de colheita, e até mesmo contribuir ao desenvolvimento de sistemas integrados de gestão da produção da exploração.

Neste sentido pretende-se desenvolver um sistema capaz de ajudar a colmatar este problema. O objetivo passa pela criação de um sistema, equipado com um algoritmo de análise e processamento de imagem que caracteriza a árvore, distingue os frutos e determina as suas dimensões, e/ou volume, e/ou peso. A previsão mais assertiva da produção com os algoritmos de análise e processamento de imagem de caracterização das árvores e de segmentação e classificação dos frutos, possibilita aos produtores uma melhor gestão da carga e previsão da produção, podendo negociar com maior confiança as quantidades e a qualidade (calibre) dos frutos na celebração de contratos de fornecimento. Esta vantagem resulta numa maior rentabilidade da produção, promoção das relações produtor/comprador e contribui para a redução do desperdício alimentar pela rejeição de produto, já que este poderá ser conduzido para outro fim atempadamente.

Atualmente, para determinar o potencial produtivo dos pomares, os modelos de base ecofisiológica, construídos a partir de dados meteorológicos, são demasiado complexos e pouco operacionais para determinar a necessidade e intensidade da monda de frutos.

---

<sup>1</sup> O calibre é determinado pelo diâmetro máximo da secção equatorial, pelo peso ou pelo número de frutos. No caso específico de pêssegos e às nectarinas que estão sujeitos a uma Norma Específica de Comercialização, o calibre mínimo é de: (1) 56 mm ou 85 g na categoria «Extra»; (2) 51mm ou 65 g nas categorias I e II (em caso de calibragem). Detalhes acerca das normas de comercialização aplicável a vários produtos hortofrutícolas podem ser encontrados em Jornal Oficial da União Europeia, 15/06/2011, que incorpora o Regulamento de Execução (UE) N.º 543/2011 da Comissão, de 7 de Junho de 2011, que estabelece regras de execução do Regulamento (CE) n.º 1234/2007 do Conselho nos sectores das frutas e produtos hortícolas e das frutas e produtos hortícolas transformados.

Por sua vez, os modelos empíricos são fáceis de utilizar e de operacionalizar, mas carecem de automatismos para determinar a capacidade das árvores e para contar os frutos. Na literatura, encontram-se vários exemplos de algoritmos de análise e processamento de imagem para a contagem de frutos nas árvores e escassos exemplos de aplicação desses algoritmos em plataformas móveis terrestres. Usualmente, configuram-se como veículos autônomos ou semiautônomos, que transportam uma câmara e percorrem as fileiras e recolhem imagens que são posteriormente processadas para efetuar a contagem de frutos. Os poucos exemplos existentes ainda não são comercialmente viáveis. Também os veículos aéreos não tripulados (*unmanned aerial vehicle* - UAV) são utilizados para a previsão de colheitas (sobretudo de cereais), existindo empresas que recolhem imagens a altitudes elevadas com câmaras multiespectrais. Contudo, os serviços oferecidos e os sensores envolvidos apresentam custos elevados e um erro inerente relevante, uma vez que a análise se baseia apenas na quantificação de radiação refletida. Uma limitação na contagem de frutos através de imagem reside na oclusão destes por folhas e ramos.

### **1.3. Objetivos e contribuição da dissertação**

O tema proposto tem como objetivo criar um algoritmo de análise e processamento de imagem que caracterize a árvore, distinga os frutos e calcule as suas dimensões, e/ou volume, e/ou peso. O objetivo deste algoritmo passa por ajudar os fruticultores a terem um método de previsão de produção mais viável que os métodos atuais. As vantagens de ter uma previsão mais precisa já foram indicadas e são fundamentais em vários aspetos sendo que favorecem o fruticultor.

À posteriori, perspetiva-se a implementação deste algoritmo num sistema robótico aéreo autônomo (drone) equipado com câmaras para recolher imagens digitais de alta resolução, possivelmente complementadas por imagens espectrais, sendo que este iria transitar nas entrelinhas do pomar de modo a captar dados de ambos os lados das árvores. Após o desenvolvimento desta tecnologia em pomares de Prunóideas, esta poderá ser adaptada a outro tipo de pomares (Pomóideas) com repercussões a nível regional e nacional.



Outro objetivo deste trabalho visa a divulgação dos métodos de agricultura de precisão e as suas vantagens, fazendo com que mais agricultores se tornem adeptos e utilizadores de tais tecnologias.

#### 1.4. Visão geral e organização da dissertação

Como já foi exposto anteriormente, o primeiro capítulo da presente dissertação é introdutório, fazendo o enquadramento do trabalho a ser desenvolvido e realçando a inclusão deste trabalho no tema da agricultura de precisão. Neste capítulo, é ainda explicado em detalhe o problema definindo os objetivos e apresentando as contribuições esperadas.

O segundo capítulo, intitulado “Estado da Arte”, contém a apresentação teórica dos conceitos e técnicas relacionadas com segmentação e processamento de imagem. São expostos também alguns algoritmos relacionados com o tema.

De seguida, no terceiro capítulo, aborda-se a secção do processo de desenvolvimento do algoritmo. São apresentados separadamente os vários processos de segmentação com base em algoritmos de *machine learning*, tais como *K-means clustering*, *K-Nearest Neighbors* (KNN) e *Support Vector Machines* (SVM's). Neste mesmo capítulo é também realizada uma descrição do estado da arte.

No quarto capítulo é feita a análise e discussão dos resultados obtidos. São apresentados os resultados obtidos para os coeficientes de correlação gerados pelo algoritmo, e ainda são apresentados os resultados de previsão da produção, práticos e teóricos, de modo a comprovar a precisão dos mesmos. É ainda apresentado um teste de validação do algoritmo.

No último capítulo, são apresentadas as conclusões gerais do trabalho tendo em consideração os resultados obtidos pelo algoritmo. É também apresentado um conjunto de sugestões de trabalhos futuros, no seguimento dos resultados obtidos nesta dissertação.



## 2. Estado da Arte

Este capítulo inicia-se com a apresentação de forma teórica das principais características para a segmentação de imagem. Componentes como cor, diferentes espaços de cor, textura e detecção de contornos são elementos fundamentais para extração de informação das imagens digitais. No seguimento dos conceitos referidos são ainda apresentados algoritmos de *machine learning* tais como *K-means clustering*, *K-Nearest Neighbors (KNN)* e *Support Vector Machines (SVM's)*. Neste capítulo é ainda realizada a descrição dos artigos científicos mais relevantes para o tema da dissertação.

Por fim, são apresentados alguns algoritmos já existentes com propósitos semelhantes ao algoritmo que se pretende desenvolver no âmbito deste trabalho.

### 2.1. Introdução

Uma imagem digital é uma representação numérica, normalmente binária, de uma imagem bidimensional. Existem dois tipos de imagens digitais. Dependendo se a resolução da imagem é fixa, pode ser tipo vetor ou tipo *raster*. Sem qualificações, o termo imagem digital, normalmente refere-se a imagens *raster*, também chamadas de imagens mapa de bits (*bitmap*). A classificação de imagens digitais pode ser feita de acordo com os valores do bitmap. Numa imagem em binário só existem apenas dois valores possíveis para cada pixel, 0 ou 1. Tipicamente as duas cores usadas nas imagens binárias são o preto e o branco, sendo que quaisquer duas cores podem ser usadas para representar imagens binárias. Numa imagem de escala de cinzentos, o valor de cada pixel transporta informação sobre a intensidade da cor. Imagens de escalas de cinzentos também conhecidas como imagens monocromáticas, são compostas exclusivamente por tons de cinza, variando de preto (pouca intensidade) para branco (elevada intensidade). Imagens de cor

contêm informação de cor para cada pixel. Para resultados visualmente aceitáveis são necessários três valores provenientes de três canais de cores, normalmente vermelho, verde e azul (formato RGB), para cada pixel. Uma imagem de cor verdadeira, é a imagem que aparece na sua forma real perante o olho humano. Uma imagem de cor falsa, é uma imagem com cores diferentes da realidade. Uma imagem de *range* representa a profundidade no valor de cada pixel (Chakraborty et al., 2017).

## 2.2. Cor

### 2.2.1. Representação espectral da luz

Em torno de 1670, O Sr. Isaac Newton descobriu que a luz branca era uma mistura de diferentes cores. Agora, sabemos que cada uma destas cores corresponde a uma frequência ou comprimento de onda de radiação eletromagnética. O ser humano consegue distinguir comprimentos de onda entre 400 e 700 nm como cores (Corke, 2011). Na Figura 2 encontra-se representado o espectro de cores visíveis em função de comprimento de onda.

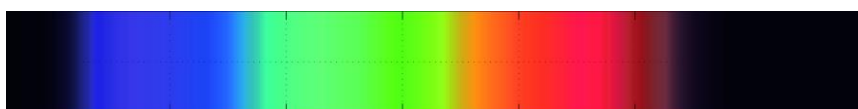


Figura 2 - Espectro de cores visíveis como função de comprimento de onda em nanômetros (Corke, 2011)

Espectro de luz: *“Durante os anos de praga de 1665-1666 Isaac Newton desenvolveu a sua teoria de luz e cor. Ele demonstrou que um prisma podia decompor luz branca num espectro de cores, e uma lente e um segundo prisma podia recompor o espectro multicolorido em luz branca. Mais importante, ele mostrou que a cor da luz não mudou quando foi refletida por diferentes objetos, do qual concluiu que a cor é uma propriedade intrínseca da luz e não o objeto”* (Corke, 2011).

Em geral, a luz observada é uma mistura de vários comprimentos de onda e pode ser representada como uma função  $E(\lambda)$  dada pela Equação 1, que descreve a intensidade em função do comprimento de onda ( $\lambda$ ) (Corke, 2011).

$$E(\lambda) = \frac{2\pi hc^2}{\lambda^5 \left( e^{\frac{hc}{k\lambda T}} - 1 \right)} \text{ Wm}^{-2}\text{m}^{-1} \quad (1)$$

Em que:

- $T$  - Temperatura absoluta da fonte (K);
- $h$  - Constante de Planck ( $6,626 \cdot 10^{-34} \text{ J} \cdot \text{s}$ );
- $k$  - Constante de Boltzmann ( $1,381 \cdot 10^{-23} \text{ J/K}$ );
- $c$  - Velocidade da luz ( $3 \cdot 10^8 \text{ m} \cdot \text{s}$ );

A fonte mais comum de luz é a incandescência, que é a emissão de luz por parte de um corpo quente tal como o sol ou um filamento de lâmpada (Corke, 2011).

### 2.2.2. Visualização de cores

O olho humano contém dois tipos de células sensíveis à luz. Células em cone que respondem a cores particulares e mostram o que se considera normal luz do dia, e células em haste (bastonetes), que são muito mais sensíveis que as células em cone, mas respondem apenas à intensidade e são mais usadas durante a noite (Corke, 2011). Na Figura 3 é apresentada uma imagem de uma eletromicrografia colorida de bastonetes e de células em cone.

Em condições normais da luz do dia, os fotorreceptores em forma de cone estão ativos e são sensíveis à cor. Os seres humanos são tricromáticos e tem três tipos de cones que correspondem a três partes diferentes do espectro. São referidos como longo (L), médio (M) e curto (S), de acordo com o pico de resposta do comprimento de onda, ou de forma mais comum como vermelho, verde e azul (Corke, 2011).

A retina do olho humano tem uma região central ou fóvea que tem apenas 0,6 mm de diâmetro e contém a maior parte das 6 milhões de células em cone: 65% são sensíveis ao vermelho; 33% são sensíveis ao verde; 2% são sensíveis ao azul. Os olhos humanos analisam subconscientemente o mundo, para construir uma imagem mental de tudo o que os rodeia. A fóvea é responsável por esse processo. Adicionalmente, ainda existem 120 milhões de bastonetes, distribuídos pela retina, que também são sensíveis ao movimento (Corke, 2011).

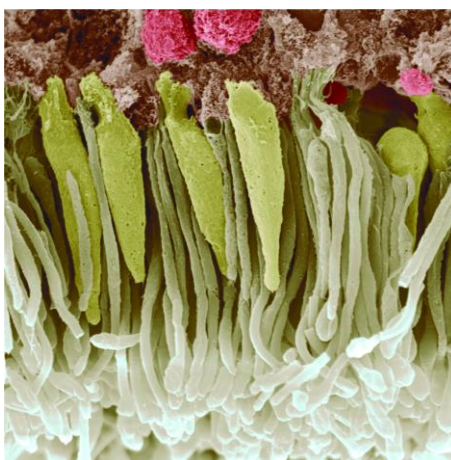


Figura 3 - Eletromicrografia colorida de bastonetes (branco) e células em cone (amarelo) (Corke, 2011).

### 2.2.3. Perspetiva de “visualização” de uma câmara digital

O sensor de uma câmara digital é análogo à retina, mas no lugar de bastonetes e células em cone possui um vetor regular de *photosites*<sup>2</sup> sensíveis à luz num chip de silicone. Cada *photosite* é da ordem de 1-10  $\mu\text{m}^2$  e transmite um sinal proporcional à intensidade de luz que atinge a sua área. Para uma câmara de cores, os *photosites*

---

<sup>2</sup> Um *photosite* é um pixel do sensor. Dependendo do design do sensor, um photosite pode conter o sistema de circuitos necessários para as múltiplas cores dos pixels. Todos os pixels dos sensores vão ter alguns dos mesmos componentes básicos. O elemento chave de um pixel é o foto-díodo. Um foto-díodo é um foto-detector semiconductor que usa as sensibilidades naturais de certos materiais, tais como Germanium e Silicon, para converter fotões em eletrões.

estão cobertos com filtros que transmitem cor vermelha, verde ou azul (Corke, 2011). Na Figura 4 encontra exemplificado um filtro Bayer, correspondente a um sensor de uma câmara digital. O valor normalizado do comprimento de onda para as cores vermelho, verde e azul encontra-se descrito na Tabela 1.

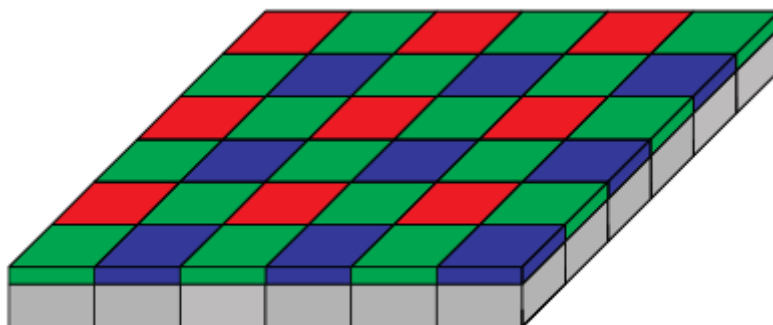


Figura 4 - Filtro Bayer. Os blocos cinzentos representam o vetor de photosites de silicone sensíveis à luz sobre o qual está um vetor de filtros vermelhos, verdes ou azuis (Corke, 2011).

Tabela 1 - Valor do comprimento de onda das cores primárias da CIE 1976 (Commission internationale de L'Éclairage) (Corke, 2011).

	vermelho	verde	azul
$\lambda(\text{nm})$	700.0	546.1	435.8

## 2.3. Espaços de Cor

### 2.3.1. Espaço de cores RGB

O espaço de cor é um modelo matemático para representar informação da cor em três ou quatro componentes de cor diferentes. Diferentes modelos de cor são usados para diferentes aplicações tais como computação gráfica, processamento de imagem, transmissões televisivas e visão computacional (Corke, 2011).

O espaço de cores RGB é definido pelas três cores primárias: vermelho, verde e azul. Os Valores de tristimulus descrevem cor, bem como brilho. Os valores relativos de tristimulus são obtidos normalizando os valores de tristimulos (Corke, 2011).

$$r = \frac{R}{R+G+B}, \quad g = \frac{G}{R+G+B}, \quad b = \frac{B}{R+G+B} \quad (2)$$

que resulta em coordenadas de cromaticidade  $r$ ,  $g$  e  $b$  que são invariáveis ao brilho em geral. Por definição:  $r + g + b = 1$ , então, uma coordenada é redundante e tipicamente apenas  $r$  e  $g$  são consideradas. Como o efeito da intensidade foi eliminado, a quantidade bidimensional  $(r, g)$  representa cor (Corke, 2011).

Na Figura 5 é apresentado o locus espectral no plano de cromaticidade  $r$ - $g$ .

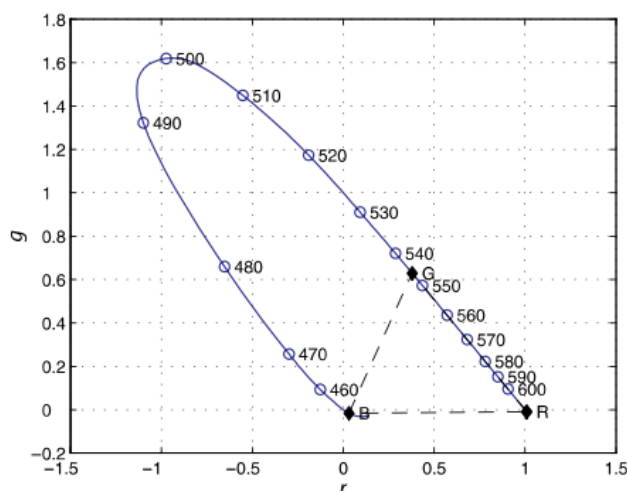


Figura 5 - Locus espectral no plano de cromaticidade  $r$ - $g$  (Corke, 2011).

O centro da lei da gravidade diz que uma mistura de duas cores é representada por uma linha entre essas duas cores no plano de cromaticidade. Uma mistura de  $N$  cores fica situada numa região delimitada por essas cores. Considerando a Figura 5, isto tem implicações significantes. Primeiro, como o estímulo de todas as cores são combinações de estímulos espectrais, o estímulo de todas as cores reais deve estar situado dentro ou no locus espectral. Segundo, qualquer cor que seja criada pela mistura das três cores primárias pode apenas estar situada no triângulo criado pelas três cores primárias - espaço de cores. É claro na figura 5, que as cores primárias CIE definem apenas um pequeno subconjunto de todas as cores possíveis - triângulo



tracejado. Muitas cores reais não podem ser criadas usando estas primárias, em particular as cores do arco-íris que se situam no locus espectral de 460-545 nm. De facto, não interessa onde estejam colocadas as cores primárias, pois não é possível produzir todas as cores. Em termos geométricos não existem três pontos dentro do espaço de cores que formem um triângulo que consiga abranger o espaço todo. Terceiro, observa-se que grande parte do espaço de cores requer uma quantidade negativa do vermelho primário que não pode ser representado (Corke, 2011).

### 2.3.2. Outros espaços de cores

Hernández-Hernández *et al.* (2016), construiu um software capaz de identificar o melhor modelo de cores dependendo do ambiente agrícola em que se pretende trabalhar. O software apresenta resultados para diferentes modelos de cor, e é baseado numa técnica de classificação de cores subjacentes e um novo método de treino. Os modelos de cor utilizados foram os mais comuns e os que normalmente apresentam melhores resultados, sendo estes HLS, HSV YCrCb, YUV, L\*a\*b\*, L\*u\*v\*, TSL, l1l2l3 e XYZ que são caracterizados de seguida.

#### 2.3.2.1. HSV e HLS

HSV e HLS são modelos de cor que se relacionam. Em ambos os casos, pode-se afirmar que os termos em que a cor é caracterizada, são os mais fáceis dos seres humanos interpretarem. O conceito de matiz é introduzido e representado por  $H$  (*Hue*), a componente de saturação representada por  $S$  (*Saturation*) e a ainda a componente do brilho esta representada por termos diferentes nos dois modelos  $V$  (*Value*) no modelo HSV e  $L$  (*Luminosity*) no modelo HLS (Hernández-Hernández *et al.*, 2016). As Equações (3), (4) e (5) permitem efetuar a transformação do espaço RGB para o espaço HSV (Hernández-Hernández *et al.*, 2016).

$$H = \arccos \frac{\frac{1}{2}(2R-G-B)}{\sqrt{(R-G)^2 - (R-G)(G-B)}} \quad (3)$$

$$S = \frac{\max(R,G,B) - \min(R,G,B)}{\max(R,G,B)} \quad (4)$$

$$V = \max(R, G, B) \quad (5)$$

### 2.3.2.2. YCrCb e YUV

YCrCb e YUV são modelos de cor igualmente interrelacionados. Estes são utilizados para aplicações de compressão e transmissão de imagens e vídeo. Nestes modelos, a decomposição da cor é feita de maneira diferente, sendo decomposta em intensidade de luz (Y), e dois canais correspondentes ao tom de cor que não são influenciados pelo brilho (Hernández-Hernández et al., 2016). A conversão entre os espaços de cor RGB e YCbCr é dada pelas Equações (6) e (7).

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0 \\ 120 \\ 120 \end{bmatrix} + \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (6)$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.000 & 0.000 & 1.400 \\ -0.169 & -0.343 & -0.711 \\ 1.000 & 1.765 & 0.000 \end{bmatrix} \begin{bmatrix} Y \\ Cb - 128 \\ Cr - 128 \end{bmatrix} \quad (7)$$

### 2.3.2.3. L\*a\*b\* e L\*u\*v\*

L\*a\*b\* e L\*u\*v\* são modelos criados com o propósito de satisfazer aplicações industriais. Foram criados pela CIE em 1976. Estes contêm também a componente de intensidade de luz (L), e duas componentes de tom de cor (a\*-b\* e u\*-v\*), que são definidas como transformações não lineares do modelo RGB. O principal objetivo destes modelos é obter uma representação da cor que seja uniforme à percepção humana (Hernández-Hernández et al., 2016).

#### 2.3.2.4. XYZ

O espaço de cores XYZ também foi definido pela CIE e é essencialmente baseado no modelo tricromático da retina humana. Consiste num sistema de primárias virtuais imaginárias conhecidas como X, Y e Z que incluem na totalidade todo o espaço de cores. X e Z tem não tem luminância. A luminância é contribuída apenas por Y. Todas as cores reais podem ser correspondidas por valores positivos destas três primárias (Corke, 2011).

Hernández-Hernández *et al.* (2016) mostraram que apesar dos canais XYZ se terem mostrado pouco promissores com resultados de classificação pouco aceitáveis, as combinações de dois canais mostraram resultados muito promissores. As coordenadas de cromaticidade correspondentes são dadas pela Equação (8).

$$x = \frac{X}{X+Y+Z}, y = \frac{Y}{X+Y+Z}, z = \frac{Z}{X+Y+Z} \quad (8)$$

#### 2.3.2.5. TSL e I1I2I3

Os espaços TSL e I1I2I3 foram criados para aplicações específicas. No caso do espaço I1I2I3, a aplicação proposta por Ohta *et al.* (1980) residiu na resolução de problemas de segmentação de cor. Para aplicações tais como reconhecimento facial foi criado o espaço TSL, que tem como componentes Matiz, Saturação e Luminosidade (Hernández-Hernández *et al.*, 2016).

### 2.4. Textura

Apesar de não haver uma definição universal para descrever a textura de uma imagem, aos olhos dos humanos a textura é uma característica perceptível e representa uma fonte de importante informação. De um modo geral, a textura pode ser referida como sendo um padrão visual complexo composto por diversas características como sub padrões, brilho, cor, declive, tamanho entre outras (Materka & Strzelecki, 1998). A análise de textura continua a ser um problema primário em processamento de imagem. Em certas aplicações tais como

reconhecimento de objetos e recolha de informação da imagem, a análise da textura continua a ser um problema fundamental (Zhang & Tan, 2002).

Os métodos de análise de textura podem ser divididos em quatro grupos:

- Estruturais;
- Estatísticos;
- Baseados em modelos;
- Transformações.

#### **2.4.1. Métodos estruturais**

Os métodos estruturais representam a textura através de primitivas (microtextura) e uma hierarquia de arranjos espaciais (macro textura), dessas mesmas primitivas definidas previamente. Para descrever a textura, além de serem definidas as primitivas, é requerida a definição das regras de colocação. A escolha de uma primitiva a partir de um grupo de primitivas e a probabilidade dessa primitiva estar colocada numa determinada localização pode ser uma função de localização ou de primitivas perto desse local (Materka & Strzelecki, 1998). As primitivas ou elementos mais comuns e mais utilizados são intensidade, área, excentricidade, orientação, alongamento, magnitude, compacidade, constante de Euler, etc. Métodos estruturais são mais adequados para descrições de macro texturas. Alguns dos métodos estruturais são (Zhang & Tan, 2002):

- Contribuição de perímetro e compacidade;
- Histograma invariante;
- Descritores de textura topológicos;
- Decomposição Morfológica.

### 2.4.2. Métodos espaciais

Ao contrário dos métodos estruturais, os métodos espaciais não tentam descrever a textura compreendendo a hierarquia estrutural da mesma. Estes representam a textura indiretamente através das propriedades não determinísticas, que dominam as distribuições entre os níveis de cinzento da imagem (Materka & Strzelecki, 1998). Muitos dos métodos espaciais são baseados no facto do sistema visual humano usar características estatísticas para a discriminação da textura incluído a estatística de primeira ordem, estatística de segunda ordem e estatística de ordem elevada (Zhang & Tan, 2002). Este facto foi descoberto por Julesz's (Julesz & Caelli, 1979).

Alguns métodos espaciais são:

- Parcelas polares e polarogramas;
- Estatística da textura de contornos em parcelas polares;
- Expansão Harmónica;
- Método digital ótico;
- Estatísticas de ordem elevada;
- Momentos invariantes;
- Método de distribuição de características;

### 2.4.3. Métodos baseados em modelos

Nestes métodos, as imagens de texturas são modeladas como métodos de probabilidade ou combinações lineares de funções base. Os coeficientes destes modelos são usados para caracterizar imagens de textura (Zhang & Tan, 2002).

Alguns destes modelos são:

- Modelos SAR e RISAR;
- Modelo de Markov;

- Modelo de Wold-like;
- Filtro Gabor multicanal;
- Piramide Steerable.

#### 2.4.4. Métodos baseados em transformações

Os métodos de análise de textura baseados em transformadas podem ser relativos às transformadas de Fourier, de Gabor e transformadas Wavelet, que representam uma imagem num espaço cujo sistema de coordenadas tem uma interpretação que é próxima às características de textura. Os métodos baseados na transformada de Fourier não permitem a obtenção de grandes resultados, visto que carecem da localização espacial. Por outro lado, os filtros de Gabor garantem meios para obter uma melhor localização espacial, contudo, a sua utilidade é relativamente limitada porque não existe um filtro de ordem 1 capaz de localizar uma estrutura espacial em texturas naturais. As transformadas Wavelet, quando comparadas com os filtros Gabor, apresentam certas vantagens, já que a variação da resolução espacial permite obter representações de textura em escalas mais favoráveis. Adicionalmente, existe uma elevada possibilidade na escolha da função, permitindo escolher a melhor função dependendo da aplicação (Materka & Strzelecki, 1998).

### 2.5. Reconhecimento de contornos

A deteção de contornos é uma operação de baixo nível utilizada em reconhecimento e processamento de imagens. As suas aplicações são diversas, mas o principal objetivo é localizar e identificar descontinuidades na imagem. Estas descontinuidades acontecem quando uma mudança abrupta na intensidade do pixel é detetada, caracterizando assim os contornos de objetos numa imagem. Estes contornos são usados para identificar objetos para propósito de segmentação e comparação (Voorhees & Poggio, 1987).

Identificar corretamente contornos numa imagem é muito útil, pois desta forma, todos os objetos podem ser localizados e certas propriedades básicas como área,

perímetro e forma podem ser medidas. Como a segmentação da imagem envolve o reconhecimento e classificação de objetos, o reconhecimento de contornos torna-se numa ferramenta fundamental (Nadernejad *et al.*, 2008).

### **2.5.1. Técnicas de reconhecimento de contornos**

De seguida são apresentados alguns dos algoritmos de deteção de contornos mais populares.

#### **2.5.1.1. Detetor de contornos Marr-Hildreth**

O detetor de contornos Marr-Hildreth era um operador de contornos muito popular antes de Canny publicar o seu artigo. É um operador baseado no gradiente que usa o operador de Laplace para obter a segunda derivada de uma imagem (Nadernejad *et al.*, 2008).

#### **2.5.1.2. Detetor de contornos Canny**

O detetor de contornos Canny é considerado, a nível global, o detetor standard da industria. Foi criado por John Canny para a sua dissertação de mestrado no MIT no ano de 1983, sendo que ainda tem melhores resultados do que muitos dos algoritmos desenvolvidos recentemente (Owens, 1997). Este método analisa o problema da deteção de contornos como sendo um problema de otimização de processamento. Assim sendo, reside numa função objetivo que tem de ser otimizada. No entanto, o problema apresentava uma solução muito complexa na forma de uma função exponencial. Assim, vários processos de otimização e aproximação foram implementados para melhorar o problema de encontrar os contornos (Nadernejad *et al.*, 2008).

### **2.5.1.3. Detecção de contornos baseada em funções booleanas e Threshold local**

Este detetor de contornos é fundamentalmente diferente da maioria dos detetores de contornos atuais. Este tem como vantagem o facto de considerar um limite (*threshold*) local e global para procurar contornos numa imagem. Contrariamente a outros detetores, este converte uma janela de pixéis num padrão binário baseado num *threshold* local. Desta forma pode aplicar máscaras pela imagem inteira para determinar se os contornos aparecem ou não num certo local. Assim, calculando o *threshold* numa base por pixel, o detetor teoricamente será menos sensível às variações da luz pela imagem (Nadernejad *et al.*, 2008).

### **2.5.1.4. Detecção de contornos usando distância euclidiana e ângulo de vetores**

Ao contrário da maioria dos detetores que trabalham numa escala de cinzentos (um canal em vez de três, RGB), este detetor trabalha com as cores. Desta forma, a imagem transporta mais informação permitindo ao detetor encontrar contornos em regiões de elevadas variações de cor, mas com baixa intensidade. Este detetor utiliza dois operadores: distância euclidiana e ângulo de vetores. A distância euclidiana é um bom operador para encontrar contornos baseados na intensidade e o ângulo de vetores é um operador que permite encontrar contornos baseados no contraste e na saturação. O detetor aplica ambos os operadores ao espaço de cores RGB de uma imagem, e depois combina os resultados dos dois baseando-se na quantidade de cor da região (Nadernejad *et al.*, 2008).

### **2.5.1.5. Detetor de contornos em espaços de cor RGB**

Outro caminho para a detecção de contornos usando espaços de cor é simplesmente alargar o uso de um detetor de contornos de intensidade para um espaço de cor. Este método procura usar os pontos fortes conhecidos dos detetores comuns e tenta superar as suas lacunas ou pontos menos fortes, fornecendo mais informação na



forma dos três canais de cor, ao contrário de apenas um canal de intensidade (Nadernejad et al. 2003)

## 2.6. Machine Learning

A *Machine Learning* estuda técnicas de aprendizagem automática com o objetivo de fazer previsões corretas baseadas em observações passadas. Por exemplo, suponha-se que se pretende desenvolver um filtro para emails que consegue distinguir *spam* de emails importantes. A melhor maneira de encarar este problema seria a seguinte: primeiro, começa-se por recolher o máximo de exemplos possíveis para ambos os casos. De seguida submetem-se estes exemplos, devidamente identificados como *spam* ou emails importantes, ao algoritmo de aprendizagem escolhido que vai automaticamente aplicar uma regra de classificação. Desta forma, dado um novo exemplo sem identificação, a regra vai tentar identificar se está perante um email *spam* ou um email importante. O objetivo reside na criação de uma regra que faça as previsões mais corretas nos novos exemplos a testar (Schapire, 2003).

De seguida é apresentada uma breve explicação de alguns dos algoritmos mais utilizados no âmbito de *machine learning*.

### 2.6.1. Algoritmo K-means Cluster

K-means é um algoritmo de aprendizagem sem supervisão que classifica um conjunto de dados num certo número de grupos ( $K$  clusters). Começa por selecionar aleatoriamente os centros de cada grupo ( $c$ ). Considerando  $X = X_1, X_2, X_3, \dots, X_n$ , os pontos do conjunto de dados e  $V = V_1, V_2, V_3, \dots, V_c$ , sendo  $c$  o número de grupos, com cada  $V_i, i \in 1, 2, \dots, c$ , obtem-se um vetor de dimensão  $n$ . As partições são calculadas usando a distância entre cada ponto  $X$  e o centro do grupo. As métricas de distâncias mais usadas são *squared Euclidean*, *Cityblock*, distância de Manhattan e coseno.

Os pontos são atribuídos ao centro do grupo, cuja distância é mínima entre todas as distâncias de todos os centros e o ponto. Os centros dos grupos são recalculados, tal como a distância entre cada ponto, para obter novos centros. Se nenhum ponto

foi reatribuído, o algoritmo pára, caso contrário os pontos são reatribuídos tal como descrito anteriormente.

### 2.6.2. K-nearest neighbors

KNN é um método supervisionado que classifica objetos baseado nos exemplos de treino mais próximos no espaço em questão. É um método simples de *machine learning* usado para prever características de qualquer tipo. Um objeto é classificado pelo voto maioritário dos seus vizinhos, sendo atribuída a classe mais comum entre os seus vizinhos mais próximos. Para cada ponto ser atribuído, o algoritmo usa a informação mais próxima para estimativa, aproveitando-se da informação do local e tomando uma decisão. A proximidade pode ser definida usando qualquer métrica de distância. As mais comuns são a distância euclidiana, a distância Cityblock ou Manhattan, distância Chebychev, coseno e distância Minkowsky (Xu *et al.*, 1992). Nas Equações 9 a 11 apresentam-se as expressões de cálculo das distâncias:

Distância euclidiana,

$$\sqrt{\sum_{i=1}^k (x_i - y_i)^2} \tag{9}$$

Distância Cityblock ou Manhattan,

$$\sum_{i=1}^k |x_i - y_i| \tag{10}$$

Distância Minkowsky,

$$(\sum_{i=1}^k (|x_i - y_i|^q))^{\frac{1}{q}} \tag{11}$$

### 2.6.3. Máquinas de suporte vetorial (*Support Vector Machine* - SVM)

As máquinas de suporte vetorial (*Support Vector Machine* - SVM) são modelos de aprendizagem supervisionados (função deduzida a partir de dados de treino identificados), utilizados para classificação binária e análise de regressão. Dado um conjunto de exemplos de treino, pertencentes a uma de duas categorias, também denominadas classes, a SVM constrói um modelo que atribui novos exemplos a uma das duas categorias. As funções de semelhança da SVM são as chamadas funções Kernel para análise de padrões. Na prática, os descritores de treino são mapeados pela função Kernel em dimensões mais elevadas até que um hiper-plano consiga separar as duas classes. As mesmas transformações são aplicadas a qualquer descritor, mapeando-o no espaço definido. Assim, o lado do hiper-plano onde o descritor foi mapeado define-se como classe. As funções Kernel mais usadas são: linear, polinomial, função de base radial (*Radial basis function* - RBF) e sigmóide (Li & Wang, 2012).

#### 2.6.3.1. Algoritmos Kernel

Os métodos Kernel utilizam técnicas desde otimização, estatística, e análise funcional para obter generalidade, flexibilidade e desempenho máximo. Estes algoritmos são diferentes das técnicas mais antigas usadas em aprendizagem máquina (*machine learning*) em vários aspetos: por exemplo, estes são baseados explicitamente em métodos teóricos contrariamente a analogias soltas/largas com sistemas de aprendizagem naturais ou outras heurísticas. Os métodos kernel trazem garantias do seu desempenho e tem construção modular que permite separadamente implementar e analisar os seus componentes. (Schölkopf & Cristianini, 2002)

## 2.7. Estudos científicos na temática

Neste ponto e no seguimento do conteúdo apresentado durante este capítulo, é relevante apresentar alguns dos algoritmos já criados e que se assemelham ao pretendido com o trabalho desenvolvido na presente dissertação.

Bulanon *et al.* (2002) criou um algoritmo de segmentação de imagem para diferenciação e localização de maçãs Fuji na árvore. Este algoritmo consiste em descobrir o valor ótimo de limite (*threshold*) de forma a conseguir separar, com o máximo de precisão possível, o objeto (fruto) do fundo (folhas, ramos, solo, céu). O estudo da imagem é feito pela cor. A análise da cor nas imagens foi feita com um método usado na transmissão dos sinais de televisão (Awcock & Thomas, 1996). Como o fruto em questão é vermelho, apenas foram utilizados os parâmetros luminância  $Y$  e diferença de vermelhos  $CR$ . As Equações (12) e (13) permitem o cálculo destes parâmetros.

$$Y = 0,3R + 0,6G + 0,1B \quad (12)$$

$$C_R = R - Y \quad (13)$$

Sendo que  $R$ ,  $G$  e  $B$  são os valores de intensidade que variam no intervalo de 0 a 255 das cores vermelho, verde e azul respetivamente.

Zhao *et al.* (2005) construíram um algoritmo para reconhecimento de frutos na árvore. As ferramentas principais foram a textura e a cor. O algoritmo começa por identificar áreas com a mesma textura, baseada no contraste. Como áreas de diferentes cores podem ter texturas semelhantes, a identificação da fruta não podia apenas ser feita com textura. Assim, uma técnica de cálculo foi adicionada para identificar cada pixel com valores de textura semelhantes. Estes valores são depois usados na técnica de deteção de contornos, ajudando assim, a isolar áreas com a mesma textura.

As imagens adquiridas são também sujeitas a uma transformação relativamente à cor. A Equação (14) é equação de vermelhos, *redness*, aplicada às imagens ajudando a salientar os contrastes entre o fruto e todo o resto.

$$r = 3R - (G + B) \quad (14)$$

Zhao *et al.* (2012) apresentaram um estudo que descreve a criação de um sistema baseado numa máquina de suporte vetorial que atua em tempo real. Este sistema foi concebido com o propósito de guiar um robô de apanha de maçã Fuji vermelha. O principal objetivo deste sistema reside em colmatar uma das principais falhas detetadas em estudos anteriores, a incapacidade de um sistema de reconhecimento para contrariar as constantes mudanças nas condições de iluminação proporcionadas pela luz natural do sol. O estudo apresentado é composto por vários processos antes de chegar a parte de treino da máquina de suporte vetorial, começando por uma melhoria gráfica da imagem através da aplicação de um filtro de vetor médio. Este método destaca os objetos em primeiro plano, neste caso as maçãs, e elimina ou enfraquece o ruído existente no segundo plano, ou seja, o fundo (Wang *et al.*, 2009). A imagem passa ainda por dois processos de segmentação, o primeiro baseado no ponto inicial de crescimento da maçã, e o segundo baseado nas propriedades da cor. São ainda aplicadas operações matemáticas morfológicas de fecho e abertura para preencher pequenos buracos e limpar pontos isolados respetivamente. Por último, são extraídas as características da cor, usando o espaço de cores HLS. A Figura 6 e Tabela 2 apresentam um exemplo das imagens usadas para o algoritmo e os resultados obtidos, respetivamente.



Figura 6 - Exemplo de reconhecimento baseado na SVM (Wang *et al.*, 2012).

A Figura 6 apresenta uma imagem exemplo do reconhecimento automático baseado na máquina de suporte vetorial. Na imagem é possível verificar que apenas as maçãs maiores e mais perto da câmara, no momento da captação de imagem, são reconhecidas. É visível que um grande conjunto de frutos com características diferentes não é reconhecido. Desta forma é possível afirmar que neste caso, as rigorosas condições impostas, permitem ao algoritmo detetar apenas as maçãs que não estão obstruídas por folhas, ramos ou mesmo ocultas por sombras.

Tabela 2 - Taxa média de reconhecimento da máquina de suporte vetorial para diferentes funções Kernel (Wang *et al.*, 2012).

	<i>Polinomial</i>	<i>RBF</i>	<i>Sigmóide</i>
<i>Taxa média de reconhecimento [%]</i>	85,7	93,3	57,2
<i>Tempo de operação [mseg]</i>	268	256	175

A Tabela 2 contém os resultados da taxa média de reconhecimento e o tempo de operação da máquina de suporte vetorial para diferentes funções Kernel. Estes resultados são obtidos com base em ambas as características, cor e forma.

Wei *et al.* (2014) criaram um método automático de extração de fruta em ambientes agrícolas complexos. O espaço de cores escolhido para trabalhar foi o OHTA. Comparado aos espaços de cor tradicionais HIS e HSV, a relação de conversão entre o espaço de cores OHTA e o espaço de cores RGB é linear (Wei *et al.*, 2014). Um algoritmo de segmentação baseado em Otsu foi o método escolhido. Este método foca-se mais nas características da cor, calculando o limite (*threshold*) mais indicado para melhor limpar o complexo fundo agrícola e extrair com maior precisão o fruto.

Song *et al.* (2014) descrevem um novo método para localizar e contar pimentos verdes em ambientes complexos em dois passos. No primeiro passo é feita a localização das frutas na imagem e no segundo passo são combinados diferentes ângulos da imagem para aumentar o rácio de deteção das frutas. A abordagem tomada para encontrar os pimentos, é baseada na combinação da localização de

pontos de interesse, aplicando um complexo descritor de características de elevada dimensão de um detalhe em torno do ponto de interesse, e a utilização de um *bag-of-words* para classificar esse detalhe.

Maldonado & Barbosa (2016) construíram um algoritmo para automaticamente contar o número de frutas na árvore. Primeiro, foram adquiridas as representações *bas-relief* através da combinação das técnicas de conversão do modelo de cores (*color model conversion*), limitação (*thresholding*), histograma, equalização (*equalization*) e filtragem espacial (*spatial filtering*) com operadores de Laplace, Sobel e Gaussian blur. Tendo em consideração que as fotos foram obtidas no exterior e são suscetíveis a variações nas condições de iluminação, foram usados métodos de ajuste do brilho, métodos esses conhecidos como transformação logarítmica (*log transformation*) e equalização do histograma de níveis de cinzento (*gray levels histogram equalization*) (Gonzales e Woods, 2000; Kurtulmus et al., 2011; Savvides & Kumar, 2003), que foram aplicados à camada valor no modelo de cores HSV.

O passo mais complexo e desafiante no processamento de imagens digitais é a segmentação da característica desejada (Maldonado e Barbosa, 2016). Isto torna-se ainda mais difícil quando a aquisição de imagem ocorre em condições não padronizadas (Zhao *et al.*, 2015). Neste caso específico, a extração dos frutos na imagem foi efetuada com a informação sobre o brilho extraída a partir da camada valor (*V*) no espaço de cores HSV e pela técnica de equalização do histograma (*histogram equalization*). O passo seguinte residiu na análise da textura recorrendo aos filtros de Laplace e de Sobel. Estas técnicas consistem na convolução das imagens por meios de máscaras predefinidas. O solo foi removido recorrendo ao intervalo de cores aplicado na camada matiz (*H*) no modelo de cores HSV. Este processo foi possível porque a diferença dos valores das cores entre o solo e o resto da imagem é elevada (Maldonado & Barbosa, 2016).

### **2.7.1. Outros estudos científicos na temática da agricultura de precisão**

Jiménez *et al.* (1999) apresentaram um estudo para deteção de frutos. O estudo apresentado era capaz de reconhecer frutas esféricas em condições naturais. O sistema de reconhecimento conta com um sensor laser (modelo de cálculo de

distância) e um algoritmo de análise de cor e forma para localizar a fruta. O cálculo da posição da fruta é efetuado após as fases de reconhecimento.

Morimoto *et al.* (2000) apresentaram um estudo com um novo método para avaliar a forma das frutas em termos quantitativos. Este estudo tinha três componentes fundamentais, o uso de atratores, dimensão fractal e redes neuronais. Foram criados perfis unidimensionais para caracterizar a forma da fruta. Os conceitos de dimensão fractal e atrator foram aplicados para identificar de forma quantitativa as irregularidades dos perfis unidimensionais. Uma rede neuronal de três camadas foi ainda aplicada para avaliar essas mesmas irregularidades.

Leemans *et al.* (2002) apresentaram um estudo para classificação de fruta de acordo com a sua qualidade exterior. Essa classificação é obtida após vários passos, como aquisição de imagem, classificação da cor do solo, segmentação das imperfeições, caracterização dessas imperfeições e por fim, a classificação das frutas.

Zhang & Wu (2012) expõem um estudo para classificação de frutos recorrendo à utilização de visão computacional e uma máquina de suporte vetorial de várias classes. O estudo apresentado consiste nos seguintes processos: remoção do segundo plano, fundo, extração das características de cor, forma, textura e análise da característica principal. Com estas características foram criadas três máquinas de suporte vetorial sendo estas Winner-Takes-All (WTA), Max-Wins-Voting (MWV) e Directed-Acyclic Graph (DAG), para as quais foram escolhidos três tipos de funções kernel diferentes: linear, polinomial e de base radial. A Tabela 3 apresenta os resultados de precisão do estudo.

Tabela 3 - Taxa média de reconhecimento [%] da máquina de suporte vetorial para diferentes funções Kernel (Zhang & Wu, 2012).

SVM	Linear	Polinomial	RBF
WTA	48,1	61,7	55,4
MWV	53,5	75,6	88,2
DAG	53,5	70,1	84,0



Na Figura 7 é possível observar as amostras de imagens dos diferentes frutos que se pretendem classificar. De notar, que como referenciado anteriormente, o segundo plano é removido na sua totalidade, fazendo com que não haja a necessidade de análise de diferentes objetos com propósito de indicar o que é fruto ou não. No caso exposto por Zhang & Wu (2012), apenas existe a necessidade de analisar as características de cada fruto.













Types	Sample Image	Types	Sample Image	Types	Sample Image
Yellow Bananas		Golden Pineapple		Black Grape	
Green Plantains		Passion Fruits		Black berry	
Tangerines		Bosc Pears		Blue berry	
Hass Avocado		Anjou Pears		Straw berry	

Figura 7 - Diferentes tipos de frutos e respectiva amostra gráfica (Zhang & Wu, 2012).

Bac *et al.* (2014) apresentaram um estudo para um braço robótico capaz da recolha autónoma de pimentos. Este braço robótico é auxiliado por iluminação artificial, fios como pistas visuais e um algoritmo com vários componentes de segmentação e análise de imagem. O principal objetivo deste braço robótico é evitar a colisão com a vegetação e recolher apenas o fruto.



## 3. Algoritmo de Reconhecimento e Previsão de Carga de Frutos

### 3.1. Introdução

Neste capítulo é apresentado todo o procedimento do desenvolvimento do algoritmo. São descritas as várias fases de análise e processamento das imagens que permitem obter os dados para treinar a SVM e também todas as características que integram a SVM, desde conjuntos de treino à função utilizada.

### 3.2. Algoritmo de reconhecimento de fruta

Um algoritmo consiste num conjunto de processos necessários para resolver um problema. No trabalho em questão, o problema consiste na identificação de frutos em árvore, neste caso pêsegos, através de fotografias. Várias adversidades são encontradas no meio natural, descritas com mais detalhe ao longo do capítulo, e como tal foi necessário submeter as imagens a vários processos de análise e processamento para esbater a influência das adversidades no processo de identificação dos frutos.

#### 3.2.1. Limite (*threshold*) de Cor

Quando se pensa em pêsegos, pensa-se em cores vivas, cores quentes como o amarelo, vermelho e laranja. Apesar de as diferentes qualidades de pêsegos que existem admitirem uma gama de cores que vai além destas três, estas não deixam

de ser as predominantes quando se pensa neste fruto. Por outro lado, quando se pensa no meio ambiente, pensa-se no azul do céu, o verde das plantas, entre outras. Assim pode-se afirmar que a cor é um dos maiores aliados na resolução deste trabalho, sendo ela uma das características fundamentais.

O limite (*threshold*) de cor foi feito através da *toolbox* do Matlab *Color Threshold*. Esta *toolbox* permite trabalhar em quatro espaços de cor diferentes, sendo estes RGB, HSV, YCbCr e  $L^*a^*b^*$ . Todos eles já foram descritos anteriormente e para este trabalho o espaço escolhido foi o HSV. Os seus três planos: contraste (*hue*), saturação (*saturation*) e brilho (*value*), permitem uma definição mais precisa das características que se pretendem separar. A Figura 8 apresenta um exemplo da aplicação da *toolbox* do MatLab de limite de cor, no plano HSV, a frutos.

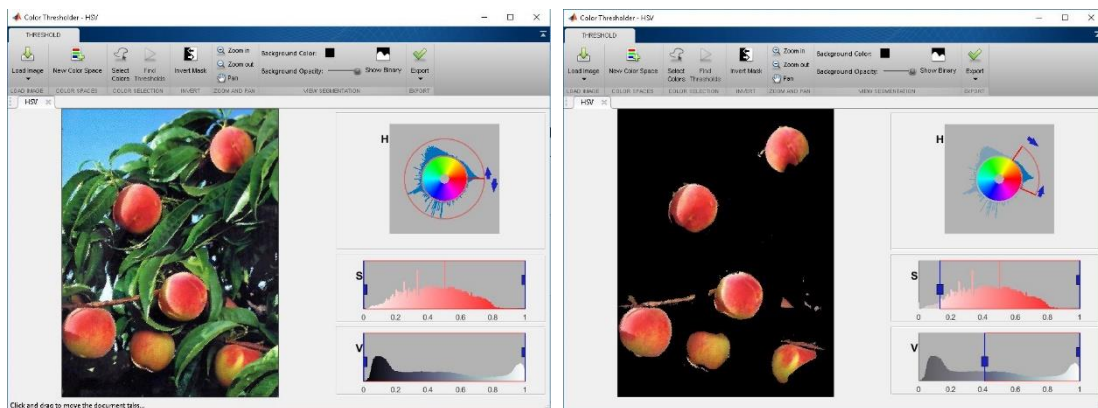


Figura 8 - Limite de cor no plano HSV.

Como se pode ver na Figura 8, a *toolbox* permite não só delimitar o contraste, mas também a saturação e o brilho. Ainda no exemplo, apesar de se verificar, como referido anteriormente, que a cor desempenha um papel fundamental, continua a haver objetos indesejados na imagem, mais especificamente pequenos ramos da árvore cuja cor é similar à dos frutos.

De seguida é apresentado o algoritmo extraído da *toolbox* usado para realizar o *threshold* dos três planos, *H*, *S* e *V*.

```
function [BW,maskedRGBImage] = createMask(RGB)
%createMask Threshold RGB image using auto-generated code from
colorThresholder app.
%-----
% Convert RGB image to chosen color space
I = rgb2hsv(RGB);
% Define thresholds for channel 1 based on histogram settings
channel1Min = 0.833;
channel1Max = 0.172;
% Define thresholds for channel 2 based on histogram settings
channel2Min = 0.197;
channel2Max = 1.000;
% Define thresholds for channel 3 based on histogram settings
channel3Min = 0.398;
channel3Max = 1.000;
% Create mask based on chosen histogram thresholds
BW = ( (I(:,:,1) >= channel1Min) | (I(:,:,1) <= channel1Max) ) & ...
      (I(:,:,2) >= channel2Min) & (I(:,:,2) <= channel2Max) & ...
      (I(:,:,3) >= channel3Min) & (I(:,:,3) <= channel3Max);
% Initialize output masked image based on input image.
maskedRGBImage = RGB;
% Set background pixels where BW is false to zero.
maskedRGBImage(repmat(~BW,[1 1 3])) = 0;
```

### 3.2.2. Segmentação e análise das imagens

Após o *threshold* aplicado no subcapítulo anterior, e com o propósito de eliminar o máximo de objetos indesejados da imagem, são aplicadas outras transformações. Estas têm como objetivo eliminar pequenas manchas e algum ruído existente. A Figura 9 apresenta um exemplo de imagem com frutos após segmentação.

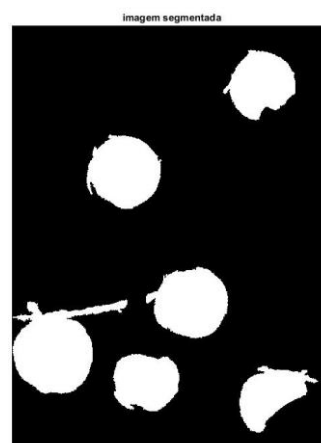


Figura 9 - Exemplo de imagem após segmentação.

No caso apresentado na Figura 9, pode-se visualizar que como o pêssgo e o ramo estão sobrepostos e dentro do mesmo espectro de cores, o algoritmo deteta ramo e pêssgo como um objeto só. Contudo, é possível verificar que o resto da imagem ficou sem pequenos objetos ou ruído. O excerto de código usado para fazer estas transformações é apresentado abaixo.

```
bw = bwareaopen(I2,3500); %remover objetos com menos de 3500 pixeis;  
se = strel('disk',1);  
bw = imclose(bw,se); %preencher pequenos espaços;  
bw = imdilate(bw,se);  
bw = imfill(bw,'holes'); %preencher buracos para facilitar a  
utilização da função regionprops;
```

### 3.2.3. Detecção de arestas

Neste subcapítulo expõe-se a próxima transformação do algoritmo que consiste na análise e detecção de arestas. Para este propósito foi escolhido um operador de Sobel. Este operador faz a análise em regiões de 3x3 de toda a imagem, sendo o operador exposto de seguida é aquele que melhores resultados apresentou.

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

Mais uma vez, um pequeno excerto de código é apresentado mostrando a transformação executada sobre a imagem.

```
Skernel_2=[-1 -1 -1 ; -1 8 -1 ; -1 -1 -1];  
Sobel = filter2 (Skernel_2, bw);
```

A Figura 10 mostra os resultados obtidos após a aplicação do operador de Sobel.

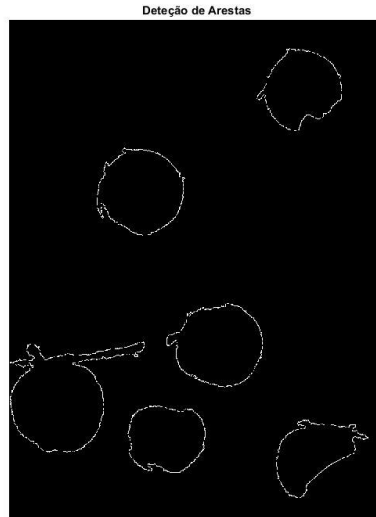


Figura 10 - Resultados obtidos como operador de Sobel.

#### 3.2.4. Probabilidade de um objeto como redondo

Esta secção do algoritmo é baseada no mesmo princípio da detecção de arestas, mas neste caso o contorno serve para calcular a área do objeto e a probabilidade desse mesmo objeto ser redondo. Este é um cálculo importante pois é com ele que se obtêm os valores usados para treinar a SVM mais a frente.

Para este cálculo são usadas duas funções do Matlab, *bwboundaries* e *regionprops*. A primeira função, tem o mesmo princípio do processo explicado no subcapítulo anterior, visto que é usada para detetar contornos. A segunda é usada para calcular a área, que é dada em número de pixéis dentro de um certo contorno, e o centro desse mesmo contorno. Com estes valores e através de fórmulas matemáticas, é estimado o perímetro, tornando assim possível calcular a probabilidade de um objeto ser redondo. A Equação (16), mostra o cálculo efetuado para obter essa percentagem.

$$Metrica = \frac{4 \times \pi \times \text{Área}}{\text{Perímetro}^2} \quad (16)$$

Os resultados obtidos a nível de contornos são em tudo semelhantes aos obtidos pelo operador de Sobel exposto no subcapítulo 3.2.3. A Figura 11 mostra um exemplo dos resultados obtidos após efetuadas as transformações explicadas ao longo do subcapítulo 3.2.

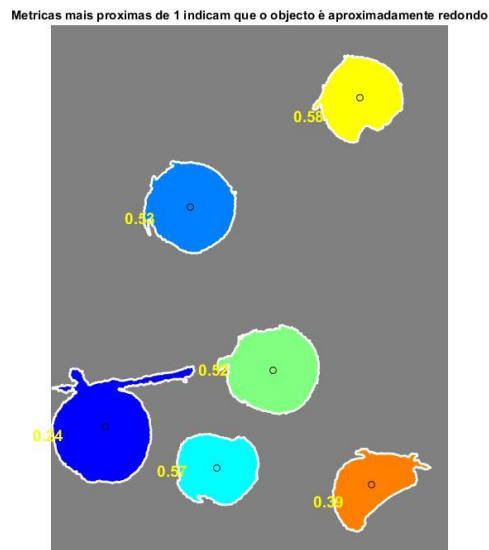


Figura 11 - Exemplo de contornos e respetiva métrica.

O trecho de código abaixo destina-se à aplicação dos processos efetuados para obter os resultados da Figura 11.

```
[B,L] = bwboundaries(bw, 'noholes');  
imshow(label2rgb(L, @jet, [.5 .5 .5])) % mostrar objectos com  
diferentes cores e desenhar os contornos;  
hold on  
for k = 1:length(B) 4  
    boundary = B{k};  
    plot(boundary(:,2), boundary(:,1), 'w', 'LineWidth', 2)  
end  
stats = regionprops(L, 'Area', 'Centroid');  
threshold = 0;  
for k = 1:length(B)  
    boundary = B{k}; % Obter as coordenadas (X,Y) dos contornos  
correspondentes à iteração k;  
    delta_sq = diff(boundary).^2; % calcular uma simples  
estimativa do perímetro do objeto;  
    perimeter = sum(sqrt(sum(delta_sq,2)));
```



```

area = stats(k).Area; % obter o calculo da área correspondente à
iteração K;
metric = 4*pi*area/perimeter^2; % Calculo da métrica
(probabilidade de ser redondo);
metric_string = sprintf('%2.2f',metric); % Mostrar resultados;
stats(k).metric=metric;
% Marcar objectos acima do threshold;
if metric > threshold
    centroid = stats(k).Centroid;
    plot(centroid(1),centroid(2), 'ko');
end
text(boundary(1,2)-
35,boundary(1,1)+13,metric_string, 'Color','y',...
'FontSize',14, 'FontWeight', 'bold');
end
title(['Metricas mais proximas de 1 indicam que o objecto é
aproximadamente redondo']);

```

Na Tabela 4 são apresentados os resultados da Figura 11 que são armazenados numa estrutura. Essa estrutura é obtida, guardando os valores dos objetos da imagem da esquerda para a direita, como se pode constatar comparando os valores das métricas na imagem com os valores das métricas que estão na Figura 11.

Tabela 4 -Resultados referentes à Figura 9.

<i>Imgref</i>	<i>Area</i>	<i>Metrica</i>
<i>Img1</i>	28968	0,2365
<i>Img1</i>	20767	0,5294
<i>Img1</i>	15385	0,5655
<i>Img1</i>	21062	0,5189
<i>Img1</i>	16822	0,5843
<i>Img1</i>	14250	0,3878

### 3.2.5. Nota Conclusiva

A Figura 12 mostra o fluxograma da primeira parte do algoritmo exposta ao longo do subcapítulo 3.2. Foi ainda adicionado um bloco condicional, sendo  $k$  o número de imagens numa certa diretoria, com o objetivo de automatizar o processo fazendo com que o algoritmo se torne numa função que analisa não só uma imagem, mas  $k$  imagens.

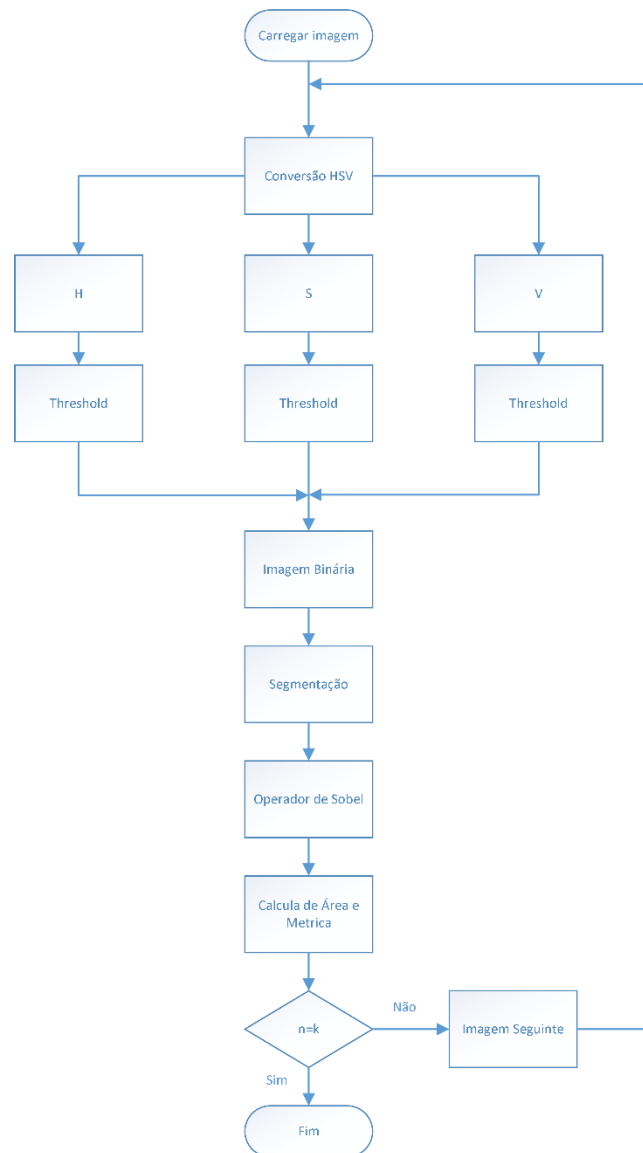


Figura 12 - Fluxograma da função SVM DATA

### 3.3. Aprendizagem máquina (*Machine Learning*) - Treino da SVM

Neste subcapítulo é descrito o processo de classificação dos resultados obtidos pela função do subcapítulo 3.2. A classificação destes resultados e a sua identificação como conjuntos de treino e conjuntos de teste, garantem os exemplos de treino

necessários para treinar a SVM de modo a obter os melhores resultados possíveis referentes à identificação de pêsegos.

### 3.3.1. Classificação de exemplos de treino

Como referido no subcapítulo 2.3.6, dado um conjunto de exemplos de treino, pertencentes a uma de duas categorias, também denominadas classes, a SVM constrói um modelo que atribui novos exemplos a uma das duas categorias. As variáveis contempladas para o treino da SVM foram a área e métrica, como se pode ver na Tabela 5. Cada linha da Tabela 5 corresponde a um exemplo de treino que pode ser classificado como zero, não pêsego, ou um, pêsego. No exemplo mostrado ao longo do subcapítulo 3.2, aquando da comparação da Figura 11 com a Tabela 5, pode-se constatar que todos os exemplos de treino correspondem a pêsegos.

A Tabela 5 mostra a classificação dos resultados para a Figura 11.

Tabela 5 - Classificação dos exemplos de treino.

<i>Imgref</i>	<i>Área</i>	<i>Métrica</i>	<i>Classificação</i>
<i>img1</i>	28968	0,2365	1
<i>img1</i>	20767	0,5294	1
<i>img1</i>	15385	0,5655	1
<i>img1</i>	21062	0,5189	1
<i>img1</i>	16822	0,5843	1
<i>img1</i>	14250	0,3878	1

A coluna da classificação é preenchida manualmente, sendo que esta será a coluna que indicará à SVM quais os exemplos de treino que correspondem a pêsegos ou a não pêsegos. É apresentado de seguida o excerto de código que corre a função apresentada no subcapítulo 3.2, criando assim a tabela que permite classificar os exemplos de treino. De forma muito resumida, este excerto de código aplica a função a todas as imagens presentes na sua diretoria.

```
D = dir('*.*jpg');
n=length(D);
Imagestruct(n).Img_name=[];
for i = 1:n
    Imagestruct(i).Img_name=D(i).name;
end
temp=1;
svm_data(293).Imgref=[];
for i=1:n
    RGB = imread(Imagestruct(i).Img_name);
    [A,C,M,F] = SVM_DATA(RGB);
    temp1=0;
    for j=temp:temp-1+length(A)
        svm_data(j).Area=A(1+temp1);
        svm_data(j).Metric=M(1+temp1);
        svm_data(j).Imgref=strcat('img',int2str(i));
        temp1=temp1+1;
    end
    print(F, '-dpdf',
strcat('img',int2str(i),'_contornos.pdf'));
    temp=length(A)+temp;
end
```

### 3.3.2. Modelo SVM e a sua aplicação

#### 3.3.2.1. Normalização dos dados

Posteriormente à classificação de todos os exemplos de treino das várias imagens e antes de selecionar os conjuntos de treino e teste, é feita uma normalização dos dados. Como podemos constatar na Tabela 5, os dados da coluna Métrica e os dados da coluna Classificação variam entre zero e um. Assim, para que os valores da coluna Área variem também apenas entre zero e um, foi feita uma normalização dos dados da Área. A normalização é feita por dois motivos: o primeiro é evitar que características em gamas de valores numéricos elevados dominem características em gamas de valores numéricos mais pequenos. Segundo, evitar dificuldades numéricas durante o cálculo. Estas dificuldades podem verifica-se com a utilização de certas funções kernel, como por exemplo as funções lineares e as funções polinomiais, em que os seus valores dependem dos produtos internos dos vetores característica, fazendo com que características com valores elevados possam causar problemas numéricos. (Hsu *et al.*, 2016). Na Tabela 6 é apresentada a Tabela 5 com os dados normalizados.

Tabela 6 - Normalização dos dados.

<i>Imgref</i>	<i>Área</i>	<i>Métrica</i>	<i>Classificação</i>
<i>Img1</i>	0,1123	0,2365	1
<i>Img1</i>	0,0755	0,5294	1
<i>Img1</i>	0,0513	0,5655	1
<i>Img1</i>	0,0768	0,51896	1
<i>Img1</i>	0,0577	0,5843	1
<i>Img1</i>	0,0462	0,3878	1

Os dados de normalização apresentados na Tabela 6 foram obtidos através da equação (17), aplicada de igual forma a todos os valores da coluna Área.

$$Normalização = \frac{Área - \min(Área)}{\max(Área) - \min(Área)} \quad (17)$$

Em que:

max - corresponde ao maior valor da coluna;

min - corresponde ao menor valor da coluna;

### 3.3.2.2. Criação do modelo SVM

Após a normalização dos dados e selecionados os conjuntos de treino e teste, é criado o modelo da SVM e feita a previsão para posterior comparação de resultados. A SVM escolhida foi de classificação binária, devido às características dos conjuntos de treino e o que cada um pode representar. De todas as funções disponibilizadas pelo software, a função escolhida para treinar a SVM foi uma função quadrática, ou polinomial de segundo grau, sendo esta a que apresentou melhores resultados para o caso em estudo. Após criado o modelo da SVM, é executada a função de classificação sobre o conjunto de teste, que mais tarde é comparado com os valores inseridos manualmente. Todo este processo está representado no pequeno excerto de código que se pode observar de seguida.

```
SVMModel = svmtrain(X,Y,'kernel_function','q4uadratic','options',  
...  
    statset('Display','final','MaxIter',500000),'tolkkt',1e-12);  
Previsao=svmclassify(SVMModel,W);  
p = Previsao';
```

Criado o modelo SVM e feita a previsão dos resultados, resta apenas proceder ao cálculo estatístico da assertividade do modelo em questão. Esta assertividade é calculada comparando os resultados gerados pelo modelo SVM com a classificação inserida manualmente. Esta comparação é feita através de coeficientes de correlação, mais especificamente os coeficientes de correlação de Kendall, Pearson e Spearman. O pequeno excerto de código apresentado abaixo mostra a utilização dos mesmos.

```
[dados_kendal,~] = corr(cmos',res','Type','Kendall');  
[dados_pearson,~] = corr(cmos',res','Type','Pearson');  
[dados_spearman,~] = corr(cmos',res','Type','Spearman');
```

### 3.3.3. Nota Conclusiva

A Figura 13 mostra um diagrama de blocos, que representa as operações feitas pelo algoritmo para obter os resultados finais, após feita toda a análise e classificação dos dados extraídos das imagens. De referir que o exemplo trabalhado ao longo do capítulo 3, mais especificamente a Figura 9, não faz parte das imagens analisadas para o modelo SVM criado neste trabalho. A figura em questão foi escolhida pela sua simplicidade para facilitar a apresentação de todos os processos efetuados ao longo do algoritmo. Durante o próximo capítulo, capítulo 4, serão apresentados exemplos reais tal como toda a informação relevante sobre as características que influenciaram a escolha dos conjuntos de treino e de teste.



Figura 13 - Diagrama de blocos do algoritmo final.

## 4. Análise e Discussão de Resultados

### 4.1. Introdução

Neste capítulo são apresentados os resultados provenientes dos testes do algoritmo. Estes testes foram efetuados nos pomares de duas cultivares diferentes de pêsego, *Diamond Princess* e *Sweet Dream*, na Quinta Nova, localizada na proximidade da localidade do Ferro. Estas duas cultivares encontram-se em condições de luminosidade diferentes. As fotografias utilizadas são fotografias de fruto em árvore, recolhidas em condições naturais e em dias diferentes, com uma câmara NIKON D3200 com 24 bits de profundidade. Nos exemplos que serão apresentados ao longo do capítulo, pode-se verificar que as árvores em questão são árvores de elevada folhagem, o que por sua vez, proporciona elevados níveis de oclusão e sombras, dificultando assim o processo de segmentação da imagem. Ao contrário da maior parte dos estudos, neste trabalho não foram estabelecidas condições para que fosse encontrado apenas um conjunto de pêsegos. Após todo um estudo focado na deteção do fruto em árvore, foi ainda adicionada uma componente para previsão de produção, na expectativa de obter um valor de produção aproximado, em toneladas por hectare, com base no volume dos pêsegos encontrados.

### 4.2. Testes e Resultados

Como referido no subcapítulo 4.1, as fotografias foram recolhidas em condições naturais e em dois dias diferentes, 5 de julho de 2017 entre as 6h28m e as 7h27m, dia #1, e dia 20 de julho de 2017 entre as 6h26m e as 7h25m, dia #2, proporcionando condições de luminosidade diferentes. Para tentar minimizar estas diferenças de luminosidade em ambos os dias, as fotografias foram tiradas durante o período da manhã e relativamente à mesma hora, antes do sol incidir nas árvores (cerca da 7h00m). Contudo, e tal como a expressão “condições naturais” indica, as diferenças

de luminosidade foram significativas, devido à variação na hora do nascer do sol. Esta variação de luminosidade provocou grandes dificuldades na parte da segmentação, fazendo que tivesse sido necessário aplicar uma correção, aumento de cerca de 40%, na luminosidade das imagens do segundo dia. As Figuras 14 e 15 mostram essas variações de luminosidade e as diferenças obtidas na segmentação da imagem nas diferentes condições.





(a)

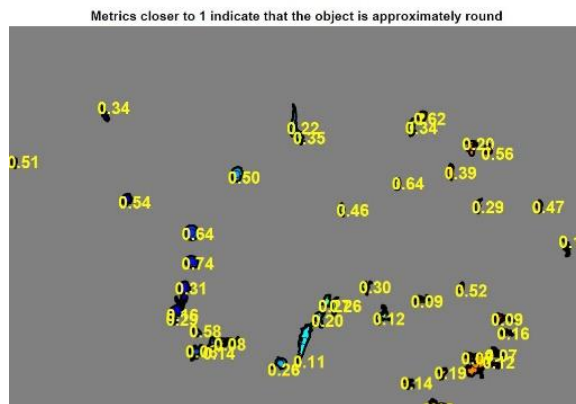


(b)

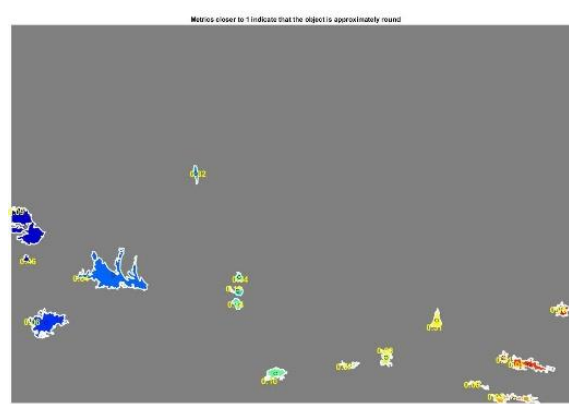


(c)

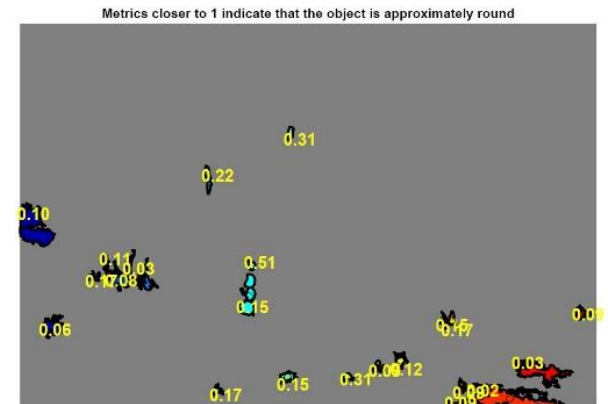
Figura 15 - Variação de luminosidade nos diferentes dias: a) dia #1, b) dia #2, c) correção da luminosidade do dia dois.



(a)



(b)



(c)

Figura 14 - Variação dos resultados de segmentação para as imagens da Figura 14: a) Segmentação da Figura 14 a), b) Segmentação da Figura 14 b), c) Segmentação da Figura 14 c).

Também referido anteriormente, as fotografias foram recolhidas em duas cultivares diferentes, dois tipos de pêsegos, o que implica frutos com diferentes características. A principal característica que os diferencia é o tamanho, como se pode visualizar na Figura 16. A fotografia da Figura 16 a) foi retirada num campo com a cultivar 1, cultivar *Sweet Dream* onde crescem frutos de diâmetro médio de 9 cm e peso que pode variar entre as 350g e as 400g, enquanto que a Figura 16 b) foi retirada no campo com a cultivar 2, cultivar *Diamond Princess* onde crescem frutos com um diâmetro médio de 7,6 cm e peso médio de 170g. Assim, para que as diferentes características entre cultivares, e as diferentes luminosidades para os dois dias não influenciassem os resultados obtidos, os testes foram separados em três blocos: cultivar 2 no dia 5 de julho de 2017 (dia #1), cultivar 1 no dia 20 de julho de 2017 (dia #2) e cultivar 2 no dia 20 de julho de 2017 (dia #2). As imagens obtidas na Cultivar 1 no dia #1 foram desprezadas face à sua reduzida qualidade. Desta forma, foram obtidos três modelos SVM, um para cada caso em estudo, de forma a obter os resultados mais precisos possíveis para cada cultivar. Os resultados referentes a cada bloco serão apresentados na secção 4.2.1.



(a)



(b)

Figura 16 - Exemplo das diferentes dimensões dos pêsegos: a) Cultivar 1, b) Cultivar 2.

#### **4.2.1. Apresentação dos resultados**

Neste subcapítulo são apresentadas as Tabelas 7 a 9, que correspondem aos três casos de estudo, sendo que foram testados dois conjuntos de treino e testes para cada caso. Na literatura incidente na temática, o tamanho do conjunto de teste varia entre 3% e 20% da amostra total, entre os diferentes autores. Com base nestes, decidiu-se o tamanho dos conjuntos de teste para os dois casos em 10% e 20%.

#### 4.2.1.1. Caso de estudo 1 (Cultivar 2, dia #1)

O primeiro caso de estudo refere-se às fotografias da Cultivar 2 recolhidas no primeiro dia. Os dados recolhidos destas fotografias correspondem a combinações de dados de área e métrica, isto é, à probabilidade do objeto ser redondo, como demonstrado na Tabela 5 da secção 3.3.1. Estes dados são inseridos numa estrutura que é aleatoriamente reestruturada antes de serem efetuados os testes. Desta forma, os 20 testes contemplados na Tabela 7 correspondem a 20 conjuntos de treino e teste, todos diferentes e totalmente aleatórios. Assim, é possível garantir que os resultados obtidos não são influenciados.

Tabela 7 - Resultados obtidos para os testes do primeiro caso: conjunto de teste 20% e 10% respectivamente lado esquerdo e lado Direito.

<i>Cultivar 2 - Dia #1</i>			<i>Cultivar 2 - Dia #1</i>		
<i>Amostra de Treino</i>	5722 - 80%		<i>Amostra de Treino</i>	6437 - 90%	
<i>Amostra de Teste</i>	1430 - 20%		<i>Amostra de Teste</i>	715 - 10%	
<i>Ensaio</i>	<i>Taxa de acerto</i>	%	<i>Ensaio</i>	<i>Taxa de acerto</i>	%
1	0,6000	60	1	0,6169	62
2	0,6100	61	2	0,6098	61
3	0,6017	60	3	0,6094	61
4	0,5997	60	4	<b>0,6660</b>	<b>67</b>
5	0,5963	60	5	0,6520	65
6	0,6424	64	6	0,6214	62
7	0,5988	60	7	0,6199	62
8	0,6116	61	8	0,6179	62
9	0,5973	60	9	0,6243	62
<b>10</b>	<b>0,6697</b>	<b>67</b>	10	0,6312	63
<i>Valor Médio</i>		61	<i>Valor Médio</i>		63

Como é possível observar na Tabela 7, a melhor taxa de acerto para ambos os casos é de 67%. Porém, verifica-se um aumento do valor médio para a taxa de acerto na segunda configuração de conjuntos. Este aumento, deve-se ao aumento de informação no conjunto de treino, garantindo assim mais informação de treino e um modelo SVM mais preciso do que o anterior.

De seguida são apresentadas as Figuras 17 e 18 correspondentes à análise da Tabela 7. A Figura 17 corresponde aos gráficos de dispersão para os conjuntos de treino e teste dos melhores ensaios de cada caso. A Figura 18 apresenta os gráficos de barras com os resultados obtidos em cada caso.

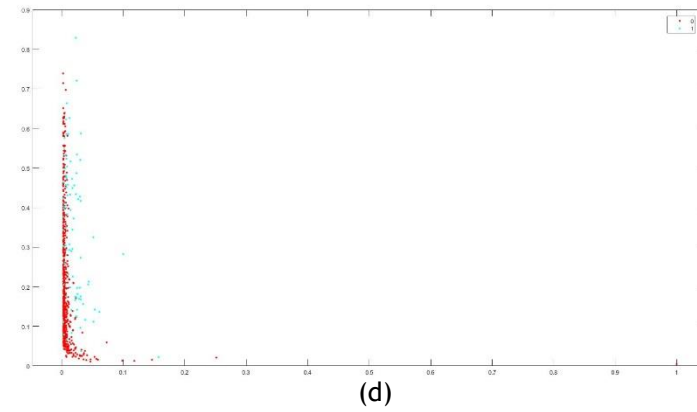
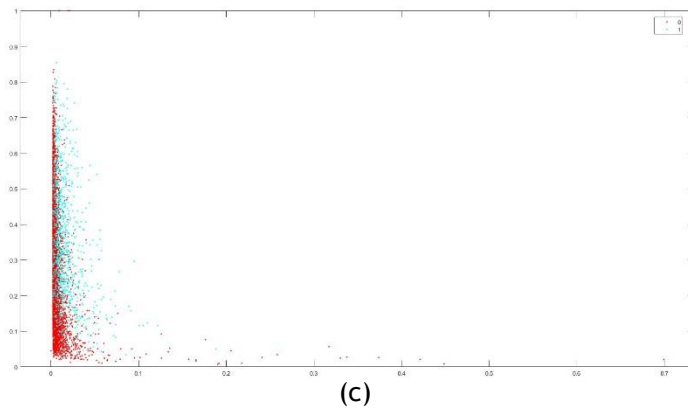
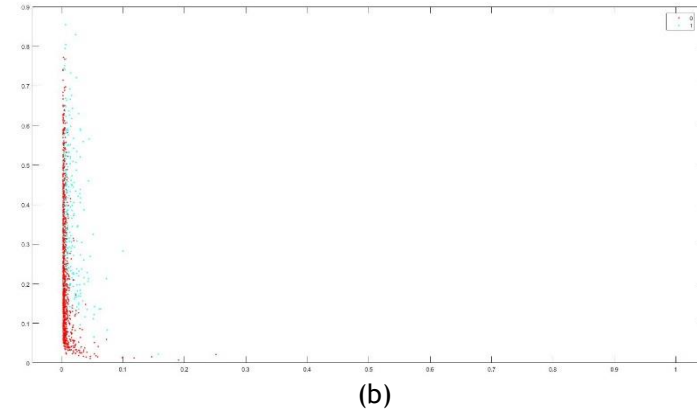
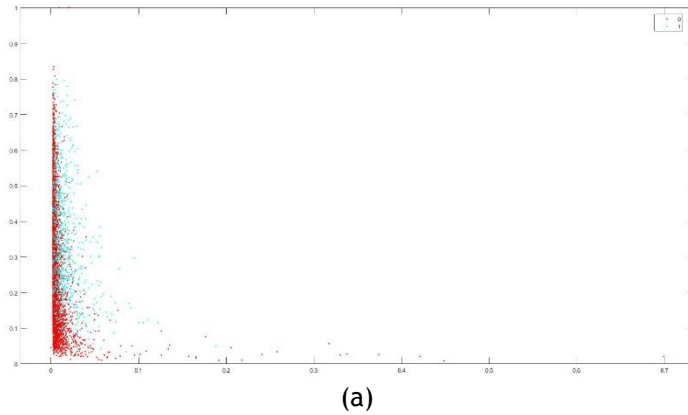
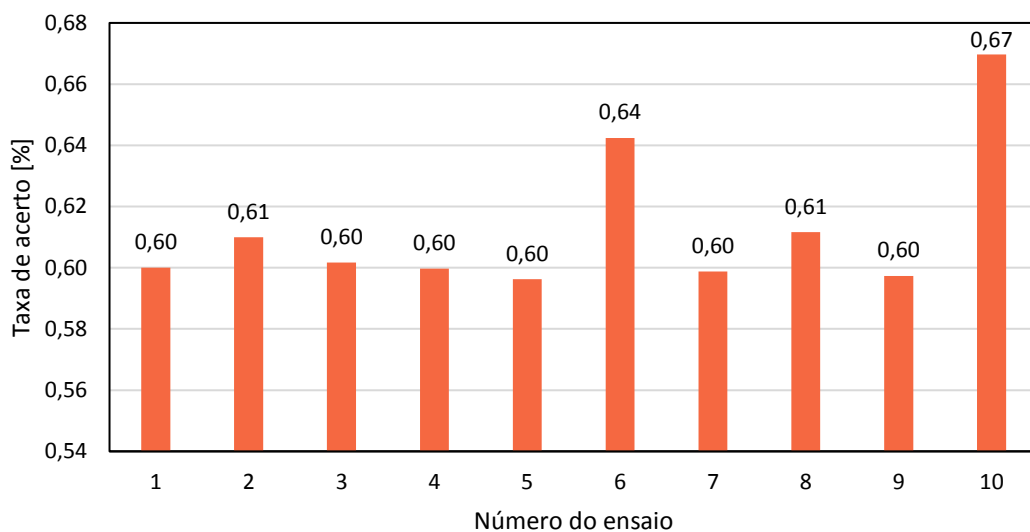
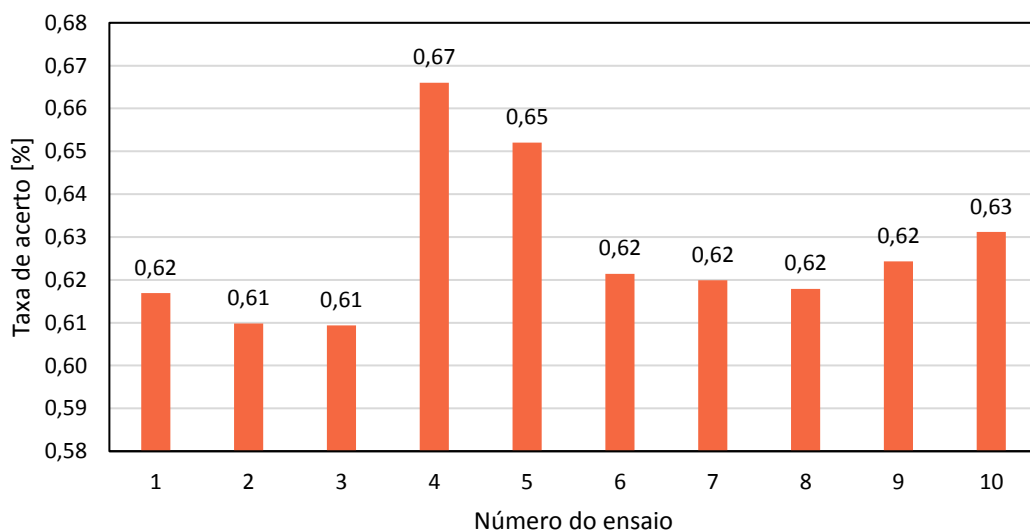


Figura 17 - Gráficos de dispersão para os melhores ensaios para ambos os casos: a) Gráfico de dispersão do conjunto de treino equivalente a 80% da amostra, b) Gráfico de dispersão do conjunto de teste equivalente a 20% da amostra, c) Gráfico de dispersão do conjunto de treino equivalente a 90% da amostra, d) Gráfico de dispersão do conjunto de teste equivalente a 10% da amostra.



(a) Cultivar 2 - dia #1 (Treino: 80%; Teste 20%).



(b) Cultivar 2 - dia #1 (Treino: 90%; Teste 10%).

Figura 18 - Representação através de gráfico de barras dos resultados obtidos na Tabela 7: a) Gráfico dos resultados obtidos para os 10 ensaios com a configuração de conjuntos 80%-20%, b) Gráfico dos resultados obtidos para os 10 ensaios com a configuração de conjuntos 90%-10%.

Como se pode observar nos gráficos apresentados na Figura 18, os coeficientes de correlação variam entre 60% e 67% para o caso com a configuração de conjuntos 80%-20%, permitindo assim obter um valor médio de 61% e entre 61% e 67% para o

caso com a configuração de conjuntos 90%-10%, obtendo desta forma um valor médio de 63%. Estes são os únicos casos de estudo com fotografias tiradas no dia #1, 5 de julho de 2017. Nestes casos de estudo, os frutos ainda não tinham atingido a época de maturação, encontrando-se assim numa maior variedade de cor e tamanhos.

#### 4.2.1.2. Caso de estudo 2 (Cultivar 1, dia #2)

O segundo caso de estudo refere-se às fotografias da Cultivar 1 recolhidas no segundo dia. Neste caso, como já foi referido anteriormente, as fotografias antes de serem segmentadas, foram sujeitas a uma correção da luminosidade, devido às baixas condições de luminosidade encontradas aquando da recolha das imagens. O processo referente à escolha aleatória dos conjuntos de treino e teste foi repetido. A Tabela 8 apresenta os resultados obtidos para esta amostra.

Tabela 8 - Resultados obtidos para os testes do segundo caso: conjunto de teste 20% e 10% respectivamente lado esquerdo e lado Direito.

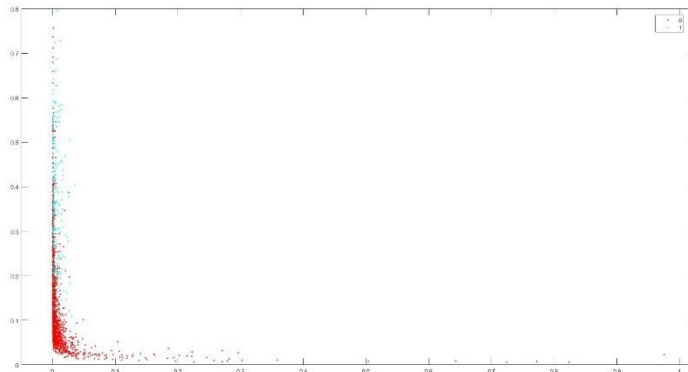
<i>Cultivar 1 - Dia #2</i>			<i>Cultivar 1 - Dia #2</i>		
<i>Amostra de Treino</i>	2300 - 80%		<i>Amostra de Treino</i>	2587 - 90%	
<i>Amostra de Teste</i>	575 - 20%		<i>Amostra de Teste</i>	288 - 10%	
<i>Ensaio</i>	<i>Taxa de acerto</i>	<i>%</i>	<i>Ensaio</i>	<i>Taxa de acerto</i>	<i>%</i>
1	0,5974	60	1	0,6307	63
2	<b>0,6372</b>	<b>64</b>	2	0,6301	63
3	0,6249	62	3	0,7078	71
4	0,6279	63	4	0,6213	62
5	0,6108	61	5	0,6108	61
6	0,6001	60	<b>6</b>	<b>0,7219</b>	<b>72</b>
7	0,6131	61	7	0,6598	66
8	0,6232	62	8	0,6727	67
9	0,6126	61	9	0,6358	64
10	0,6298	63	10	0,6337	63
<i>Valor Médio</i>		62	<i>Valor Médio</i>		65

Como é possível observar na Tabela 8, existe uma diferença considerável na melhor taxa de acerto de ambos os casos. Na primeira configuração de conjuntos, o melhor

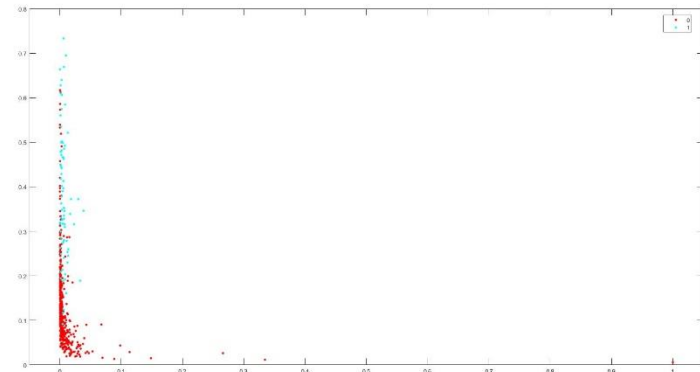
ensaio apresentou uma taxa de acerto de 64%, enquanto que para a segunda configuração de conjuntos, o melhor ensaio apresentou uma taxa de acerto de 72%. Novamente, os melhores valores, são obtidos no segundo caso de estudo quando os conjuntos de treino são maiores, ou seja, o modelo SVM criado no segundo caso de estudo contem mais informação, concretizando assim uma classificação mais precisa. Os valores médios para a taxa de acerto, neste caso, situam-se nos 62% e 65%, respetivamente, para o primeiro e segundo casos de estudo.

De seguida são apresentadas as Figuras 19 e 20 correspondentes à análise da Tabela 8. A Figura 19 corresponde aos gráficos de dispersão para os conjuntos de treino e teste do melhor ensaio de cada caso. A Figura 20 apresenta os gráficos de barras com os resultados obtidos em cada caso.

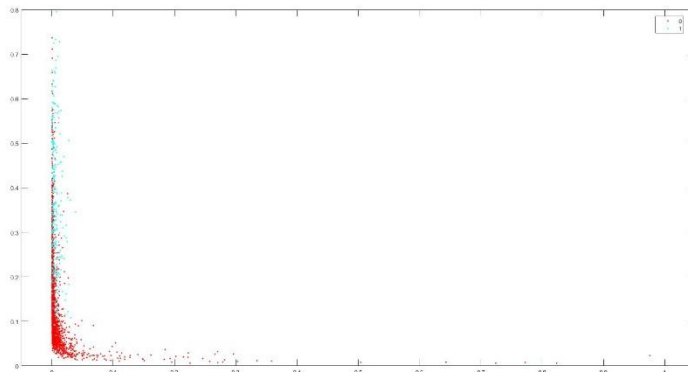




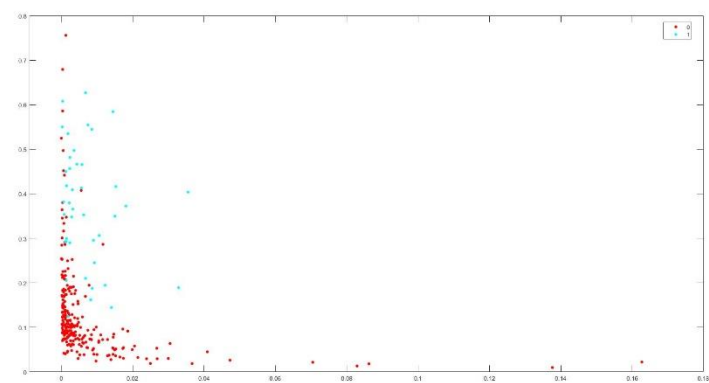
(a)



(b)

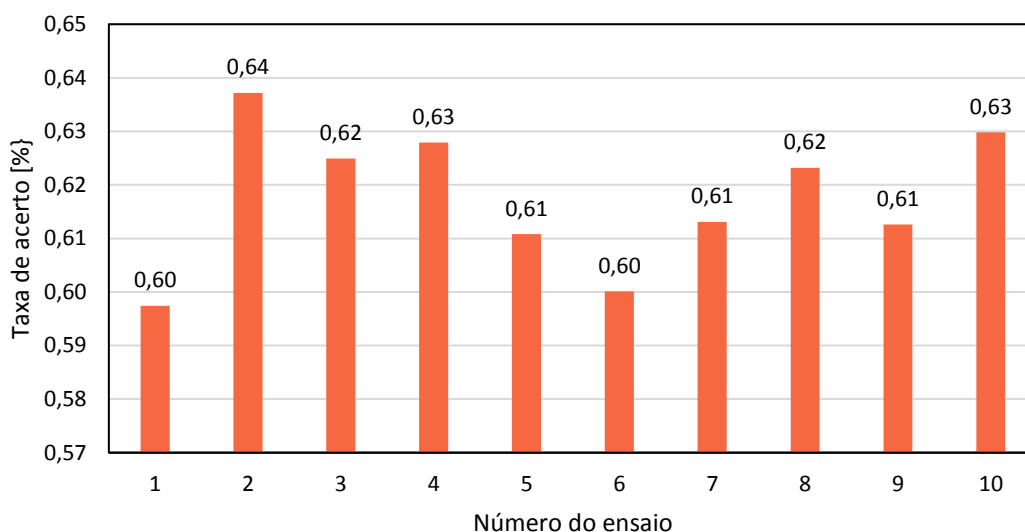


(c)

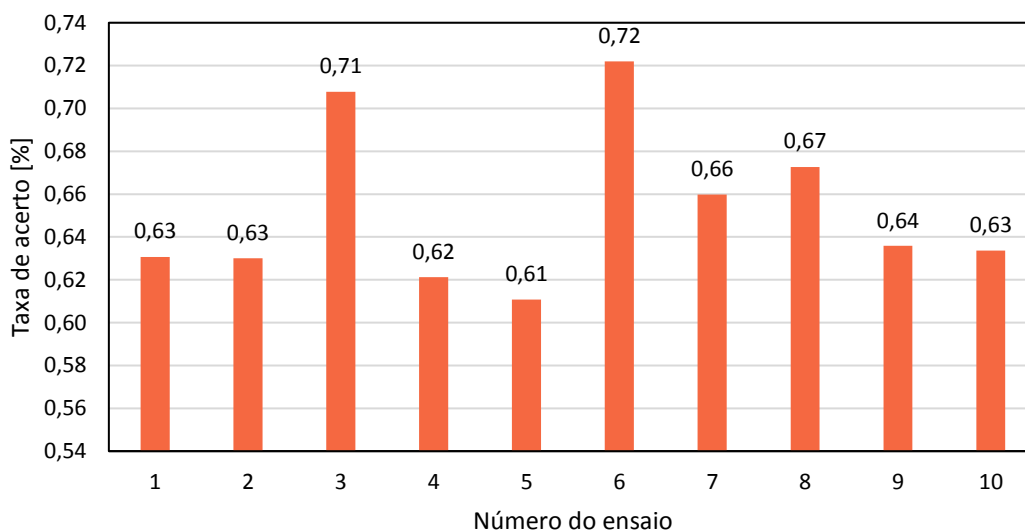


(d)

Figura 19 - Gráficos de dispersão para os melhores ensaios para ambos os casos: a) Gráfico de dispersão do conjunto de treino equivalente a 80% da amostra, b) Gráfico de dispersão do conjunto de teste equivalente a 20% da amostra, c) Gráfico de dispersão do conjunto de treino equivalente a 90% da amostra, d) Gráfico de dispersão do conjunto de teste equivalente a 10% da amostra.



(a) Cultivar 1 - dia #2 (Treino: 80%; Teste 20%).



(b) Cultivar 1 - dia #2 (Treino: 90%; Teste 10%).

Figura 20 - Representação através de gráfico de barras dos resultados obtidos na Tabela 8: a) Gráfico dos resultados obtidos para os 10 ensaios com a configuração de conjuntos 80%-20%, b) Gráfico dos resultados obtidos para os 10 ensaios com a configuração de conjuntos 90%-10%.

Como se pode observar nos gráficos apresentados na Figura 20, os coeficientes de correlação variam entre 60% e 64% para o caso com a configuração de conjuntos 80%-20%, permitindo assim obter um valor médio de 62%. Para a configuração de

conjuntos 90%-10%, obtiveram-se valores dos coeficientes de correlação entre 61% e 72%, obtendo desta forma um valor médio de 65%.

Estes resultados foram obtidos a partir das fotos da cultivar 1 recolhidas no dia #2, 20 de julho de 2017. Para esta cultivar não são realizados cálculos de previsão no sub-capítulo 4.3, pois alguns dos pêssegos já tinham sido colhidos previamente à recolha das fotos.

#### 4.2.1.3. Caso de estudo 3 (Cultivar 2, dia #2)

O terceiro caso de estudo refere-se às fotografias da Cultivar 2 recolhidas no segundo dia. Tal como no segundo caso de estudo, estas fotografias fazem parte do grupo que sofreu a correção de luminosidade. Mais uma vez, os conjuntos de treino para todos os testes foram escolhidos com o mesmo método sendo totalmente aleatórios. A Tabela 9 apresenta os resultados para esta amostra.

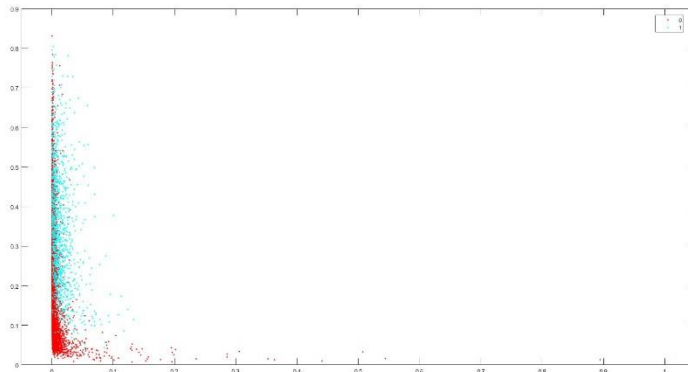
Tabela 9 - Resultados obtidos para os testes do terceiro caso: conjunto de teste 20% e 10% respectivamente lado esquerdo e lado Direito.

<i>Cultivar 2 - Dia #2</i>			<i>Cultivar 2 - Dia #2</i>		
<i>Amostra de Treino</i>	6630 - 80%		<i>Amostra de Treino</i>	7459 - 90%	
<i>Amostra de Teste</i>	1657 - 20%		<i>Amostra de Teste</i>	829 - 10%	
<i>Ensaio</i>	<i>Taxa de acerto</i>	<i>%</i>	<i>Ensaio</i>	<i>Taxa de acerto</i>	<i>%</i>
1	0,6639	66	1	0,6621	66
2	0,6556	66	2	0,6819	68
3	0,6416	64	3	0,6645	66
4	0,6409	64	4	<b>0,7127</b>	<b>71</b>
5	0,6462	65	5	0,6760	68
6	0,6404	64	6	0,6942	69
7	0,6579	66	7	0,6654	67
8	0,6629	66	8	0,6795	68
9	<b>0,6678</b>	<b>67</b>	9	0,6607	66
10	0,6441	64	10	0,6790	68
<i>Valor Médio</i>		65	<i>Valor Médio</i>		68

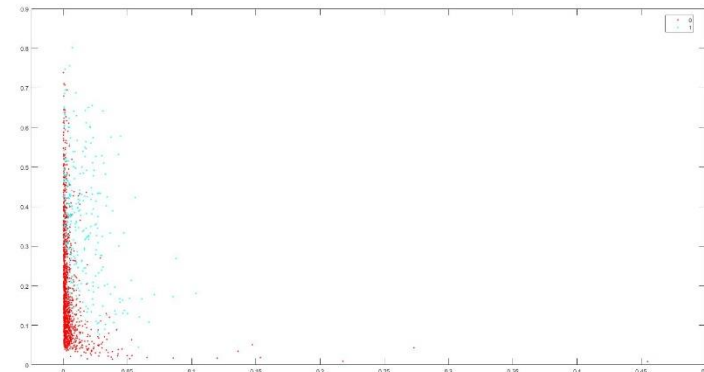
A Tabela 9 apresenta os resultados referentes ao terceiro e último caso de estudo. Ambas as configurações de conjuntos apresentam resultados com valores para a taxa de acerto maioritariamente mais elevados do que os casos de estudo

anteriores. Novamente, é possível verificar que a segunda configuração de conjuntos, coincidente com um maior número de elementos para o conjunto de treino, apresenta um ensaio com taxa de acerto acima dos 70%. Em suma, é possível afirmar, que ao longo dos testes para os diferentes casos de estudo, um padrão foi observado, mostrando que o aumento das taxas de acerto está diretamente relacionado com o aumento dos conjuntos de treino, permitindo à SVM classificações mais precisas dos dados inseridos para teste.

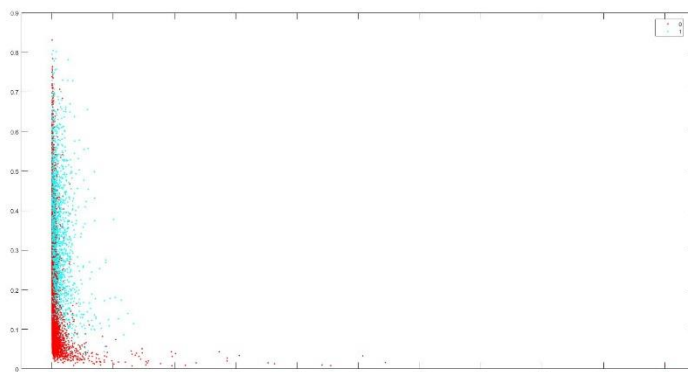
De seguida são apresentadas as Figuras 21 e 22 correspondentes à análise da Tabela 9. A Figura 21 corresponde aos gráficos de dispersão para os conjuntos de treino e teste do melhor ensaio de cada caso. A Figura 22 apresenta os gráficos de barras com os resultados obtidos em cada caso.



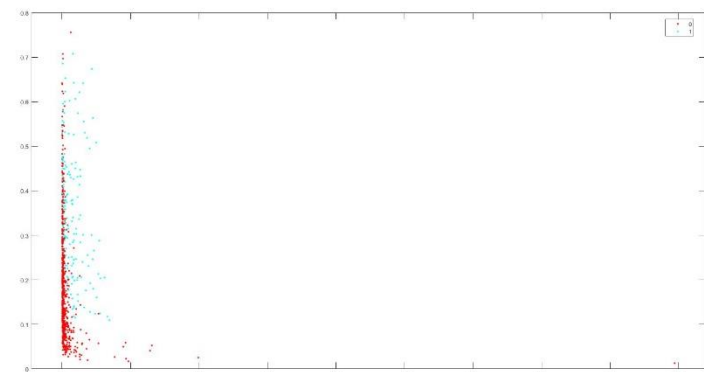
(a)



(b)

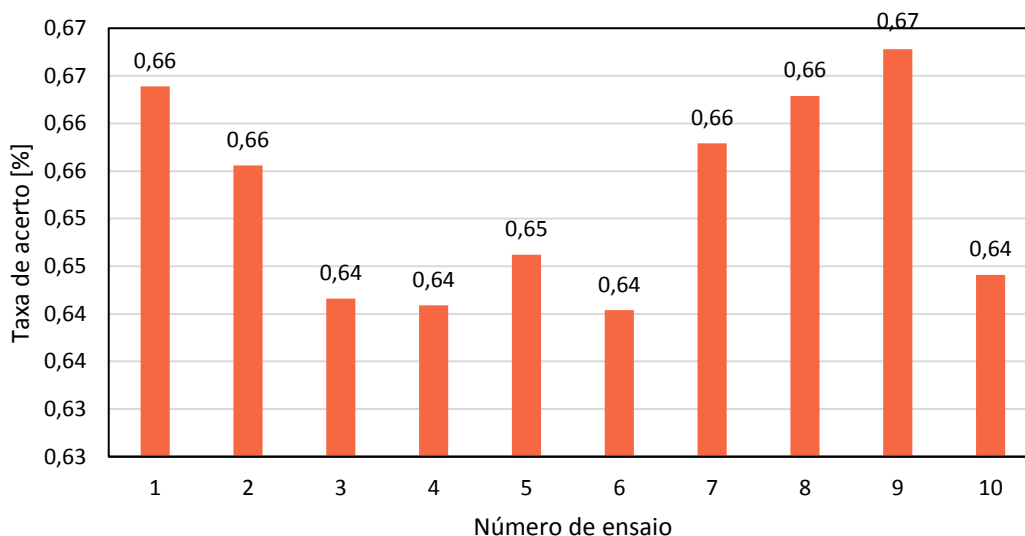


(c)

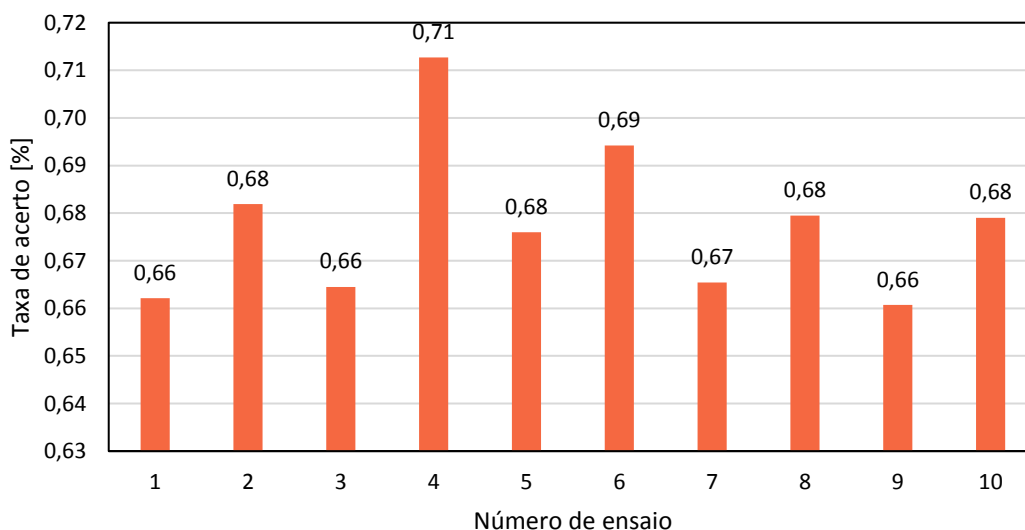


(d)

Figura 21 - Gráficos de dispersão para os melhores ensaios para ambos os casos: a) Gráfico de dispersão do conjunto de treino equivalente a 80% da amostra, b) Gráfico de dispersão do conjunto de teste equivalente a 20% da amostra, c) Gráfico de dispersão do conjunto de treino equivalente a 90% da amostra, d) Gráfico de dispersão do conjunto de teste equivalente a 10% da amostra.



(a) Cultivar 2 - dia #2 (Treino: 80%; Teste 20%).



(b) Cultivar 2 - dia #2 (Treino: 90%; Teste 10%).

Figura 22 - Representação através de gráfico de barras dos resultados obtidos na Tabela 9: a) Gráfico dos resultados obtidos para os 10 ensaios com a configuração de conjuntos 80%-20%, b) Gráfico dos resultados obtidos para os 10 ensaios com a configuração de conjuntos 90%-10%.

Como se pode observar nos gráficos apresentados na Figura 22, os coeficientes de correlação variam entre 64% e 67% para o caso com a configuração de conjuntos 80%-20%, permitindo assim obter um valor médio de 65%. O caso com a configuração de conjuntos 90%-10%, apresenta valores dos coeficientes de correlação a variar entre 66% e 71%, obtendo desta forma um valor médio de 68%.

Esta cultivar foi a que apresentou coeficientes de correlação com valores médios mais elevados para as fotos recolhidas no dia 20 de julho de 2017. No dia em questão, os pêssegos já tinham atingido a maturação, apresentando assim características de cor, tamanho e forma mais uniformes. No subcapítulo 4.3 são apresentados os cálculos da previsão de produção para esta cultivar.

### 4.3. Previsão de produção

A cultivar escolhida para apresentar a previsão de produção com o algoritmo desenvolvido foi a cultivar 2, *Sweet Dream*. Esta cultivar apresenta características muito específicas, principalmente a nível de tamanho. Estas características têm uma relação peso-diâmetro que pode chegar a 400 g e 9 cm respetivamente (Simões, 2016). Para o cálculo da previsão de produção foram consideradas as seguintes características:

- Peso médio 400 g;
- Diâmetro médio 9 cm;
- Massa específica de 1,05 g/cm<sup>3</sup>

O peso e o diâmetro escolhidos foram baseados na relação peso-diâmetro apresentada por Simões (2016). Foi escolhido o limite superior de valores para peso e diâmetro, tendo como perspetiva ajudar a balancear a obstrução proporcionada pelo excesso de folhagem. As estimativas da previsão da produção são apresentadas em toneladas por hectare [ton/ha].

#### 4.3.1. Previsão de Produção - Algoritmo

Para obter um valor de previsão de produção para o campo com a Cultivar 2, foram necessários ter em conta certos dados sobre o terreno. A cultivar tem um compasso de 12,375 m<sup>2</sup> e 100 plantas. Sabendo que o valor do compasso corresponde à relação entre o espaço das entre linhas e o espaçamento entre plantas, conclui-se que o terreno tem uma área de 1237,5 m<sup>2</sup>.

#### 4.3.1.1. Compasso

Como referido anteriormente, o compasso corresponde à relação entre o espaço das entrelinhas do pomar e o espaçamento entre as plantas em cada linha do pomar. Assim, é possível calcular a área que ocupa cada planta.

$$C = E_L \cdot E_P \Leftrightarrow C = 4,5 \cdot 2,75 \Leftrightarrow C = 12,375 \text{ m}^2$$

Em que:

- $C$  - Compasso [ $\text{m}^2$ ];
- $E_L$  - Espaço entre linhas [m];
- $E_P$  - Espaçamento entre plantas [m].

#### 4.3.1.2. Cálculo da massa específica

O valor da densidade obtido foi  $1,05 \text{ g/cm}^3$  para pêsegos com diâmetro de 9 cm e peso de 400 g. De seguida é apresentado o cálculo efetuado para obter este valor.

$$V = \frac{4}{3} \cdot \pi \cdot r^3 \Leftrightarrow V = \frac{4}{3} \cdot \pi \cdot 4,5^3 \Leftrightarrow V = 381,7 \text{ cm}^3$$

$$\rho = \frac{P}{V} \Leftrightarrow \rho = \frac{400}{381,7} \Leftrightarrow \rho = 1,05 \text{ g/cm}^3$$

Em que:

- $V$  - Volume [ $\text{cm}^3$ ];
- $r$  - Raio [cm];
- $\rho$  - Massa específica [ $\text{g/cm}^3$ ];
- $P$  - Peso [g];

#### 4.3.1.3. Estimativa da Previsão de Produção

A estimativa da produção foi realizada com base nos dados obtidos no ensaio número 4, da Tabela 9, para a configuração de conjuntos 90%-10%. O coeficiente de correlação obtido para este ensaio situou-se em 71%. Sabendo que o intervalo de valores de produção pode variar entre 20 [t/ha] e 58 [t/ha] (Simões & Carvalho,



2008), este será considerado o intervalo de referência para as estimativas a calcular.

- Valor do peso para 10% da amostra dado pela SVM: 296 980 g;

A estimativa para 100% da amostra vem.

$$\begin{array}{rcl} 296\,980\text{ g} & \rightarrow & 10\% \\ x & \rightarrow & 100\% \end{array}$$

$$x = \frac{296\,980 \cdot 100}{10} \Leftrightarrow x = 296\,980 \cdot 10 \Leftrightarrow x = 2\,969\,800\text{ g} \Leftrightarrow x \approx 2,97\text{ ton}$$

Como as fotografias recolhidas não correspondiam à totalidade do campo da cultivar 2, foi necessário calcular a proporção de terreno fotografado para mais tarde fazer a estimativa da área total. Sabendo que a primeira linha do pomar tem 11 plantas e apenas metade da linha foi fotografada, e que a última linha do pomar tem 20 plantas e, novamente, apenas metade foi fotografada, concluímos que 18 das 100 plantas da cultivar não foram fotografadas, ou seja, apenas 82 plantas foram fotografadas. Assim, com o cálculo do compasso apresentado anteriormente é possível calcular a área de terreno que foi fotografada.

$$A_f = 82 \cdot 12,375 \Leftrightarrow A_f = 1014,75\text{ m}^2$$

Em que:

- $A_f$  - Área fotografada [ $\text{m}^2$ ].

Por fim, e com todos os dados necessários calculados, é possível fazer uma estimativa da previsão da produção para um hectare.

$$\begin{array}{rcl} 2,97\text{ ton} & \rightarrow & 1\,014,75\text{ m}^2 \\ x & \rightarrow & 10\,000\text{ m}^2 \end{array}$$

$$x = \frac{2,97 \cdot 10\,000}{1\,014,75} \Leftrightarrow x = 29,26\text{ [t/ha]} \Leftrightarrow x \approx 29,3\text{ [t/ha]}$$

Concluindo assim que a previsão obtida pela máquina é de 29,3 [t/ha], que coincide com o intervalo de valores apresentado anteriormente.

#### 4.3.2. Previsão de Produção - Teórica

Para a previsão teórica foi feita a contagem manual dos pêsegos na fotografia apresentada na Figura 23.



Figura 23 - Contagem manual de pêsegos.

Sendo que os pêsegos contabilizados correspondem apenas a um dos lados da planta, o número de pêsegos contabilizados, 62 pêsegos, foi multiplicado por dois para obter um numero mais próximo do número real de frutos produzidos por cada planta. Considerando o peso de 400 g por fruto, referido no início do capítulo 4.3, foi calculado o peso produzido em cada planta e multiplicado pelo número de plantas num hectare.

$$P_{planta} = 124 \cdot 400 \Leftrightarrow P_{planta} = 49\ 600\text{ g} \Leftrightarrow P_{planta} = 49,6\text{ kg}$$

$$NA_h = \frac{100 \cdot 10\ 000}{1\ 237,5} \Leftrightarrow NA_h = 808\text{ plantas}$$

$$P_{total} = 49,6 \cdot 808 \Leftrightarrow P_{total} = 40\ 076\text{ kg/ha} \Leftrightarrow P_{total} = 40,1\text{ ton/ha}$$

Em que:

- $P_{\text{planta}}$  - Peso médio de cada planta [kg];
- $NA_h$  - Número de plantas por hectare;
- $P_{\text{total}}$  - Peso estimado para 808 plantas, ou seja, 1 hectare [ton/ha].

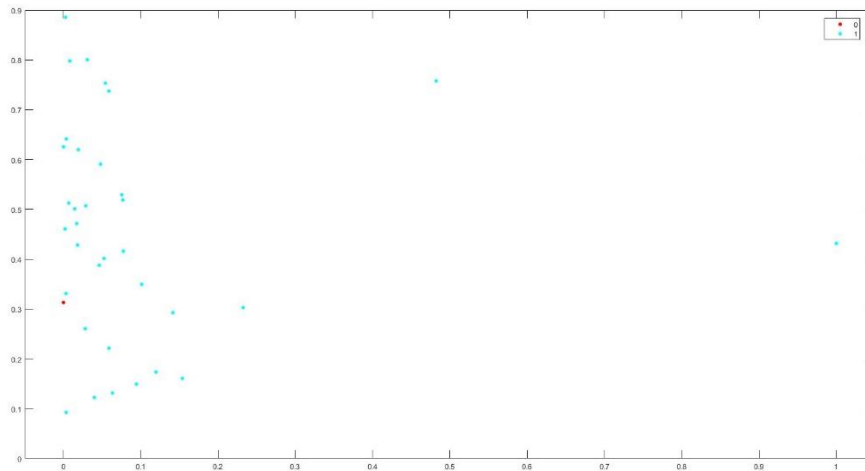
#### 4.4. Discussão dos resultados

Ao longo do Capítulo 4 foram apresentados os resultados do reconhecimento de pêsegos obtidos com a máquina de suporte vetorial apresentada no Capítulo 3. Estes resultados podem ser analisados pelos gráficos das Figuras 18, 20 e 22 que correspondem aos dados das Tabelas 7 a 9, respectivamente. Estes dados correspondem aos coeficientes de correlação entre os resultados dados pela SVM e as classificações desses mesmos dados inseridos manualmente. Como se pode observar no subcapítulo 2.7 e 2.7.1, os resultados obtidos pelos trabalhos científicos expostos na literatura são consideravelmente melhores que os obtidos no presente estudo. No entanto, e como já foi referido anteriormente, as características dos trabalhos apresentados nesses subcapítulos são diferentes das expostas no presente estudo. Em ambos os casos expostos pelos diferentes autores, a segmentação das imagens ou é aplicada de forma a procurar características muito específicas, ou as frutas são colocadas em superfícies planas de cor sólida de modo a não conter informação desnecessária, ou ainda são fotografias de pequena amplitude de modo muito mais localizado, procurando frutas não obstruídas por folhas ou ramos. No caso em estudo, as condições são totalmente diferentes. O estudo apresentado propõe uma máquina de suporte vetorial para localização de frutos em árvore. Frutos em árvore implicam condições naturais. Todas as fotografias recolhidas foram captadas sob contrastes e luminosidades naturais, sem foco numa parte específica da planta, mas sim, tentando captar o maior número de frutos possíveis, com o objetivo de posteriormente usar a SVM para identificar todos os pêsegos possíveis. Como resultado, o intervalo que a SVM considera ou não pêsego é muito mais alargado, fazendo com que haja uma ligeira redução na taxa de reconhecimento, que varia dos 60% aos 72% entre os diferentes casos de estudo. Contudo, e considerando todas as condições aqui expostas, os resultados estão dentro do esperado para uma primeira abordagem ao problema em questão.

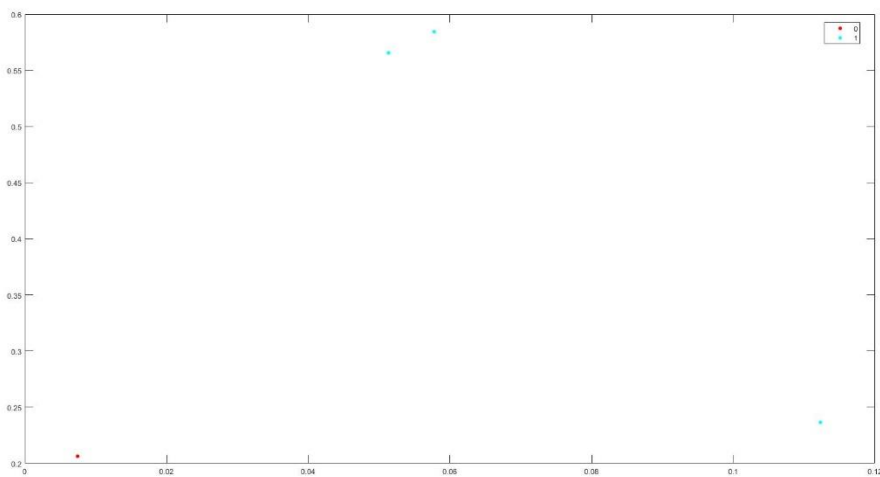
Ainda ao longo do Capítulo 4, foram apresentadas as estimativas prática e teórica para a previsão de produção da cultivar 2, *Sweet Dream*. O valor da estimativa obtido pela máquina de suporte vetorial foi de 29,3 [t/ha]. Tendo em consideração o intervalo de valores de produção, entre 20 e 58 [t/ha], a estimativa obtida pelo algoritmo está correta. Os valores de produção anuais para uma cultivar de pessegueiro, são valores que variam consideravelmente, pois é um valor que depende de muitos fatores da área agrícola, desde a disponibilidade de sol, de água, ataques de pragas, etc. Para a previsão obtida, é possível afirmar que foi um ano com uma produção média aceitável. O valor da estimativa obtido através do cálculo teórico foi de 40,1 [t/ha]. Mais uma vez, verifica-se que o valor estimado coincide com o intervalo de valores de produção. Neste caso, o cálculo foi realizado contemplando o número de plantas e calculando o valor médio do peso dos frutos produzidos em cada planta. Conclui-se assim, que os valores obtidos para as estimativas, estão em conformidade com os valores esperados.

#### **4.5. Nota Conclusiva**

Uma pequena simulação foi feita para comprovar que as características escolhidas para treinar a máquina de suporte vetorial são válidas. Esta simulação é composta por uma amostra de trinta e nove elementos, recolhidos de dez figuras semelhantes à figura utilizada ao longo do Capítulo 3, ou seja, com um segundo plano que contém menos informação, sem oclusão de frutos ou sombras e frutos com cores mais uniformes. Os conjuntos de treino e teste, foram novamente, escolhidos de forma aleatória com a configuração de 90% treino e 10% teste. Como seria de esperar, os resultados foram em tudo semelhantes aos da literatura, obtendo coeficientes de correlação que variam entre 85% e 100%. As Figuras 23 e 24 apresentam os gráficos de dispersão para um ensaio com um coeficiente de correlação de 100%, e exemplos das imagens usadas para a recolha da informação, respetivamente.



(a) Dispersão do conjunto de treino.



(b) Dispersão do conjunto de teste.

Figura 24 - Graficos de dispersão do ensaio de validação das características de treino da SVM: a) conjunto de treino, b) conjunto de teste.





Figura 25 -Exemplo de figura usada para extração de informação do ensaio de validação.

Apesar do cenário encontrado para a recolha das fotografias ter apresentado vários desafios, proporcionados pelas condições adversas, os resultados obtidos mostram que é possível fazer reconhecimento aproximadamente total dos frutos em árvore, de uma determinada cultivar, sujeitos apenas às condições naturais.

## 5. Conclusões

### 5.1. Conclusões

O objetivo desta dissertação visa a criação de um algoritmo de reconhecimento de pêssegos em ambiente natural, para previsão de produção de uma cultivar. O algoritmo tem como objetivo distinguir os frutos do resto dos elementos da fotografia, de forma a calcular as suas dimensões e respetivo peso. Futuramente, propõe-se a aplicação do presente algoritmo num sistema robótico aéreo autónomo (drone), visando a criação de um sistema totalmente autónomo na recolha e análise dos dados.

No presente estudo, as fotografias recolhidas em ambiente natural de árvores de fruto, são sujeitas a vários processos de segmentação. Numa primeira instância é feito um pré processamento, delimitado o espaço de cores que se pretende reconhecer, aplicando operações matemáticas morfológicas e fazendo reconhecimento de arestas. Após a extração de todas as características necessárias, um método de classificação através de uma máquina de suporte vetorial é aplicado para reconhecimento de frutos (pêssegos) em árvore. Os resultados obtidos com recurso à SVM são considerados válidos, com coeficiente de correlação de 71%, considerando todas as adversidades encontradas e explicadas, como características de luminosidade adversas, excesso de folhagem, entre outros. No entanto, com a elevada quantidade de informação recolhida, de certo modo, essas adversidades foram colmatadas até um certo ponto. Estas mesmas adversidades encontradas tiveram um papel fundamental nos valores encontrados para a estimativa de previsão de produção. A discrepância encontrada entre a estimativa teórica e prática desses valores foi influenciada pela estratégia escolhida para estimar a previsão. O cálculo da previsão é feito calculando o volume da esfera das formas aproximadamente redondas encontradas pela SVM, que na realidade podem estar ocultas por folhas e ramos. Adicionalmente, a câmara usada não possui abertura de

lente suficiente para fotografar uma planta na sua totalidade no espaço disponível, pelo que não só as fotos recolhidas não contêm todos os pêseços que eram visíveis, como os pêseços que aparecem, em grande parte, não aparecem na sua totalidade. Assim, e apesar das adversidades, a estimativa prática, de 29,3 [t/ha], encontra-se dentro do intervalo de valores entre o qual pode variar a produção anual dum hectare de pessegueiros. Desta forma, é possível concluir que o presente estudo atingiu os objetivos preconizados, confirmando mais uma vez que as máquinas de suporte vetorial são métodos de classificação muito viáveis, apresentando bons resultados mesmo em situações com condições bastante adversas.

### 5.2. Sugestões de trabalhos futuros

Os resultados obtidos neste estudo são promissores, fazendo com que trabalhos futuros na área deixem quase de ser sugestões, mas sim uma necessidade, começando logo pela recolha das fotografias. Para uma validação e previsão da produção mais precisa, será necessário recorrer a câmaras mais avançadas, marcar árvores para contagem de frutos, definir distâncias e realizar medições no terreno de modo a obter locais específicos para adquirir fotografias com o mesmo ângulo e equidistantes às árvores.

Para melhorar o reconhecimento da máquina de suporte vetorial, algoritmos de segmentação mais complexos podem ser aplicados, como por exemplo, algoritmos que contemplam o fator de crescimento dos pêseços, algoritmos que analisem outras características dos pêseços, além de cor, tamanho e forma, sendo assim possível obter novas características para classificação. Outra possível melhoria na classificação, pode ser obtida através da criação de uma “SVM adaptativa”, considerando diferentes características de classificação consoante a altura em que a recolha das fotografias é feita. Desta forma, seria possível ter em conta outros fatores, como a percentagem de queda dos frutos, avaliar a taxa de crescimento dos pêseços entre as diferentes alturas da recolha das fotografias e aplicar algoritmos de controlo de qualidade para identificação de doenças ou outros fatores, que prejudiquem o que crescimento do fruto. Todos estes fatores permitiriam novas características de treino, tornando o reconhecimento mais preciso.



Para a estimativa da previsão de produção, uma estratégia diferente pode ser adotada. O presente estudo usa volume e massa específica para prever o peso. Contudo, com um pouco de trabalho de campo, uma redução da folhagem das plantas, permitiria obter melhores fotografias e chegar a um número mais aproximado do número real de pêssegos por árvore, permitindo atingir uma estimativa mais precisa da produção anual de uma cultivar específica. Neste caso seria possível considerar também localização geoespacial. A redução na folhagem faria com que fotos em ambos os lados da planta causassem repetição do mesmo fruto. Com a utilização de localização geoespacial, essa a repetição de frutos deixaria de ocorrer.



## Referências Bibliográficas

Awcock, G, & Thomas R (1996). Applied Image Processing. McGraw-Hill Inc., New York

Bac, C.W., Hemming, J., & Van Henten, E.J. (2014). Stem localization of sweet-pepper plants using the support wire as a visual cue. *Computers and Electronics in Agriculture*, 105, 111-120.

Bulanon, D., Kataoka, T., Ota, Y., & Hiroma, T. (2002). AE—automation and emerging technologies: a segmentation algorithm for the automatic recognition of Fuji apples at harvest. *Biosystems Engineering*, 83(4), 405-412

Chakraborty, S., Roy, M., & Hore, S. (2017). A Study on Different Edge Detection Techniques in Digital Image Processing. In N. Dey, A. Ashour, e P. Patra (Eds.), Feature Detectors and Motion Detection in Video Processing (pp. 100-122).

Coelho, J.P.C, & Silva, J. R. M. (2009). “Agricultura de Precisão”, em *Agricultura de Precisão*, 1ª ed. Lisboa, Portugal: AJAP, 2009, pp. 2-9.

Hsu C.-W., Chang, C.-C., & Lin, C.-J. (2016). A Practical Guide to Support Vector Classification.

Hernández-Hernández, J.L., García-Mateos, G., González-Esquivá, J.M., Escarabajal-Henarejos, D., Ruiz-Canales, A., & Molina-Martínez, J.M. (2016). Optimal color space selection method for plant/soil segmentation in agriculture. *Computers and Electronics in Agriculture*, 122, 124-132.

Jiménez, A.R., Jain, A.K., Ceres, R., & Pons, J.L. (1999). Automatic fruit recognition: a survey and new results using Range/Attenuation images. *Pattern Recognition*, 32, 1719-1736.

Julesz B., & Caelli, T. (1979). On the limits of fourier decompositions in visual texture perception. *Perception*, 8, 69-73.

Zhao, J., Tow, J., & Katupitiya, J. (2005). On-tree Fruit Recognition Using Texture Properties and Color Data. In 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems (pp. 3993-3998).

Leemans, V., Magein, H., & Destain, M.-F. (2002). On-line Fruit Grading according to their External Quality using Machine Vision. *Biosystems Engineering*, 83, 397-404.

Li, X., & Wang, Y. Y. (2012). Image Quality Assessment Method Based on Support Vector Machine and Particle Swarm Optimization. *Advances in Computer Science and Information Engineering*, 353-359.

Materka, A., & Strzelecki, M. (1998). Texture Analysis Methods - A Review. Technical University of Lodz, Institute of Electronics, COST B11 report, Brussels.

Morimoto, T., Takeuchi, T., Miyata, H., & Hashimoto, Y. (2000). Pattern recognition of fruit shape based on the concept of chaos and neural networks. *Computers and Electronics in Agriculture*, 26, 171-186.

Nadernejad, E., Sharifzadeh, S., & Hassanpour, H. (2008). Edge Detection Techniques: Evaluations and Comparisons. *Applied Mathematical Sciences*, 31(2), 1507-1520.

Ohta, Y., Kanade, T., & Sakai, T. (1980). Color information for region segmentation. *Comput. Graphics Image Process*, 13(3), 222-241.

Corke, P. (2011). "Computer Vision", *Robotics, Vision and Control*, 3th ed. Berlin, Germany: Springer.

Owens, R., "Lecture 6", Computer Vision IT412, 10/29/1997. [http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL\\_COPIES/OWENS/LECT6/node2.html](http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/OWENS/LECT6/node2.html)

Schapiro, R. E. (2003). Nonlinear Estimation and Classification. *Lecture Notes in Statistics*, 171, 149-171.

Schölkopf B., & Cristianini N. (2002). Support Vector Machines and Kernel Methods. The New Generation of Learning Machines. *Ai Magazine*, 23(3), 31-41.

Simões, M.P. (2016). + *pêssego Guia Prático da Produção*. Covilhã. Centro Operativo e Tecnológico Hortofruticultura Nacional (COTHN).

Simões, M.P. e Carvalho, M.L. 2008. Relatório final do projecto Agro 452. Escola Superior Agrária de Castelo Branco - Instituto Politécnico de Castelo Branco, Castelo Branco, 28 pp (Anexos, 183 pp).

Song, Y., Glasbey, C., Horgan, G., Polder, G., Dieleman, J., & van der Heijden, G. (2014). Automatic fruit recognition and counting from multiple images. *Biosystems Engineering*, 118, 203-215.

Voorhees, H., & Poggio, T. Detecting textures and texture boundaries in natural images. *ICCV 87*:250-25,198.

W., Zhao, D., Cheng, F., Xu, B., Zhang, Y., & Wang, J. (2012). Automatic recognition vision system guided for apple harvesting robot. *Computers & Electrical Engineering*, 38(5), 1186-1195.

Maldonado Jr, W., & Barbosa, J. C. (2016). Automatic green fruit counting in Orange trees using digital images. *Computers and Electronics in Agriculture*, 127, 572-581

Wei, X., Jia, K., Lan, J., Li, Y., Zeng, Y., & Wang, C. (2014). Automatic method of fruit object extraction under complex agricultural background for vision system of fruit picking robot. *Optik - International Journal for Light and Electron Optics*, 125:19, 5684-5689

Xu, L., Krzyzak A., & Suen, C. Y., (1992). Methods of combining multiple classifiers and their applications to handwriting recognition. *IEEE Transactions on Systems, Man, and Cybernetics*, 22(3), 418-435.

Zhang J., & Tan T. (2002). Brief review of invariant texture analysis methods. *Pattern Recognition*, 35(3), 735-747.

Zhang, Y., & Wu, L. (2012). Classification of fruits using computer vision and a multiclass support vector machine. *Sensors*, 12(9), 12489-12505.