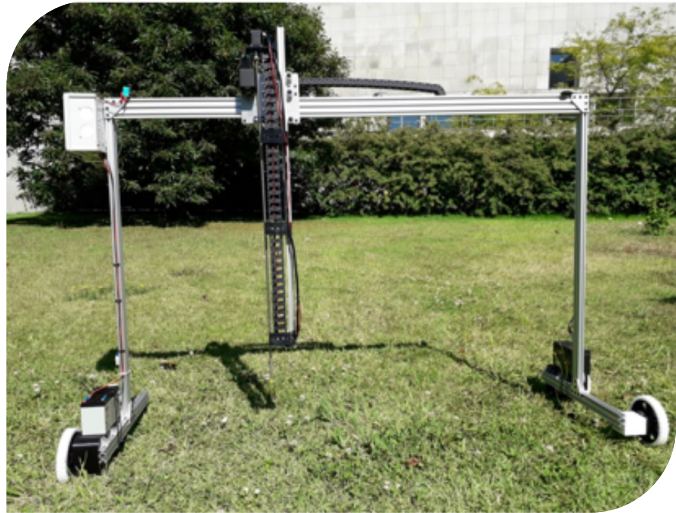




PixelCropRobot, protótipo robótico de baixo custo para fenotipagem em horticultura

Leandro de Almeida Rodrigues

Mestrado em Engenharia Agronómica

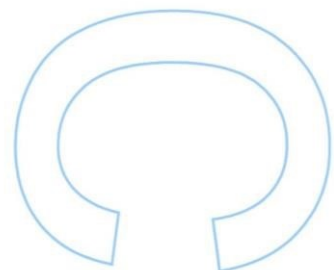
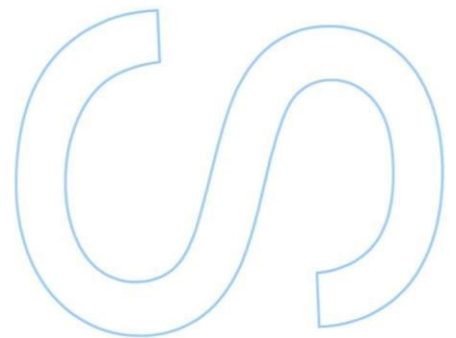
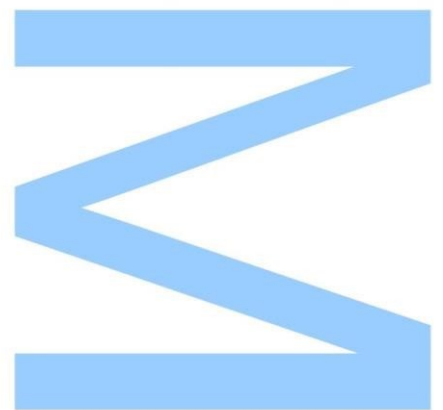
Departamento de Geociências, Ambiente e Ordenamento do Território
2021

Orientador

Professor Doutor Mário Campos Cunha, Professor Associado,
Faculdade de Ciências da Universidade do Porto (FCUP)

Coorientador

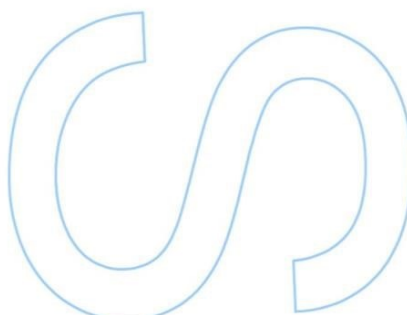
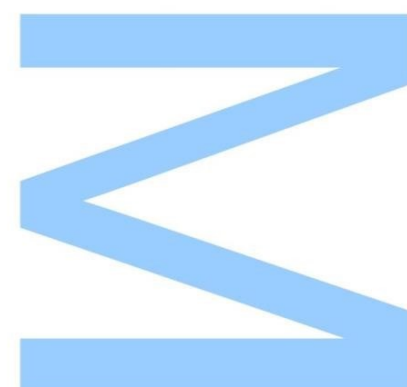
Doutor Filipe Neves dos Santos, Investigador Sénior,
Instituto de Engenharia de Sistemas e Computadores,
Tecnologia e Ciência (INESCTEC)



As correções determinadas pelo júri, e só essas, foram efetuadas.

A Presidente do Júri,

Porto, ____/____/____



“Tu tornas-te eternamente responsável por aquilo que cativas”

Antoine de Saint-Exupéry

Agradecimentos

Permitam-me que dirija uma palavra de agradecimento a todos aqueles que me apoiaram nos últimos meses na elaboração deste trabalho. Um sentido obrigado a todos.

De forma especial, gostaria de agradecer ao meu orientador, Professor Doutor Mário Cunha que tudo fez pelo sucesso deste trabalho, pelos ensinamentos transmitidos, pelo apoio, e pela total disponibilidade, palavras de incentivo e opiniões durante a realização da minha Dissertação. A forma entusiasta como o Professor partilha o conhecimento é contagiante e não deixa ninguém indiferente. Daí a minha resposta, prontamente, afirmativa, quando foi lançado o desafio de desenvolver este trabalho. Foi um ótimo “treinador”.

Ao Doutor Filipe Santos, do INESC TEC, pela oportunidade em trabalhar junto da equipa do CRIIS, ao deixar-me sempre à vontade para usufruir do espaço e dos recursos do laboratório, permitiu-me interagir com temáticas fora da minha zona de conforto e desenvolver novas competências. A dedicação e o apoio foram fundamentais para o sucesso deste trabalho.

Um agradecimento especial à instituição que me acolheu, à Faculdade de Ciências da Universidade do Porto, que me permitiu criar bases de conhecimento suficientes e que, no meu entender, culminam neste trabalho.

À minha família, em particular aos meus pais e à minha irmã, agradeço e reconheço a paciência, o apoio a diversos níveis, que tiveram em todos os momentos, particularmente nos mais difíceis. Sempre se mostraram presentes, apesar da distância, e me deram alento para continuar o trabalho, sem eles nada disto era possível. Foram, e são, um pilar fundamental, na minha formação académica em particular, mas na Vida em geral. Obrigado por me terem deixado sonhar e lutar por um objetivo que sendo meu, se tornou nosso. Espero merecer o vosso orgulho

Aos membros do CRIIS, agradeço a estima, a disponibilidade e as sugestões que foram surgindo durante as nossas reuniões semanais. O meu agradecimento especial ao Francisco Terra, a disponibilidade que teve para esclarecer as dúvidas que foram surgindo. Foi um gosto partilhar conhecimento e colaborar para o sucesso deste projeto. Agradeço ainda ao Sandro Magalhães, ao Pedro Moura e ao Daniel Silva.

Aos sócios da IAAS Portugal e de todos os seus comités em particular aos da IAAS Porto, que partilharam comigo diversas experiências, promovendo o intercâmbio de conhecimento, ideias e (in)formação. FamíliaAS para sempre!

À equipa do Agroportal, ao José Diogo, ao José Maria e ao Paulo, pelo desafio lançado e todo o apoio. Agradeço por sempre me deixarem à vontade para conciliar da forma mais conveniente as tarefas na plataforma com o desenvolvimento deste trabalho.

À Fernanda Martins, responsável pela gestão agrícola do Campus Agrário de Vairão, por sempre se mostrar disponível para “deitar um olho” aos meus ensaios, garantindo que as plantas estavam sempre “felizes”.

Aos meus amigos, agradeço toda a atenção, carinho e disponibilidade, sem eles teria sido muito difícil suportar as adversidades que foram surgindo, sou, também, melhor pessoa à custa deles. Particularmente, à Avó e à Tia, à Van e à Kika, a família que o Porto me ofereceu em 2016 e que foram e continuam a ser essenciais em todo o meu percurso. Que junto delas possa continuar a ser “idiota”. Ao Germano Moreira, que acompanhou de perto este trabalho e sempre se mostrou disponível para apoiar, principalmente, nos momentos em que a tecnologia não colaborava. Ao grupo de boleias da Gi que tornaram as viagens para Vairão mais alegres. Ao meu Soil por nunca ter deixado de brilhar junto de mim.

O meu sincero agradecimento a todos os demais que direta ou indiretamente apoiaram este projeto.

Um bem-haja a todos!

Leandro de Almeida Rodrigues

Resumo

A evolução da investigação no âmbito agrícola, os avanços científicos e tecnológicos nas áreas da robótica, visão computadorizada e Inteligência Artificial permitem o desenvolvimento de operações cada vez mais precisas. Nomeadamente, no que diz respeito à monitorização do crescimento vegetal, é possível avaliar, de forma automática e em tempo real, as características das plantas, principalmente aquelas apresentadas como resultado da interação dinâmica entre a componente genética e o ambiente. Neste trabalho desenvolvemos um protótipo robótico com visão computadorizada. Este protótipo de robô cartesiano está equipado com sensores óticos para a captação *in situ* de imagem de forma autónoma e dinâmica de plantas e solo. Esta informação tem aplicação na fenotipagem de características das plantas para a fundamentação de práticas agrícolas. Foi obtido e anotado um conjunto de imagens (4123) de oito culturas hortícolas (acelga, alface, cenoura, coentro, espinafre, nabo, rabanete e rúcula) em diferentes estádios fenológicos. São aplicados modelos *Deep Learning* (*Single Shot Multibox Detector* e *You Only Look Once*) para processar a informação contida nas imagens de diferentes culturas hortícolas com o objetivo de detetar e classificar (tendo como base uma escala fenológica). O desempenho dos modelos em análise foi avaliado independentemente para a deteção e classificação fenológica de cada cultura. Foi ainda avaliada a capacidade de generalização (versatilidade) para as várias culturas em estudo, considerando a resolução de cada modelo e as características das imagens. São aplicadas diferentes técnicas de fotogrametria às imagens bidimensionais captadas, possibilitando a reconstrução tridimensional de plantas individuais e de grupos de plantas. A estrutura robótica foi desenhada, construída e validada em ambiente laboratorial, mostrando bom desempenho de movimentos. Os modelos *Deep Learning* analisados desempenharam convenientemente ambas as operações de deteção e classificação nas oito culturas em estudo, apresentando um valor médio de *Average Precision* próximo de 81% na deteção, em relação à classificação o valor médio de Exatidão Balanceada ficou próximo dos 82%. O modelo YOLO v4 apresentou o desempenho mais consistente tanto na deteção como na classificação, em comparação com os restantes modelos. Em relação à reconstrução tridimensional das plantas, considera-se que através dos resultados gerados (nuvens de pontos e modelos de superfícies) seja possível avaliar o arranjo geométrico e características relacionadas, parametrizando o crescimento vegetal a diferentes níveis espaciais. A plataforma robótica desenvolvida em conjunto com as informações recolhidas e processadas podem ter diferentes aplicações de fenotipagem (nomeadamente no suporte de operações culturais), teste de sensores, uma forte vertente pedagógica, entre outras.

Palavras-chave – Agricultura de Precisão, Classificação, *Deep Learning*, Deteção, Fotogrametria, Robótica

Abstract

The evolution of agricultural research, scientific and technological advances in robotics, computer vision and Artificial Intelligence allows the development of increasingly precise operations. In what concerns plant growth monitoring, it is possible to evaluate, in an automatic way and in real-time, the characteristics of the plants, mainly those presented as a result of the dynamic interaction between the genetic component and the environment. In this work, we develop a robotic prototype with Computer Vision. This cartesian robot prototype is equipped with optical sensors for autonomous and dynamic *in situ* image capture of plants and soil. This information has application in phenotyping plant characteristics to inform agricultural practices. A set of images (4123) of eight horticultural crops (arugula, carrot, chard, coriander, lettuce, radish, spinach and turnip) at different phenological stages was obtained and annotated. Deep Learning models (Single Shot Multibox Detector and You Only Look Once) are applied to process the information contained in the images of different horticultural crops to detect and classify (based on a phenological scale). The performance of the models under analysis was independently evaluated for the detection and classification of each crop. The generalisation capacity for the various crops considered was also assessed, considering the resolution of each model and the characteristics of the images. Different photogrammetry techniques are applied to the captured two-dimensional images, enabling the three-dimensional reconstruction of individual plants and groups of plants. The robotic structure was designed, built and validated in a laboratory environment, showing good movement performance. The Deep Learning models analysed performed conveniently both detection and classification operations in the eight crops evaluated, presenting mean Average Precision value close to 81% in detection, regarding classification the average Balanced Accuracy value was close to 82%. The YOLO v4 model presented the most consistent performance in both detection and classification, compared to the other models. Regarding the three-dimensional reconstruction of the plants, it is considered that through the generated results (point clouds and surface models) it is possible to evaluate the geometric arrangement and related characteristics, parameterising the plant growth at different spatial levels. The developed robotic platform together with the collected and processed information may have different applications of phenotyping (namely in the support of cultural operations), sensor testing, a strong pedagogical component, among others.

Keywords – Classification, Deep Learning, Detection, Photogrammetry, Precision Agriculture, Robotics

Créditos

Neste trabalho foram incluídos dados, metodologias e resultados publicados em diferentes produtos científicos em que o autor desta dissertação teve uma participação nas diferentes fases da investigação: planeamento, implementação e disseminação. Os produtos de comunicação científica foram:

- Apresentação do artigo intitulado “PixelCropRobot, a cartesian multitask platform for microfarms automation”, no âmbito do 1º Simpósio Internacional de Desenvolvimentos em Robótica Inteligente e Automação Industrial (IRIA) organizado pelo Instituto de Tecnologia Indiano (IIT).
- Publicação do artigo intitulado “*PixelCropRobot, a cartesian multitask platform for microfarms automation*” no IEEE Xplore: Terra, F.; Rodrigues, L.; Magalhães, S.; Santos, F.; Moura, P. and Cunha, M. PixelCropRobot, a cartesian multitask platform for microfarms automation. In Proceedings of the International Symposium of Asian Control Association on Intelligent Robotics and Industrial Automation (IRIA), Goa, India, 20-22 September, 2021; pp. 382-387. doi:<https://doi.org/10.1109/IRIA53009.2021.9588786>.
- Publicação do repositório de imagens do projeto PixelCropRobot na plataforma Zenodo: Rodrigues, L.; Terra, F.; Magalhães, S.; Santos, F.; Moura, P. and Cunha, M. PixelCropRobot dataset: images of vegetables crops in different phenological stages taken in greenhouses. doi:<https://doi.org/10.5281/zenodo.5574515>.

Índice

1. Introdução	1
2. Revisão Bibliográfica	4
2.1. Robótica Agrícola	4
2.1.1. Classificação	4
2.1.2. Automatização da produção hortícola em pequena escala	8
2.2. Fenotipagem vegetal com imagens RGB	11
2.2.1. <i>Deep Learning</i> para deteção e classificação automática de plantas	13
2.2.2. Reconstrução tridimensional de plantas por fotogrametria	18
3. PixelCropRobot – design, montagem e testes	20
4. Material e Métodos	24
4.1. Aquisição, processamento e análise de imagens das culturas	25
4.1.1. Instalação das culturas	25
4.1.2. Captação de imagens RGB	26
4.1.3. Anotação do repositório	28
4.2. Modelos <i>Deep Learning</i> para deteção e classificação fenológica de plantas	31
4.2.1. <i>Augmentation</i> do repositório	31
4.2.2. Divisão do repositório	32
4.2.3. Treino dos modelos SSD e YOLO	34
4.2.4. Avaliação do desempenho dos modelos <i>Deep Learning</i>	37
4.3. Técnicas de fotogrametria para reconstrução tridimensional de plantas	43
4.3.1. Planta individual	45
4.3.2. Grupo de plantas	48
5. Resultados	49
5.1. Desempenho dos modelos SSD e YOLO	49
5.1.1. Alface (LACSA)	49
5.1.2. Resultados gerais das culturas estudadas	64
5.2. Reconstrução tridimensional	66
5.2.1. Planta individual	66
5.2.2. Grupo de plantas	68
6. Discussão	73
6.1. PixelCropRobot	73
6.2. Deteção e classificação fenológica de plantas	74
6.3. Reconstrução tridimensional	76
7. Conclusão	77
8. Referências Bibliográficas	78

Índice de Tabelas

Tabela 1 – Classificação de robôs agrícolas de acordo com os sistemas de mobilidade/locomoção e manipulação/atuação. Adaptado de: [31].	6
Tabela 2 – Descrição dos <i>Technology Readiness Level</i> (TRL). Adaptado de: [33].	6
Tabela 3 – Robôs terrestres de pequeno tamanho para agricultura de pequena escala com produção de hortícolas. Adaptado de: [17,31,50,51].	10
Tabela 4 – Conceito relevantes associados ao <i>Machine Learning</i> . Adaptado de: [76,77].	13
Tabela 5 – Culturas hortícolas selecionadas neste estudo.	25
Tabela 6 – Número de imagens e código EPPO correspondente a cada cultura.	26
Tabela 7 – Classes definidas para cada cultura de acordo com a sua fenologia e ciclo cultural.	30
Tabela 8 – Distribuição de referências pelas classes consideradas nas culturas em estudo nas amostras de treino, validação e teste. Apenas são referidas as imagens não transformadas por <i>augmentation</i> .	33
Tabela 9 – Parâmetros e métricas calculadas para a avaliação do desempenho dos modelos de deteção em estudo nas amostras de validação e teste considerando os problemas de deteção e classificação.	42
Tabela 10 – Confiância que otimiza o F1. Parâmetro calculado na amostra de validação sem <i>augmentation</i> .	50
Tabela 11 – Resultados de várias métricas para os modelos em estudo, considerando todas as deteções e a confiança otimizada, na amostra de teste sem <i>augmentation</i> .	52
Tabela 12 – Matriz de confusão para os modelos SSD e YOLO em estudo na amostra de teste sem <i>augmentation</i> .	54
Tabela 13 – Resultados das métricas exatidão, <i>precision</i> e <i>recall</i> para cada classe de alface para os modelos SSD e YOLO em estudo. Avaliação na amostra de teste sem <i>augmentation</i> .	55

Tabela 14 – *Average Precision* e *Mean Average Precision* dos modelos em estudo. Avaliação executada na amostra de teste sem *augmentation*. Representação das imagens em escala de cinzentos. Resultados para a confiança 0%. 56

Tabela 15 – Valor de confiança que otimiza o valor de F1, na amostra de validação sem *augmentation*. 56

Tabela 16 – Resultados das métricas consideradas para os modelos em estudo. Cálculo executado para todas as detecções e na melhor confiança calculada, na amostra de teste sem *augmentation*. 59

Tabela 17 – Matriz de confusão para os modelos SSD e YOLO em estudo para a confiança otimizada na amostra de teste sem *augmentation*. 61

Tabela 18 – Resultados das métricas exatidão, *precision* e *recall* para cada classe de alface para os modelos SSD e YOLO em estudo. Avaliação na amostra de teste sem *augmentation*. 62

Tabela 19 – Resultados das métricas calculadas na amostra de teste para a confiança otimizada. Métricas calculadas para a detecção: *Average Precision*, *Precision*, *Recall* e F1. Métricas calculadas para a classificação: Exatidão Balanceada e F1 Médio. Resultados gerais para as oito culturas (acelga, alface, cenoura, coentro, espinafre, nabo, rabanete e rúcula). 64

Índice de Figuras

- Figura 1 – Subsistemas e operações de um robô agrícola. Setas preenchidas representam transferência de informação e ações, setas tracejadas representam ligações conceituais. Adaptado de: [2]. 5
- Figura 2 – Exemplos de robôs de pequenas dimensões. 9
- Figura 3 – Funcionamento típico de um sistema de *Machine Learning*. Adaptado de: [79]. 14
- Figura 4 – Características de uma flor. Esquerda: descritores botânicos, direita: descritores de um algoritmo *Deep Learning*. Adaptado de: [80]. 15
- Figura 5 – Desenho digital do PixelCropRobot. Adaptado de: [6]. 1 – Sistema Universal de Ferramentas, 2 – ponte, 3 – guia vertical, 4 – trave da roda, 5 – trave da roda, 6 – guia vertical. 20
- Figura 6 – Elementos que compõem um código QR. Adaptado de: [109]. 21
- Figura 7 – Processo de localização do PixelCropRobot através de códigos QR. Deste processo realçam-se a posição da câmara (no centro da imagem) e a orientação do robô (em graus). Adaptado de: [6]. 22
- Figura 8 – Estrutura do PixelCropRobot. Adaptado de: [6]. 1 – Sistema Universal de Ferramentas, 2 – ponte, 3 – guia vertical, 4 – trave da roda, 5 – trave da roda, 6 – guia vertical. 23
- Figura 9 – Talhões onde foi realizada a instalação das culturas. a) sementeira individual de cada cultura, b) sementeira a lanço da mistura das sementes de todas as culturas estudadas. Imagens captadas a 9 de março, 7 dias após a sementeira. 25
- Figura 10 – Conversão de imagem RGB (à esquerda) em escala de cinzentos (*grayscale*) (à direita). 27
- Figura 11 – Exemplo do interface da plataforma CVAT e a anotação de uma planta de alface. 28
- Figura 12 – Técnicas de *augmentation* aplicadas ao repositório de imagens do PixelCropRobot. 32

Figura 13 – Processo de treino e avaliação do desempenho dos modelos SSD e YOLO em estudo. Adaptado de: [122].	36
Figura 14 – Representação da métrica <i>Intersection Over Union</i> . Adaptado de: [114].	37
Figura 15 – Método para gravação de vídeo para reconstrução tridimensional de plantas individuais. a) rabanete, b) coentro.	43
Figura 16 – Método para gravação de vídeo para a reconstrução tridimensional do grupo de plantas.	44
Figura 17 – Exemplo da imagem de rabanete proposto para reconstrução tridimensional.	45
Figura 18 – Imagem da planta de coentro antes do processamento no PlantCV.	46
Figura 19 – Exemplificação do processo de detecção e classificação de plantas: a) detecção, verde: referências, vermelho: previsões do modelo. b) anotações de referência, c) classificação, cores diferentes significam classes diferentes.	49
Figura 20 – Evolução do F1 com o aumento da confiança, na amostra de validação sem <i>augmentation</i> .	50
Figura 21 – Evolução do número de: a) Verdadeiros Positivos, b) Falsos Positivos e c) Falsos Negativos com o aumento da confiança, na amostra de validação sem <i>augmentation</i> .	51
Figura 22 – Curvas <i>precision</i> × <i>recall</i> para a: a) confiança de 0% e b) confiança otimizada. Avaliação executada na amostra de teste sem <i>augmentation</i> .	53
Figura 23 – Evolução do F1 com o aumento da confiança, na amostra de validação sem <i>augmentation</i> .	57
Figura 24 – Evolução do número de: a) Verdadeiros Positivos, b) Falsos Positivos e c) Falsos Negativos com a evolução da confiança, na amostra de validação sem <i>augmentation</i> .	58
Figura 25 – Curvas <i>precision</i> × <i>recall</i> para a: a) confiança de 0% e b) confiança ótima. Avaliação executada na amostra de teste sem <i>augmentation</i> .	60
Figura 26 – Reconstrução tridimensional de rabanete. a) imagem original, b) nuvem dispersa de pontos, c) nuvem densa de pontos, d) modelo tridimensional.	66

Figura 27 – Reconstrução tridimensional de coentro. a) imagem original, b) região de interesse, c) nuvem dispersa de pontos, d) nuvem densa de pontos, e) modelo tridimensional. 67

Figura 28 – Reconstrução do talhão na fase inicial do ciclo cultural. a) nuvem de pontos dispersa, b) nuvem de pontos densa, c) modelo tridimensional. 69

Figura 29 – Reconstrução do talhão numa fase avançada do ciclo cultural. a) nuvem de pontos dispersa, b) nuvem de pontos densa, c) modelo tridimensional, d) pormenor do modelo tridimensional. 70

Figura 30 – Talhões 42 dias após a sementeira. a) sementeira individual de cada cultura, b) sementeira a lanço da mistura das sementes de todas as culturas. 71

Figura 31 – Reconstrução tridimensional dos talhões. a) nuvem de pontos dispersa, b) nuvem de pontos densa, c) modelo tridimensional 72

Lista de Abreviaturas

2D – Bidimensional

3D – Tridimensional

AP – *Average Precision*

BBCH – *Biologische Bundesanstalt, Bundessortenamt and Chemical Industry*

CNN – *Convolutional Neural Network*

Colab – *Google Colaboratory*

DL – *Deep Learning*

EB – Exatidão Balanceada

FN – Falso Negativo

FP – Falso Positivo

IA – Inteligência Artificial

IOU – *Intersection Over Union*

mAP – *Mean Average Precision*

ML – *Machine Learning*

MVE – *Multi-View Environment*

MVS - *Multi-View Stereo*

ODM – OpenDroneMap

p. ex. – por exemplo

px. – pixéis

RGB – *Red Green Blue*

SfM – *Structure from Motion*

SSD – *Single Shot Multibox Detector*

SUF – Sistema Universal de Ferramentas

VN – Verdadeiro Negativo

VP – Verdadeiro Positivo

YOLO – *You Only Look Once*

1. Introdução

A evolução da investigação no âmbito agrícola, os avanços científicos e tecnológicos nas áreas da robótica, visão computadorizada e Inteligência Artificial (IA) tem vindo a desenvolver suporte para operações agronómicas cada vez mais precisas [1]. Robôs agrícolas são capazes de realizar a generalidade das operações inerentes a um sistema de cultura, num ambiente normalmente pouco estruturado e bastante dinâmico, devido, principalmente a alterações espaço-temporais [2]. A visão computadorizada tem um forte impacto na agricultura, capaz de suportar a tomada de decisão do produtor agrícola [3]. Na generalidade dos casos, os sistemas de visão computadorizada incluem uma fase de captação de imagem, através de câmaras e outros sensores óticos. Após a recolha desta informação é necessário o seu processamento e análise.

As técnicas de *Machine Learning* (ML) estão frequentemente associadas à visão computadorizada e, principalmente, à IA, permitindo que máquinas aprendam por si, com base na experiência. Para tal, são desenvolvidos algoritmos que adquirem conhecimento, através da informação (por eles) analisada, criando modelos capazes de prever e tomar decisões de forma precisa e em tempo real. Esta estratégia tem vasta aplicação, tanto nas decisões agronómicas relacionadas com o desempenho da planta num dado ambiente como para suportar operações culturais [4].

A combinação destas tecnologias permite a obtenção de informação espaço-temporal de elevado débito e dimensionalidade acerca do sistema de cultura. Com isto, é possível avaliar as características das plantas, principalmente aquelas apresentadas como resultado da interação dinâmica entre a componente genética (hereditariedade) e o ambiente – a medição sistemática destas características denomina-se fenotipagem [5].

Neste trabalho propomos um protótipo robótico, o PixelCropRobot, equipado com sensores óticos responsáveis pela captação de imagem para:

- i) deteção (identificação da planta em relação a outros elementos presentes no ambiente) e classificação automática dos estados fenológicos de diferentes culturas agrícolas;
- ii) reconstrução tridimensional a diferentes escalas espaciais (planta individual e grupo de plantas).

Os objetivos específicos deste trabalho são:

- Desenvolver a plataforma robótica PixelCropRobot, desde o seu planeamento e desenho experimental até à validação da estrutura em laboratório;
- Construir um repositório de imagens, vídeos e documentação associada, de várias culturas hortícolas para i) deteção e classificação fenológica e ii) reconstrução tridimensional a diferentes escalas espaciais;
- Testar a tecnologia *Deep Learning* (DL) para a deteção e classificação de culturas hortícolas através das imagens do repositório;
- Aplicar diferentes métodos de fotogrametria para a reconstrução tridimensional e avaliar a versatilidade de diferentes *softwares* específicos para o efeito.

A plataforma robótica foi desenvolvida em colaboração com Terra [6],[7]. Como referido anteriormente, a deteção e classificação de plantas foi realizada através da estratégia DL, baseada em ML, uma abordagem mais robusta, com capacidade de aprendizagem superior e portanto, mais precisa em ambientes complexos [8]. Por sua vez, a reconstrução tridimensional foi desenvolvida através de técnicas de fotogrametria aplicadas a imagens bidimensionais. Ao avaliar, de forma não destrutiva, o arranjo geométrico das superfícies de uma planta, é possível parametrizar o seu desenvolvimento a diferentes níveis espaciais (p. ex.: órgãos, planta completa, canópia de um conjunto de plantas), podendo estimar características como a geometria da planta ou a sua área foliar [9]. À semelhança do relatado anteriormente, também esta informação é importante para suportar, em tempo real, as operações culturais (p. ex.: pulverização, rega ou fertilização de precisão).

Este trabalho é uma das etapas necessárias à construção, validação e, putativa produção de um robô agrícola. Pretende-se, no futuro, que o PixelCropRobot possa ser utilizado como plataforma de testes para o suporte ao desenvolvimento de novas tecnologias (p. ex.: câmaras, sensores, órgãos atuadores e manipuladores).

Pela sua dimensão e versatilidade, esta plataforma sensorizada tem particular interesse em agricultura de pequena escala, próxima dos locais de consumo. Entre outros fatores, esta estratégia de produção de alimentos, pode garantir a qualidade do produto, associada à eficiência económica e à responsabilidade ambiental da produção, nomeadamente pela redução de emissões e custos relacionados com o transporte e conservação dos alimentos [10,11]. Este modo de produção dedica-se, geralmente, ao

cultivo de hortícolas¹ (olericultura) para consumo próprio, frequentemente associada a uma atividade de lazer e de contacto com a natureza, com reduzida expressão mercantilista [12,13]. Muitas vezes, a produção deste tipo de hortas não é suficiente para garantir o rendimento do produtor, o qual é complementado com outra atividade fora da exploração. Por isso, são desenvolvidas numa lógica de tempo parcial, suscetíveis a longos períodos de ausência do produtor [14].

Estes eventuais períodos de ausência poderão ser eliminados através de sistemas autónomos e precisos, tais como o PixelCropRobot, que possibilitem a monitorização constante da horta, nomeadamente a supervisão do crescimento das culturas ao longo do seu ciclo cultural e a identificação das restantes plantas que possam interferir no rendimento destas culturas, as infestantes.

Poderá ainda ser utilizada para ações de sensibilização, conscientização e educação de comunidades estudantis, por exemplo no que diz respeito à tecnologia agrícola ou produtores agrícolas nas temáticas relacionadas com a Agricultura de Precisão (AP) e gestão racional de recursos, e ainda para aplicações em âmbito científico, principalmente no que toca à avaliação da dependência de determinado fenótipo face a uma relação específica entre o genótipo e o ambiente. Estudam-se neste trabalho diferentes vertentes de uma Agricultura evoluída, economicamente eficiente e ecologicamente sustentável.

¹ Horticultura no sentido restrito incluindo apenas as plantas herbáceas destinadas à alimentação humana.

2. Revisão Bibliográfica

2.1. Robótica Agrícola

2.1.1. Classificação

De acordo com a ISO 8373:2012 um robô é um equipamento multifuncional e programável desenvolvido para mover materiais, ferramentas e equipamentos especializados para executar diferentes tarefas e operações. A Federação Internacional de Robótica considera uma classificação baseada nos domínios da indústria onde os robôs operam, mencionando a “classificação A” como sendo Agricultura, Floresta e Pescas [15].

De acordo com Gonzalez-de-Santos, Fernández et al. [16] as principais diferenças funcionais entre robôs industriais e agrícolas prendem-se com o modo como estes se movem em relação aos restantes elementos: i) em ambiente industrial, os equipamentos são dispostos em posições fixas enquanto os materiais se movem nas linhas de produção, ii) em ambiente agrícola, frequentemente, o robô move-se em relação às plantas ou outro material que assume uma posição fixa. A deslocação de um robô agrícola, é um verdadeiro desafio, por se tratar de um ambiente dinâmico e pouco estruturado, com variações do relevo do solo, das condições de luminosidade, períodos de oclusão, entre outros fatores [17].

Os robôs agrícolas são enquadrados em dois grupos principais: robôs móveis (terrestres ou aéreos) ou manipuladores, normalmente associados a robôs móveis (manipuladores móveis), que atuam sobre a produção agrícola [16]. Os robôs móveis terrestres são capazes de realizar a generalidade das operações inerentes a um sistema de cultura: sementeira [18], controlo de infestantes [19-21], deteção de pragas e doenças [22,23], análise da cultura (p. ex.: monitorização do vigor) [24,25], colheita [26,27], entre outras. Por forma a rentabilizar o equipamento robótico este poderá realizar diversas operações específicas a uma cultura ou generalizadas a um grupo de culturas. Estes equipamentos são habitualmente denominados de robôs de múltiplos propósitos [28-30].

Normalmente, associado ao manipulador existe o sistema de sensorização, responsável pela recolha de informação relacionada com a cultura que pode ser aplicada na tomada de decisão ao nível de determinada operação (p. ex.: deteção de infestantes, maturação de frutos, deteção de sinais e sintomas relacionados com pragas

ou doenças). Para além disso, é importante que o robô conheça a sua posição face ao restante ambiente, que detete facilmente obstáculos, evitando o contacto com estes. Para tal, o sistema de locomoção e atuação deve ser adequado à produção e à operação a executar [2]. Estes sistemas devem interatuar entre si. A relação entre robô e humano é fundamental, por motivos de segurança e manutenção. Bechar e Vigneault [2] afirmam que os sistemas mencionados desempenham um papel fundamental no desempenho de um robô agrícola e por isso, podem ser considerados como operações complementares ou coadjuvantes que este deverá ser capaz de executar (Figura 1).

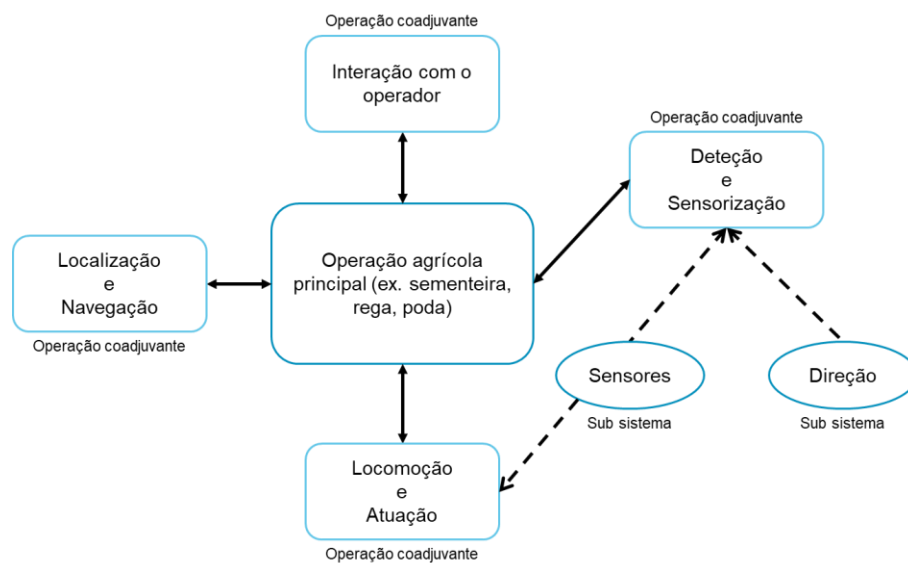


Figura 1 – Subsistemas e operações de um robô agrícola. Setas preenchidas representam transferência de informação e ações, setas tracejadas representam ligações conceptuais. Adaptado de: [2].

É possível classificar os diferentes robôs agrícolas com base nas características dos seus sistemas. Ribeiro, Andujar et al. [31] descrevem um sistema de classificação baseado na autonomia demonstrada por um robô agrícola tendo em conta os seus sistemas de mobilidade/locomoção e manipulação/atuação (Tabela 1). Em ambos os casos a escala é definida de 1 – totalmente dependente a 5 – totalmente autónomo. No primeiro caso, avalia-se a autonomia de locomoção e mobilidade. No segundo, a avaliação recai sobre a autonomia na seleção das operações agrícolas a realizar. É ainda considerada a sensibilidade do robô face à envolvente.

Tabela 1 – Classificação de robôs agrícolas de acordo com os sistemas de mobilidade/locomoção e manipulação/atuação. Adaptado de: [31].

Nível	Mobilidade/Locomoção	Manipulação/Atuação
1	Totalmente dependente, dirigível, mas sem sensibilidade face à envolvente	Totalmente dependente, desenvolve ações fixas sem sensibilidade face à envolvente
2	Semi dependente, dirigível e com sensibilidade parcial à envolvente	Semi dependente, desenvolve ações fixas com sensibilidade parcial face à envolvente
3	Intermédio, parcialmente dirigível de forma autónoma, com sensibilidade intermédia ou nula à envolvente	Intermédio, parcialmente independente, com sensibilidade intermédia ou nula à envolvente
4	Semi autónomo, direção autónoma, com sensibilidade intermédia à envolvente	Semi autónomo, seleção de ações autónoma e sensibilidade parcial face à envolvente
5	Totalmente autónomo, direção autónoma, compreende totalmente a envolvente	Totalmente autónomo, seleção de ações autónoma, sensibilidade total à envolvente

Os robôs agrícolas podem ser classificados de acordo com o *Technology Readiness Level* (TRL) [32] correspondente. O TRL é um indicador do nível de maturidade tecnológica. Numa escala de 0 a 9 (Tabela 2) são enquadradas diferentes tecnologias em quatro grupos distintos, desde a sua fase inicial (conceito) até ao momento em que a tecnologia está pronta para ser comercializada (produção).

Tabela 2 – Descrição dos *Technology Readiness Level* (TRL). Adaptado de: [33].

Nível	Grupo	Fase	Descrição
TRL 0	Conceito	Ideia	Conceitos ainda não foram testados
TRL 1		Pesquisa Básica	Identificação do estado da arte
TRL 2		Formulação da Tecnologia	Conceção de possíveis aplicações
TRL 3	Protótipo	Pesquisa Aplicada	Testes laboratoriais e prova de conceito
TRL 4		Teste em Escala Reduzida	Validação da tecnologia em ambiente controlado
TRL 5		Teste em Escala Piloto	Modelo validado em ambiente simulado
TRL 6	Validação	Protótipo em Teste	Situação próxima à do desempenho esperado
TRL 7		Demonstração	Protótipo analisado em ambiente operacional
TRL 8	Produção	Fase Pré-comercial	Tecnologia pronto e validada em ambiente real
TRL 9		Aplicação da Tecnologia	O produto está pronto para ir para o mercado

Como já referido anteriormente, os robôs agrícolas deverão ser capazes de operar em ambientes pouco ou nada estruturados e garantir a mesma qualidade do trabalho executado pelos métodos comuns (p. ex.: manuais, maquinaria agrícola supervisionada por operador). No caso particular dos robôs agrícolas terrestres, este é considerado um dos principais constrangimentos ao seu desenvolvimento e distribuição comercial generalizada [34]. Um robô desenvolvido para a produção agrícola considera-se viável se a sua aplicação cumprir pelo menos um dos seguintes critérios [2]:

- Aumenta ou mantém a produtividade e a qualidade das operações;
- Permite aumentar a ergonomia e diminuir a penosidade de tarefas ou operações agrícolas;
- Permite a execução de tarefas perigosas para o produtor ou até impossíveis de realizar por processos mecânicos ou manuais alternativos (p. ex.: em situações de escassez de mão-de-obra);
- Tantos os custos de investimento como os operacionais do sistema robótico, a produtividade e qualidade dos produtos resultantes, quando comparados a processos alternativos (mecânicos ou manuais), devem garantir a competitividade da exploração agrícola no mercado em que se insere;
- Minimiza a incerteza da produção agrícola;
- Garante benefícios nos âmbitos ambiental ou social [35].

De acordo com Bechar e Vigneault [36], a tendência aponta para a redução do tamanho dos robôs agrícolas, geralmente com custos inferiores. Tipicamente, sistemas robóticos mais pequenos consomem menos energia e apresentam um menor impacto ambiental (p. ex.: menor compactação do solo). De facto, esta é a principal vantagem da utilização de robôs de pequeno tamanho, em particular quando estes executam operações que requerem movimentação contínua do robô. Este facto é exacerbado em produções com intensidade de atuação fitotécnica elevada, como é exemplo a produção de horticolas.

2.1.2. Automatização da produção hortícola em pequena escala

Diversas empresas dispõem de robôs de pequena dimensão para gerir diferentes sistemas de produção de hortícolas, alguns deles apresentados na Figura 2. A *Iron Ox* [37], uma *startup* americana criada em 2018, integra robôs na produção hidropónica de manjerição e outras hortícolas. O *Grover* desloca-se de forma autónoma, realiza o transporte das plantas (dispostas em tabuleiros) para junto do *Ada* (Figura 2a), um braço robótico articulado que realiza o controlo de pragas e doenças. A empresa *Sky Greens* [38] dispõe de um sistema que apoia a produção vertical de hortícolas, através de calhas de alumínio conectadas à estrutura robótica. O movimento hidráulico autónomo da estrutura garante a homogeneidade de luminosidade, rega e nutrientes ao longo das calhas. A *Agbotic* [39] dispõe de um pórtico robótico que realiza a generalidade das operações em horticultura protegida.

A *Easton Robotics* [40] encontra-se a desenvolver o *ERMMI* (Figura 2b), um robô autónomo, desenhado para operar em explorações de pequena dimensão. Será responsável por mobilizar o solo, semear, aplicar fertilizantes e colher. A empresa francesa Naïo Technologies [41] apresenta o *Oz* (Figura 2c), um robô de propósitos múltiplos que dispõe de um conjunto de ferramentas que realizam a sementeira, o controlo de infestantes, operações de logística, mobilização e armação do solo). O *Tertill* (Figura 2d) [42] é considerado o primeiro robô capaz de realizar o controlo de infestantes em pequenas hortas residenciais. A Bosch e Amazone desenvolveram o *BoniRob* (Figura 2e) [43] capaz de detetar infestantes através de IA, realizando a sua remoção de forma mecânica.

A empresa americana FarmBot [44] apresenta o *FarmBot Genesis* (Figura 2f), um sistema *open-source* para produção semiautomática de pequena escala. Consiste num robô cartesiano de três eixos (semelhante a uma impressora 3D), *software* e documentação associada e ainda um repositório de dados agrícolas constantemente atualizado pelos consumidores [45]. Pelas características referidas (ser *open-source* e uma plataforma cartesiana móvel já consolidada no mercado), o *FarmBot Genesis* tem relevância particular no âmbito deste trabalho. O seu *hardware* permite que ferramentas como injetores de sementes, emissores de água e fertilizantes, supressores de infestantes e diversos sensores, executem movimentos precisos num sistema de coordenadas com posicionamento rigoroso em relação ao solo e às plantas. O sistema baseia-se na colocação do equipamento em guias fixas para controlo de trajetória que permite localizar com precisão as plantas, possibilitando a execução das operações

culturais da forma mais conveniente à espécie e à fase de desenvolvimento da planta [46].

Por sua vez, o projeto *ROMI* [47] tem como principal objetivo a conscientização de produtores agrícolas para a adoção de tecnologia. Os principais desenvolvimentos deste projeto são: o *ROMI Rover* (Figura 2g), robô para controlo autónomo e mecânico de infestantes e o *cablebot*, equipamento responsável pela recolha de imagem e operações de fenotipagem. Estes equipamentos apesar de independentes, interatuam entre si numa plataforma colaborativa [48].

Existem diversos robôs ainda em desenvolvimento por parte da comunidade científica, por exemplo: Grimstad e From [29] apresentam o *Thorvald II* (Figura 2h), um robô modular, altamente reconfigurável, capaz de operar em diversos sistemas de cultura. As diversas configurações deste robô foram testadas em ambientes distintos (por exemplo: mobilização do solo, fenotipagem de cereais, operações de logística e monitorização). Axbom e Ralsgård [49] apresentam o protótipo de um robô (Figura 2i) leve, de baixo custo, capaz de detetar e eliminar infestantes, por aplicação de herbicida.

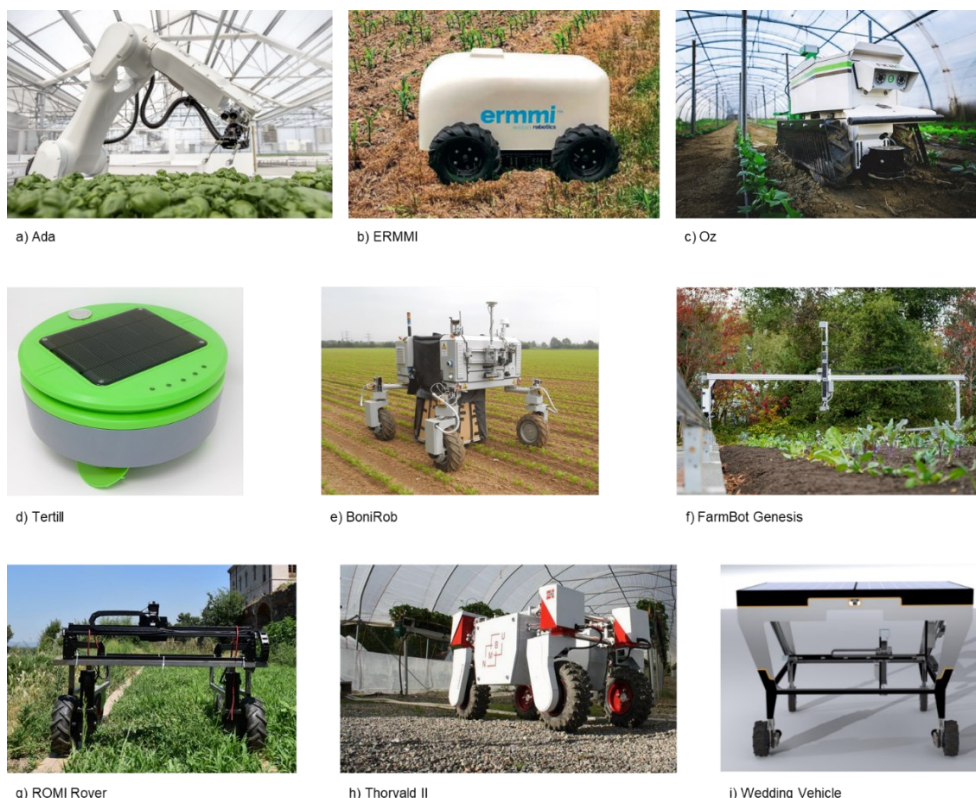


Figura 2 – Exemplos de robôs de pequenas dimensões.

A Tabela 3 resume as características dos principais robôs que operam em sistemas agrícolas de pequena escala com foco na produção de hortícolas.

Tabela 3 – Robôs terrestres de pequeno tamanho para agricultura de pequena escala com produção de hortícolas. Adaptado de: [17,31,50,51].

Robô	Operação										Configuração e Movimentos	TRL	Classificação de Ribeiro, Andujar et al. [31]*	
	M	S	R	I	F	Fito	Feno	L	C	Mob			Mobilidade	Manipulação
Ada [37]							X				A	9	s.s.	3
AgBot [52]				X	X	X					R	6	4	2
Agbotic [39]	X	X	X								P c/ rodas	9	2	2
Agrobot [53]	X				X	X	X				R	6	4	2
BoniRob [43]				X		X	X				R	8	4	3
Di-Wheel [54]		X		X		X					R	7	4	2
ERMMI [40]		X			X				X	X	R	5	4	2
FarmBot Eden [55]		X	X	X	X		X				Po	5	2	3
FarmBot Genesis v1.5 [56]		X	X	X	X		X				C c/ guias	Comercial	2	3
Greenbot [57]	X	X		X	X						R	Comercial	5	3
Grover [37]								X			P c/ rodas	9	4	3
Jackal [58]	X										R	Comercial	4	3
Land Care Robot [59]	X										R	Comercial	4	2
Multifunctional Robot [60]			X								P fixo	5	3	2
Multipurpose Robot [61]				X		X					R	4	4	2
Oz [41]	X	X		X							R	Comercial	4	2
RobotGreen [62]	X										A	5	s.s.	3
ROMI Rover [47]				X			X				C c/ rodas	7	4	2
Tertill [42]				X							R	Comercial	4	2
Thorvald II [29]	X						X	X			R	7	4	3
Weeding Vehicle [49]				X							R	7	3	2

F: Fertilização, Feno: Fenotipagem, Fito: Aplicação de Produtos Fitofarmacêuticos, I: Controlo de Infestantes, L: Logística, M: Monitorização, R: Rega, S: Sementeira, C: Colheita, Mob: Mobilização do solo

P c/ rodas: Pórtico com rodas, A: Articulado, R: Rodas, Po: Polar, C c/ guias: Cartesiano com guias, C c/ rodas: Cartesiano com rodas, P fixo: Pórtico fixo

TRL: *Technology Readiness Level*, s.s.: sem significado

*a classificação de Ribeiro, Andujar et al. [31] pode ser verificada na Tabela 1.

2.2. Fenotipagem vegetal com imagens RGB

Como mencionado anteriormente, a fenotipagem de plantas consiste na análise quantitativa e/ou qualitativa das características das plantas, principalmente aquelas apresentadas como resultado da interação entre a componente genética e o ambiente. A análise poderá ser realizada a diferentes níveis espaciais (p. ex.: células, tecidos, órgãos, plantas) e considerar a evolução das características ao longo do tempo [63].

No contexto agronómico, as características fenotípicas são principalmente estudadas para:

- i) Caracterização estrutural, a análise de características como a altura das plantas, o índice de área foliar, entre outros, apoia a tomada de decisões agronómicas, através, por exemplo da monitorização do crescimento vegetal e previsão do rendimento [64];
- ii) Detecção de plantas ou órgãos destas, a análise de características como a cor, forma ou temperatura das planta e/ou dos seus órgãos (p. ex.: frutos, ramos) apoia na execução automática de operações culturais (p. ex.: controlo de infestantes, poda, colheita) [65,66];
- iii) Avaliação fisiológica, a análise de características ao nível fisiológico (p. ex.: défices nutricionais, concentração de clorofila) [67] permite otimizar o desempenho dos processos relacionados com o metabolismo.

Para que seja possível a análise das características fenotípicas é necessário sistemas equipados com sensores para recolha de informação, normalmente sobre a forma de imagens. Neste trabalho, serão utilizadas imagens RGB (*Red Green Blue*), uma vez que se trata de uma técnica de baixo custo, que permite obter informação acerca da cor, textura e arranjo geométrico (bidimensional) das plantas. É uma técnica comumente aplicada para operações de caracterização e deteção [5,68]. São ainda utilizadas neste trabalho imagens em escala de cinzentos (*grayscale*), uma representação das imagens RGB com o objetivo de particularizar o contraste entre os objetos presentes na imagem [69].

Após a captação de imagens é necessário a sua organização e processamento, tendo em vista a extração de características de interesse. Diversas técnicas associadas à visão computadorizada permitem um processamento conveniente das imagens [70,71].

No âmbito deste trabalho recorremos a duas técnicas para processamento de imagens RGB:

- Algoritmos DL: o objetivo passa pela deteção (identificação da planta em relação a outros elementos presentes na imagem) e classificação (baseada em descritores fenológicos) de culturas hortícolas. Estes algoritmos têm capacidade de extrair características de forma automática com elevado desempenho [72];
- Técnicas de fotogrametria: neste caso, o objetivo passa pela reconstrução tridimensional a diferentes escalas espaciais (plantas individuais e grupos de plantas). Em particular foi aplicado o método *Structure from Motion* (SfM) que permite a construção de um modelo tridimensional (3D) de um objeto ou ambiente, através de imagens RGB bidimensionais (2D) [73].

2.2.1. *Deep Learning* para detecção e classificação automática de plantas

Os algoritmos de ML, um dos ramos da IA, são capazes de aprender determinado comportamento baseado na experiência. Muitas vezes associado ao ML surge a visão computorizada, permitindo que os sistemas vejam, interpretem e compreendam a informação, normalmente através de imagens digitais [74,75]. Na Tabela 4 são apresentados alguns conceitos associados ao ML, que facilitam a descrição e compreensão dos processos apresentados neste trabalho.

Tabela 4 – Conceito relevantes associados ao *Machine Learning*. Adaptado de: [76,77].

Conceito	Descrição
Exemplo	dados aplicados no processo de aprendizagem
Característica	conjunto de atributos associados a um exemplo
Hiperparâmetro	parâmetros específicos (<i>inputs</i>) do algoritmo
Conjunto de hipóteses	funções responsáveis pela relação entre as características e as anotações das características de interesse no sistema em análise
Anotação	valor associado a cada exemplo
<i>Bounding box</i>	representação geométrica da anotação
Função loss	medida que avalia a diferença entre a anotação verdadeira e aquela prevista pelo algoritmo
<i>Epoch</i>	valor associado ao número de vezes que todos os exemplos são submetidos no algoritmo
Amostra de teste	exemplos utilizados para avaliar o desempenho do algoritmo
Amostra de treino	exemplos utilizados para treinar o algoritmo
Amostra de validação	exemplos que permitem o ajuste dos parâmetros do algoritmo

Uma das principais características que distingue os algoritmos de ML, é a forma como acontece a aprendizagem. Esta pode ser:

- supervisionada, são conhecidos os *inputs* e os *outputs*. Isto é, a cada exemplo é concedida uma anotação, que pode ser na forma de *bounding box*. O algoritmo fica responsável por encontrar a melhor correspondência entre o *input* e o *output*, através das características extraídas [78];
- sem supervisão, neste caso, não são fornecidas informações acerca dos *inputs*, sendo o algoritmo responsável pela sua estruturação;
- por reforço, os algoritmos que seguem este tipo de aprendizagem, atuam pelo princípio de tentativa e erro, pela procura da melhor decisão que produz o melhor *output* [79].

De forma geral, um sistema de ML compreende duas fases fundamentais: treino e teste (Figura 3). Na primeira fase, após o processamento dos dados – a amostra de treino – é selecionado um conjunto de características que seguirão para os algoritmos responsáveis pelo treino do sistema. Para ajustar o conjunto de hipóteses aplicado e melhorar o desempenho do algoritmo, na fase de teste é introduzida nova informação – a amostra de teste [79].

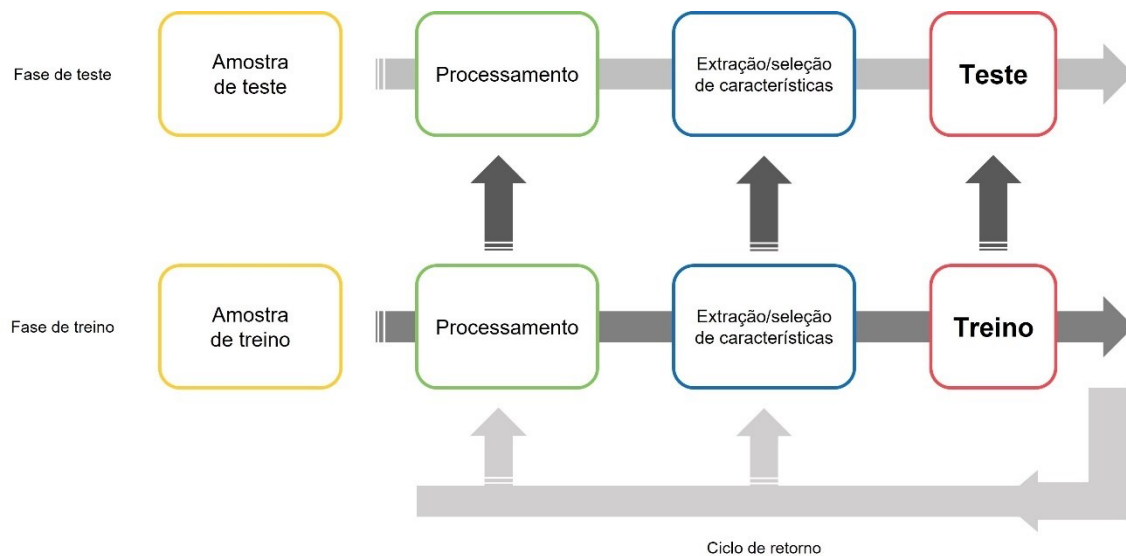


Figura 3 – Funcionamento típico de um sistema de *Machine Learning*. Adaptado de: [79].

Em especial os algoritmos DL, de aprendizagem supervisionada, asseguram a transformação dos dados com representação hierárquica, garantindo o devido processamento das imagens dadas como *input* com elevado desempenho e a extração automática de características, superando as restantes técnicas comumente utilizadas [8,72]. Para exemplificar, a Figura 4 compara as características botânicas de uma flor e algumas das consideradas por um algoritmo DL.

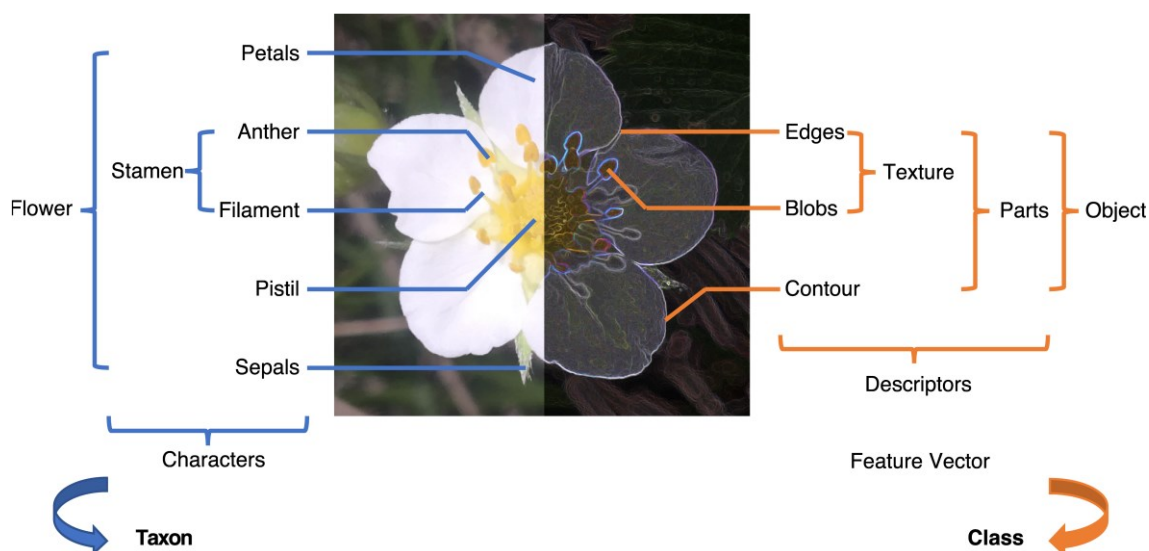


Figura 4 – Características de uma flor. Esquerda: descritores botânicos, direita: descritores de um algoritmo *Deep Learning*. Adaptado de: [80].

Dentro do conjunto de algoritmos DL, importa mencionar o grupo *Convolutional Neural Networks* (CNN), de acordo com Kamilaris e Prenafeta-Boldú [81] este conjunto de algoritmos é, provavelmente, aquele com maior popularidade e sucesso no contexto agrícola. A estrutura hierarquizada e a elevada capacidade de aprendizagem permite que estes algoritmos desempenhem devidamente a deteção e classificação de plantas e outros elementos, inclusive em situações complexas, como normalmente acontece no ambiente agrícola [82,83].

Para que o algoritmo seja capaz de detetar e classificar vários objetos, nomeadamente plantas, numa só imagem, é necessário construir um modelo de deteção de objetos. Estes modelos incluem um CNN, responsável pela extração de características (de forma automática, como já foi referido) e um algoritmo de classificação (classificador) que prevê a posição dos objetos e lhes atribui a classe correspondente [77]. No âmbito deste trabalho foram testadas duas arquiteturas distintas de modelos de deteção, *Single Shot Multibox Detector* (SSD) e *You Only Look Once* (YOLO). Estas serão descritas na secção 4.2.3.

Na detecção e classificação de plantas, é possível dividir em dois grupos principais as plantas que se pretendem detetar ou classificar: infestantes e culturas agrícolas. São enumerados por Hasan, Sohel et al. [84] um conjunto de modelos DL já aplicados para detecção de infestantes a partir de imagens digitais. Em particular, em relação aos modelos CNN, são referidos *AlexNet*, *VGG*, *GoogLeNet*, *ResNet 50*, entre outros.

Para a detecção e classificação de culturas são desenvolvidos e aplicados diferentes modelos de detecção DL. No âmbito deste trabalho são particularmente relevantes aqueles que apresentem um CNN como algoritmo de extração de características e que as operações de detecção e classificação sejam feitas a partir de imagens RGB.

Por exemplo, Zheng, Kong et al. [85] aplicam um conjunto de modelos CNN (*Faster R-CNN*, *SSD*, *RFB*, *YOLOv2*, *YOLOv3* e *RetNet*) para a detecção de diversas culturas. Os autores desenvolveram um repositório de imagens onde constam 30 categorias de diferentes culturas hortícolas num total de cerca de 30000 imagens. Os autores referem que a precisão média dos modelos considerados nas culturas de alface, nabo e espinafre foi próxima de 97%, 74% e 64%, respetivamente. Por sua vez, Chechliński, SiemiąTkowska et al. [86] desenvolvem um modelo (tendo por base o algoritmo *VGG*) responsável pela detecção de alface através de imagens de campo, a precisão média foi de 96 a 98% nos diferentes ensaios.

Chavan e Nandedkar [87] desenvolvem um modelo, a partir do algoritmo *VGG* para extração de características. O modelo foi aplicado na classificação das culturas beterraba, milho e trigo, através do *Plant Seedlings Dataset* [88], um repositório com 5544 imagens de doze espécies (as três culturas referidas e nove espécies de infestantes). Todas as imagens foram captadas numa fase inicial do ciclo vegetativo (BBCH² 10-14). Os autores referem uma exatidão próxima de 94%. De forma semelhante, Dyrmann, Karstoft et al. [89] desenvolvem um modelo capaz de detetar e classificar plantas na fase “cotilédones” de 22 espécies, tanto culturas como infestantes. É reportada pelos autores uma precisão média próxima de 86%.

É proposto por Knoll, Czymmek et al. [90] um modelo capaz de classificar cenoura e diferentes espécies de infestantes, com precisão próxima de 98%. Czymmek, Harders et al. [91] apresentam um modelo capaz de detetar e classificar as plantas de forma

² BBCH Monograph (<https://www.julius-kuehn.de/media/Veroeffentlichungen/bbch%20epaper%20en/page.pdf>). Último acesso em: 24/04/2021.

simultânea. Um sistema capaz de detetar infestantes junto da cultura (cenoura) em tempo real sem segmentação. A precisão média do modelo aplicado foi próxima de 75%.

Milioto, Lottes et al. [92] apresentam um modelo capaz de distinguir beterraba, de infestantes e do solo. Por sua vez, Lottes, Behley et al. [93] apresentam um sistema de classificação baseado em imagens sequenciais. Através da sobreposição de diferentes imagens é realizada a classificação da cultura (beterraba), infestantes e solo. Em ambos os casos é utilizado o *BoniRob* [43] para a captação de imagens.

Chaaro e Antón [94] implementam um modelo capaz de detetar espinafre e duas infestantes: dente de leão e amor de hortelão. A precisão média foi próxima de 78% para a cultura, 45% para o dente de leão e 53% para o amor de hortelão. A tecnologia desenvolvida foi testada no *FarmBot Genesis*, anteriormente apresentado.

Como já referido anteriormente, importa, no âmbito deste trabalho que os modelos de deteção sejam capazes de identificar a fase do ciclo cultural de cada planta, realizando assim a classificação baseada em descritores fenológicos. Yalcin [95] aplica um modelo CNN (algoritmo *AlexNet* para extração de características) na classificação fenológica de trigo, cevada, lentilha, algodão, pimento e milho. O autor refere que o repositório de imagens utilizado apresenta cerca de 400 imagens por cultura. É reportado pelo autor um desempenho médio próximo de 83%. Por sua vez, Samiei, Rasti et al. [96] pretenderam averiguar a faculdade germinativa de luzerna e trevo vermelho, para tal, testaram o desempenho de diferentes modelos (*VGG16*, *ResNet50*, *DenseNet121*) na classificação de três fases distintas do início do ciclo cultural (emergência do cotilédono, abertura do cotilédono e desenvolvimento da primeira folha). É reportada pelos autores uma precisão média próxima de 80%.

Todos os trabalhos referidos anteriormente realizam a deteção de plantas através de imagens RGB bidimensionais. Contudo, são referidas por Agarwal, du Terrail et al. [77] outras modalidades de deteção de objetos, não só através de imagens bidimensionais com resoluções espectrais distintas (p. ex.: multi ou hiperespectral [97]), mas também a partir de vídeos ou informação tridimensional. Em relação a esta última modalidade, serão avaliadas neste trabalho, diferentes técnicas para reconstrução tridimensional de plantas através de imagens RGB bidimensionais.

2.2.2. Reconstrução tridimensional de plantas por fotogrametria

De acordo com Li, Guo et al. [71] a fenotipagem de características das plantas baseada em sistemas tridimensionais engloba três etapas principais: i) aquisição de imagens 3D, ii) processamento das imagens 3D e iii) análise das imagens 3D com o objetivo de avaliar determinadas características fenotípicas. Paulus [9] refere um conjunto de características fenotípicas que podem ser avaliadas através de técnicas de reconstrução tridimensional, a diferentes níveis espaciais, por exemplo ao nível da planta individual (p. ex.: comprimento da folha [98]) ou a um grupo de plantas (p. ex.: altura das plantas [99]). Como já referido, as características avaliadas têm aplicação tanto nas decisões agronómicas relacionadas com o desempenho da planta num dado ambiente (p. ex.: através da análise do crescimento ao longo do ciclo cultural) como para suportar operações culturais.

A aquisição de imagens 3D pode ser realizada de forma direta ou a sua construção ser estabelecida através de imagens 2D. No âmbito deste trabalho são consideradas técnicas de triangulação, nomeadamente a técnica SfM [100]. Esta técnica permite a reconstrução tridimensional de um espaço ou objeto a partir de imagens RGB bidimensionais [71,73]. A combinação de imagens sequenciais em diferentes perspetivas desse objeto, permite a sua projeção tridimensional através de uma nuvem de pontos [101]. Pela utilização dos parâmetros de calibração calculados pelo SfM, o método *Multi-View Stereo* (MVS) [102] reconstrói a cena tridimensional, aumentando consideravelmente a quantidade de pontos, tornando a nuvem mais densa. Por último, a nuvem é convertida numa malha de superfícies através de processos de reconstrução específicos.

Wintz, Lahlou et al. [103] desenvolveram um modelo capaz de reconstruir plantas de tomate. O sistema descrito pelos autores permite avaliar o desenvolvimento de diferentes órgãos desta cultura (principalmente folhas e caules). Liu, Acosta-Gamboa et al. [104] descrevem um sistema, de baixo custo, capaz de executar a reconstrução tridimensional de plantas, através de imagens RGB. O sistema apresentado divide-se em: i) uma câmara para captação de imagens e uma plataforma rotatória, ii) um método de segmentação automática (através da plataforma PlantCV [105]) e iii) um modelo de reconstrução tridimensional (a partir da técnica SfM). Yang e Han [106] apresentam um conjunto de métodos que permitem o processamento de vídeo (captado através de um *smartphone*) e a criação de nuvens de pontos para medição de parâmetros fenotípicos (altura da planta, número e comprimento de folhas) e avaliação do crescimento de hortícolas. Martinez-Guanter, Ribeiro et al. [107] apresentam um método, através de

SfM, capaz de realizar a reconstrução de plantas de milho, beterraba e girassol. A reconstrução foi realizada a partir do *software* Agisoft Metashape³. Após a comparação dos resultados obtidos com outros métodos, os autores concluem que, tendo em conta as culturas que disponham e as suas características, o método de reconstrução SfM, a partir de imagens RGB não é eficaz. Jay, Rabatel et al. [108] avaliam a altura das plantas e a área foliar através da reconstrução das linhas onde se encontram as culturas instaladas. Pelos resultados reportados, os autores referem que os parâmetros estudados foram devidamente medidos com valores próximos da realidade. Assim, concluem que o método SfM pode ser aplicado para a reconstrução tridimensional de culturas agrícolas instaladas em linha. Não obstante os bons resultados do método SfM apresentados por estes autores, a sua operabilidade em sistemas agrícolas pouco estruturados é ainda pouco conhecido, nomeadamente: condições ambientais de aquisição das imagens, posicionamento de aquisição relativamente ao objeto, eficiência de amostragem, variabilidade entre aquisições com diferentes sensores com características diferentes, recursos computacionais e tempo de execução e nível de automatização (aquisição, processamento, produção de informação e interface com o utilizador), entre outros.

³ Agisoft Metashape (<https://www.agisoft.com/>). Último acesso em: 25/08/2021.

3. PixelCropRobot – design, montagem e testes

O PixelCropRobot é um robô agrícola cartesiano de três eixos móveis desenvolvido no âmbito deste trabalho [7]. Pretende-se que seja capaz de realizar diversas operações agrícolas: i) sementeira, ii) rega, iii) controlo de infestantes, e principalmente, iv) monitorização de culturas com base em diferentes características fenotípicas.

Tendo como base de estudo as tecnologias já existentes no mercado, o PixelCropRobot foi desenhado com inspiração particular no *FarmBot Genesis v1.5*⁴. O desenho da solução foi elaborado com recurso ao *software Fusion360*⁵ e pode ser observado na Figura 5.

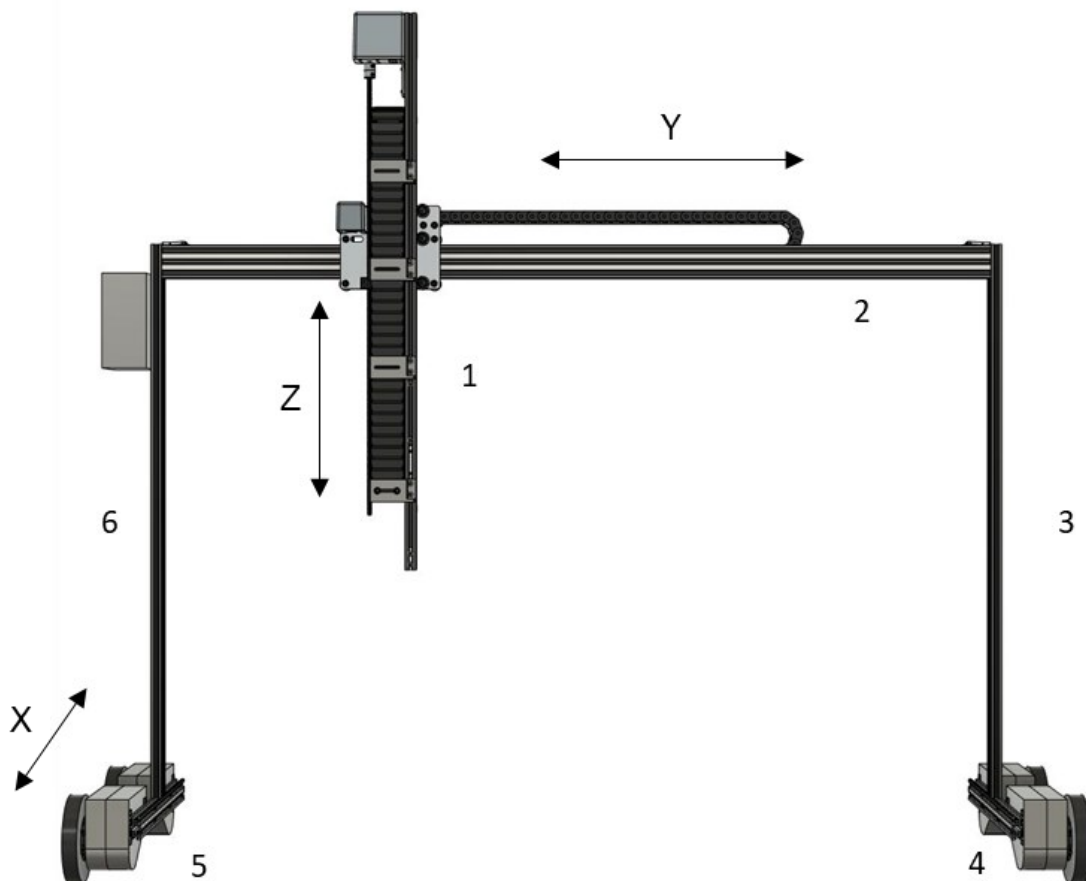


Figura 5 – Desenho digital do PixelCropRobot. Adaptado de: [6]. 1 – Sistema Universal de Ferramentas, 2 – ponte, 3 – guia vertical, 4 – trave da roda, 5 – trave da roda, 6 – guia vertical.

⁴ FarmBot - Open Source CNC Farming (<https://farm.bot/>). Último acesso em: 22/07/2020.

⁵ Fusion360 (<https://www.autodesk.com/products/fusion-360/overview>). Último acesso em: 15/06/2021.

A estrutura deste robô consiste em seis perfis de alumínio *V-slot*, numerados de 1 a 6 (Figura 5):

- 1, o Sistema Universal de Ferramentas (SUF), o perfil ao qual se conectam os diversos órgãos atuadores, tubagens, sensores óticos e demais componentes eletrónicos.
- 2, na horizontal, a ponte, um perfil mais forte, que garante rigidez e força para suportar o SUF;
- 3 e 6, dispostos na vertical, as guias verticais, conferem suporte à ponte e à caixa de eletrónica;
- 4 e 5, as traves da roda, suportam as respetivas rodas;

Os movimentos realizados pelas rodas representam o eixo X, os movimentos de translação do SUF sobre a ponte o eixo Y e os movimentos em altura do SUF o eixo Z.

O controlo do robô é conseguido com recurso a uma *Raspberry Pi 4*⁶ conectada a uma *AgloT*⁷, *hardware* desenvolvido no INESC TEC que confere a capacidade de atuação em potência dos motores. Adicionalmente, foi integrada no circuito uma nova *PCB (Printed Circuit Board)*, desenhada com o fim de controlar individualmente os motores de passo, presentes nos eixos Y e Z (Figura 5).

As aplicações que visam o controlo da estrutura foram desenvolvidas tendo como base a arquitetura *ROS 2*⁸. Nesse seguimento e, considerando a representação em grafo dos processos em execução, foi necessário projetar um conjunto de nós, os quais estão interligados a partir da publicação e subscrição de tópicos ou através de serviços.

O sistema de localização do PixelCropRobot é desenvolvido através de códigos QR (Figura 6). Estruturalmente, cada código apresenta três padrões distintos: deteção, alinhamento e *timing*.



Figura 6 – Elementos que compõem um código QR. Adaptado de: [109].

⁶ Raspberry Pi 4 (<https://www.raspberrypi.org/products/raspberry-pi-4-model-b/>). Último acesso em: 16/06/2021.

⁷ *AgloT* (<http://agiot.inesctec.pt/>). Último acesso em: 25/07/2021.

⁸ Robot Operating System (<https://docs.ros.org/en/foxy/index.html>). Último acesso em: 28/08/2021.

O robô é capaz de reconhecer e decodificar os códigos durante o seu movimento (em particular os seus padrões). O processo de deteção é realizado pela câmara Raspberry Pi Camera Module v2⁹ na resolução 1280 × 720 píxéis (px.). A câmara encontra-se fixa a uma altura de 92 cm, são utilizados códigos QR quadrados com 9,5 cm de comprimento, dispostos a uma distância de 20 cm entre eles, garantindo sempre três códigos no campo de visão da câmara. O processo de localização do PixelCropRobot é exemplificado na Figura 7.

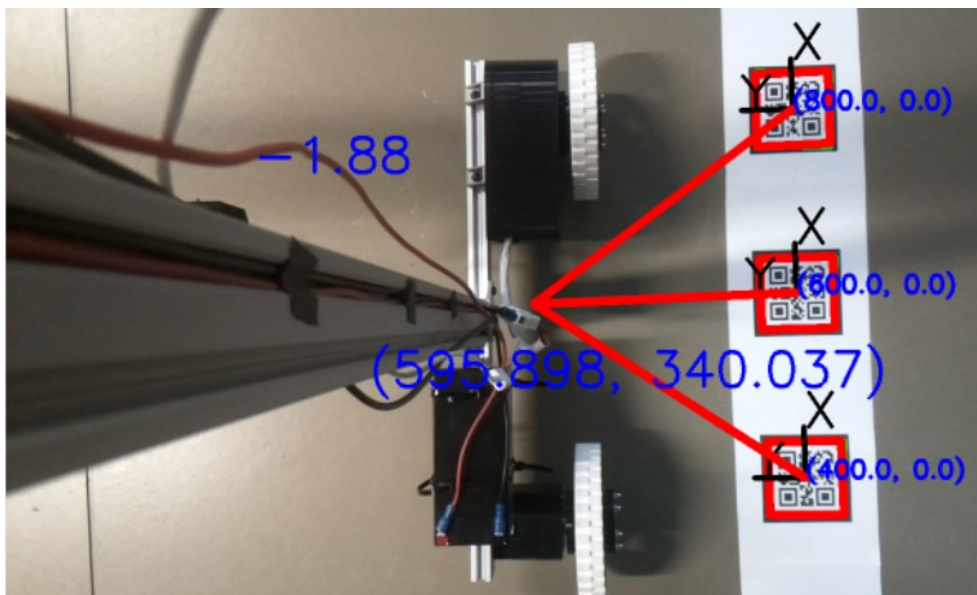


Figura 7 – Processo de localização do PixelCropRobot através de códigos QR. Deste processo realçam-se a posição da câmara (no centro da imagem) e a orientação do robô (em graus). Adaptado de: [6].

Através de uma abordagem *bottom-up*, procedeu-se durante a montagem do PixelCropRobot, a diversos testes. Os componentes eletrónicos foram testados individualmente antes de serem integrados no conjunto. Após a validação do controlo dos motores e do módulo da câmara, do funcionamento dos diferentes sistemas e a comunicação entre os mesmos, deu-se início aos testes do robô por completo.

⁹ Raspberry Pi Camera Module v2 (<https://www.raspberrypi.org/products/camera-module-v2/>). Último acesso em: 16/06/2021

Destes testes importa referir a velocidade máxima associada aos movimentos de cada eixo:

- 7,7 cm/s no eixo X, assegurando a deteção dos códigos QR;
- 2 cm/s no eixo Y (1/4 de *step*), o SUF percorre uma distância de cerca de 1,2 m neste eixo;
- 0,4 cm/s no eixo Z, minimizando ao máximo a oscilação do varão roscado, responsável pelos movimentos deste eixo.

A Figura 8 apresenta a estrutura do PixelCropRobot, no Anexo I surge a lista de materiais necessários à construção do robô, na sua maioria são componentes disponíveis para compra (*off-the-shelf*).

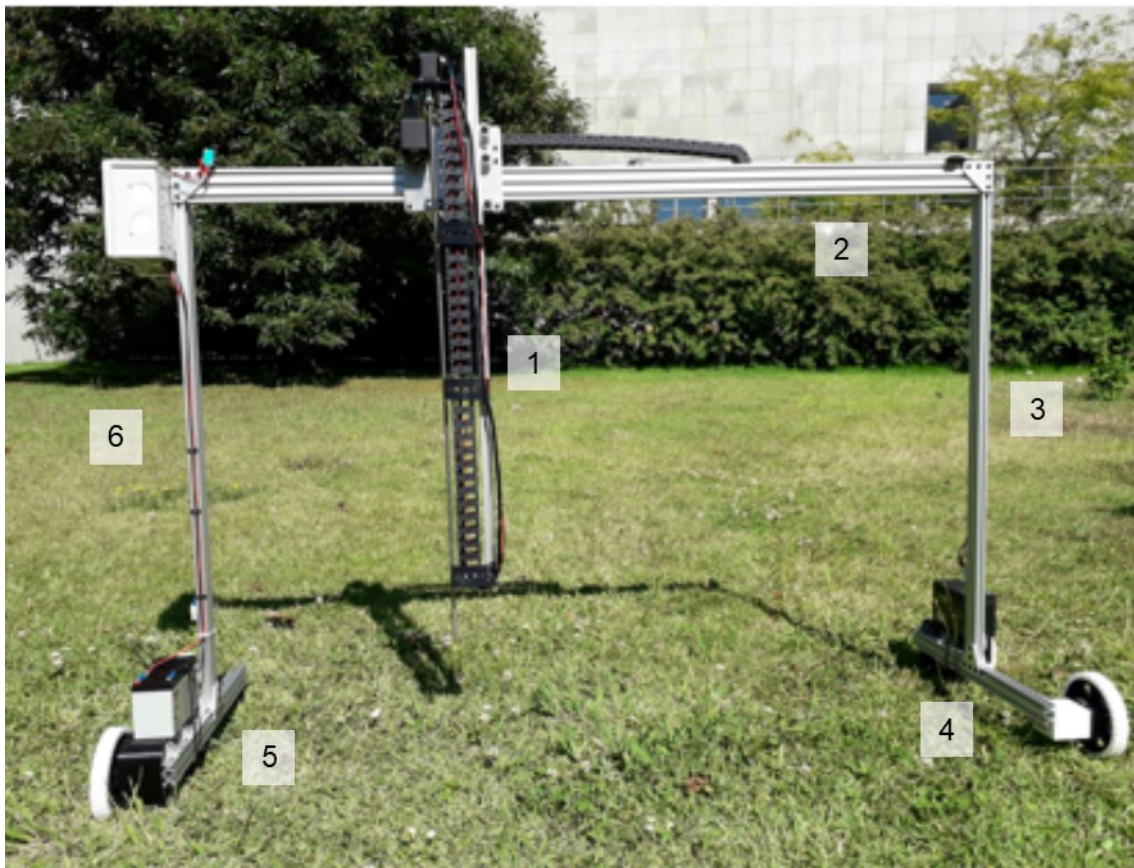


Figura 8 – Estrutura do PixelCropRobot. Adaptado de: [6]. 1 – Sistema Universal de Ferramentas, 2 – ponte, 3 – guia vertical, 4 – trave da roda, 5 – trave da roda, 6 – guia vertical.

4. Material e Métodos

O número reduzido de repositórios de imagens de acesso público é um dos principais constrangimentos ao desenvolvimento de técnicas de visão computadorizada e algoritmos de detecção e classificação de plantas e seus órgãos, úteis para operações associadas à AP (p. ex.: controlo de infestantes, detecção de frutos) [110]. A construção de repositórios de imagens anotadas com diferentes características fenotípicas, entre outras, ao permitir que os pares possam utilizar esta informação para o desenvolvimento e treino de novos modelos de detecção, é um processo altruísta de grande contributo para a comunidade científica

Lu e Young [110] enunciam diversos repositórios de imagens com acesso livre para diferentes operações de visão computadorizada. Na maioria dos casos não são captadas imagens de culturas relevantes (no âmbito deste trabalho), noutras são captadas imagens com características não avaliadas neste trabalho. Mais recentemente, Hasan, Sohel et al. [84] apresentam vários repositórios aplicados na detecção de culturas e infestantes. Apesar de alguns destes repositórios conterem imagens das culturas em estudo [111-113], não apresentam imagens das classes fenológicas relevantes para a classificação.

Face à limitação de imagens disponíveis para o propósito do nosso trabalho, procedemos à criação de um repositório de imagens próprio, já publicado e disponível na plataforma *Zenodo*: Rodrigues, L.; Terra, F.; Magalhães, S.; Santos, F.; Moura, P. and Cunha, M. PixelCropRobot dataset: images of vegetables crops in different phenological stages taken in greenhouses. doi:<https://doi.org/10.5281/zenodo.5574515>. Para além das imagens, constam neste repositório um conjunto de vídeos (úteis para a reconstrução tridimensional de plantas) e diversa documentação (a maioria em linguagem *Python*) que permite operar sobre os dados. Seguidamente, são descritos os processos efetuados para a construção do repositório, iniciando com a instalação das culturas selecionadas.

4.1. Aquisição, processamento e análise de imagens das culturas

4.1.1. Instalação das culturas

No âmbito deste trabalho, foram consideradas oito culturas hortícolas (Tabela 5), selecionadas de acordo com os seguintes critérios: ciclo cultural curto (entre um e três meses), baixa intensidade de operações culturais e baixa pressão de pragas e doenças e com ampla utilização na alimentação.

Tabela 5 – Culturas hortícolas selecionadas neste estudo.

Nome comum	Nome comum (EN)	Nome científico
Acelga	Swiss chard	<i>Beta vulgaris</i> subsp. <i>vulgaris</i> var. <i>cicla</i> L.
Alface	Lettuce	<i>Lactuca sativa</i> L.
Cenoura	Carrot	<i>Daucus carota</i> subsp. <i>sativus</i> (Hoffm.) Schübl. & Martens
Coentro	Coriander	<i>Coriandrum sativum</i> L.
Espinafre	Spinach	<i>Spinacia oleracea</i> L.
Nabo	Turnip	<i>Brassica rapa</i> L.
Rabanete	Radish	<i>Raphanus sativus</i> L.
Rúcula	Arugula	<i>Eruca vesicaria</i> subsp. <i>sativa</i> (Mill.) Thell.

Estas culturas foram instaladas em dois talhões, com área de 4,5 m² cada. A instalação foi realizada numa estufa localizada no Campus Agrário de Vairão (Figura 9).

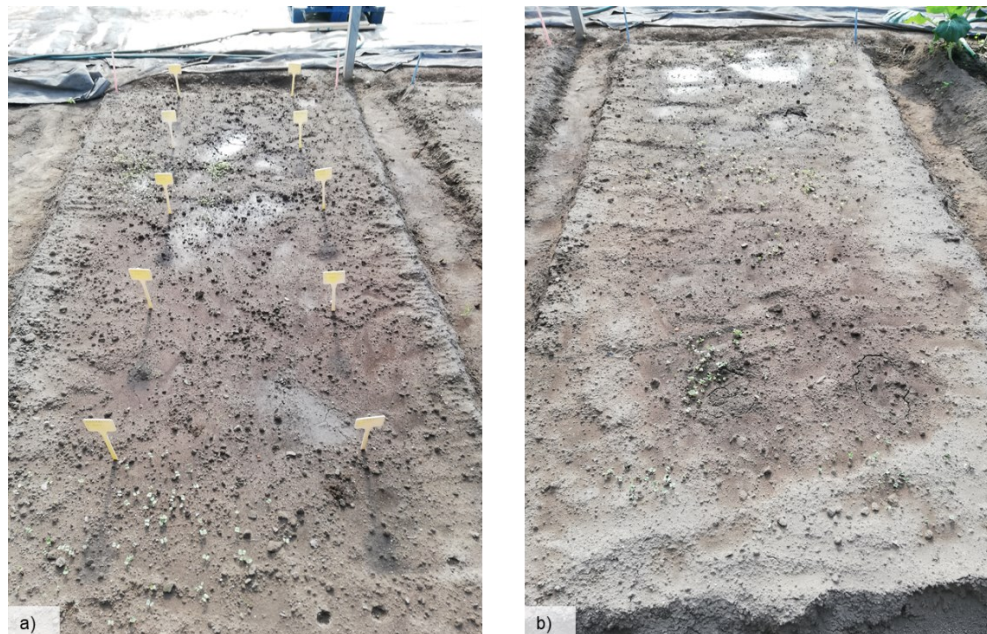


Figura 9 – Talhões onde foi realizada a instalação das culturas. a) sementeira individual de cada cultura, b) sementeira a lanço da mistura das sementes de todas as culturas estudadas. Imagens captadas a 9 de março, 7 dias após a sementeira.

O talhão apresentado na Figura 9a foi dividido em secções. Em cada secção foi semeada uma cultura. No outro talhão (Figura 9b) realizou-se a sementeira a lanço da mistura das sementes de todas as culturas. Através da comparação visual dos talhões conseguimos compreender a influência do crescimento de cada cultura nas restantes e o impacto das modalidades de sementeira no crescimento de infestantes. Em ambos os talhões surgiram, principalmente, as seguintes infestantes: *Galinsoga parviflora*, *Senecio vulgaris*, *Cyperus rotundus*, *Fumaria officinalis*, *Portulaca oleracea*, *Ranunculus repens*, *Poa annua*. A remoção manual das infestantes ocorreu numa fase inicial do seu crescimento e foi precedida da captação de imagens.

4.1.2. Captação de imagens RGB

As imagens RGB foram captadas com recurso a um *smartphone* (*Huawei Mate 10 Lite*), com resolução de 16 megapixéis, 3456×4608 px., no modo Profissional (sem flash, autofócus contínuo, ISO e velocidade do obturador automáticos). A recolha de imagens decorreu em diferentes períodos do dia, com condições de luminosidade variáveis, favorecendo a variabilidade das imagens recolhidas [3]. A Tabela 6 apresenta o número de imagens correspondente a cada cultura. É referido na Tabela 6 o código EPPO¹⁰ correspondente a cada cultura.

Tabela 6 – Número de imagens e código EPPO correspondente a cada cultura.

Nome comum	Código EPPO	Nº de imagens
Acelga	BEAVV	454
Alface	LACSA	1426
Cenoura	DAUCS	533
Coentro	CORSA	321
Espinafre	SPQOL	270
Nabo	BRSTR	313
Rabanete	RAPSR	494
Rúcula	ERUVE	312

Para além das imagens RGB, as imagens de alface sofreram uma transformação em escala de cinzentos (*grayscale*). Com esta transformação pretende-se, numa primeira instância perceber a versatilidade dos modelos previamente treinados com imagens RGB. Serve ainda para atestar a viabilidade do PixelCropRobot captar imagens *grayscale*, tipicamente mais leves e menos influenciadas pelas condições de luminosidade.

¹⁰ European and Mediterranean Plant Protection Organization (<https://www.eppo.int/>). Último acesso em: 7/05/2021.

A conversão das imagens RGB foi efetuada através da aplicação dos seguintes pesos a cada uma das bandas do espectro eletromagnético: $(0.2989 \times Red + 0.5870 \times Green + 0.1140 \times Blue)$, utilizando o seguinte *script* desenvolvido em linguagem *Python*:

```
1 import os
2 import cv2
3 import numpy as np
4 import argparse
5 img_folder = ''
6 final = ''
7 images = os.listdir(img_folder)
8 for i in images:
9     name,h = os.path.splitext(img_folder + '/' + i)
10    img = cv2.imread(img_folder + '/' + i)
11    (B, G, R) = cv2.split(img)
12    merged = cv2.merge([(0.1140*B + 0.5870*G + 0.2989*R), (0.1140*B + 0.5870*G +
13    0.2989*R), (0.1140*B + 0.5870*G + 0.2989*R)])
14    cv2.imwrite(final + '/' + i, merged)
```

A Figura 10 apresenta um exemplo da transformação de imagens RGB para a escala de cinzentos.



Figura 10 – Conversão de imagem RGB (à esquerda) em escala de cinzentos (*grayscale*) (à direita).

Para uniformizar o repositório, procedeu-se à alteração do nome de cada imagem captada, de acordo com o código EPPO correspondente (Tabela 6) e a data de criação dessa imagem. Foi ainda reduzido o tamanho de cada imagem, em quatro vezes (para 864×1152 px.) para facilitar o processamento. Por exemplo, uma imagem de alface captada no dia 22 de junho apresenta o nome da seguinte forma: *LACSA_Jun_22_x_864_1152.jpg*.

4.1.3. Anotação do repositório

No âmbito deste trabalho foram treinados modelos de detecção com aprendizagem supervisionada, ou seja, foi necessário anotar, manualmente, o repositório, indicando através de *bounding boxes* retangulares a posição e a classe fenológica de cada planta. Estas anotações são consideradas como referências. O processo de anotação foi realizado na plataforma CVAT¹¹, cuja interface está representada na Figura 11.

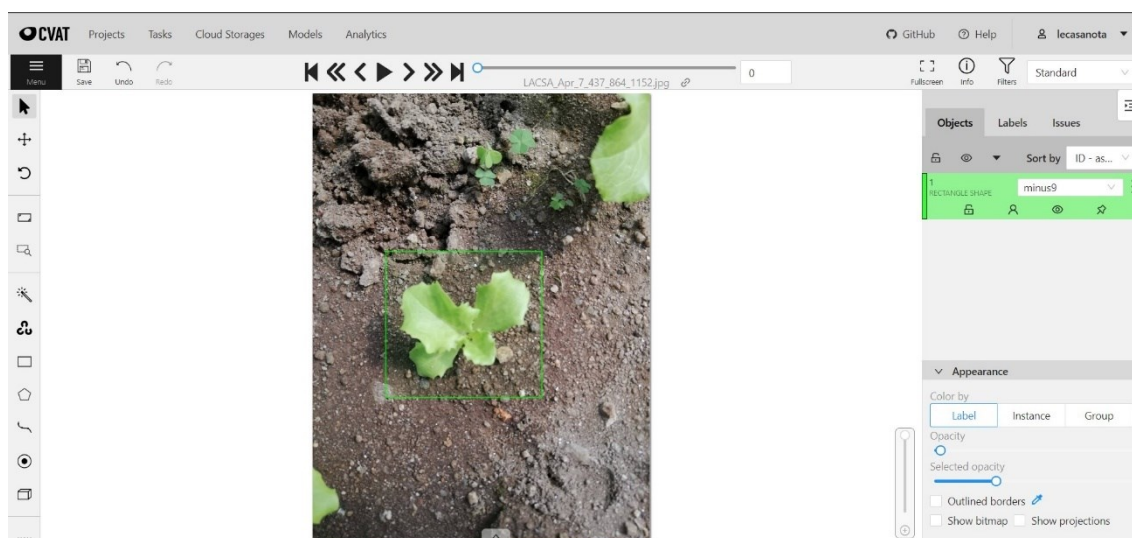


Figura 11 – Exemplo do interface da plataforma CVAT e a anotação de uma planta de alface.

Esta plataforma permite criar diferentes formatos de anotação, tendo sido selecionados os formatos Pascal VOC e o formato YOLO para a anotação das imagens do repositório criado no âmbito deste trabalho. O formato Pascal VOC consiste num ficheiro *.xml* para cada imagem, onde pode constar uma ou várias anotações das plantas presentes na imagem; para cada anotação é referida a classe correspondente. Para as imagens que não apresentam plantas das classes relevantes é criado um ficheiro em branco.

O formato YOLO consiste num ficheiro *.txt* que transmite as anotações da imagem correspondente. Cada linha do ficheiro representa uma anotação. Para além do identificador numérico da classe correspondente, contém as coordenadas da *bounding box*. Este formato necessita de um ficheiro adicional onde a cada identificador numérico corresponda uma classe [114]. De seguida, é apresentado um exemplo de cada formato de anotação para a mesma imagem (Figura 11). Ambos os exemplos dizem respeito à representação dos ficheiros obtidos na plataforma CVAT.

¹¹ Computer Vision Annotation Tool (<https://cvat.org/>). Último acesso em: 23/08/2021.

Pascal VOC:

```
<?xml version="1.0"?>
- <annotation>
  <folder>LACSA</folder>
  <filename>LACSA_Apr_07_6_864_1152.jpg</filename>
- <source>
  <database>Unknown</database>
  <annotation>Unknown</annotation>
  <image>Unknown</image>
</source>
- <size>
  <width>864</width>
  <height>1152</height>
  <depth/>
</size>
<segmented>0</segmented>
- <object>
  <name>minus9</name>
  <occluded>0</occluded>
  - <bndbox>
    <xmin>197.5</xmin>
    <ymin>441.9</ymin>
    <xmax>564.05</xmax>
    <ymax>749.77</ymax>
  </bndbox>
</object>
</annotation>
```

YOLO:

```
1 0.4395543981481481 0.5163498263888889 0.42424768518518513 0.2672482638888889
```

Para cada espécie foram definidas várias classes de acordo com a fase do ciclo cultural, tendo em conta a escala fenológica BBCH¹² (um sistema que permite a codificação homogénea dos estados fenológicos de plantas), garantindo que diferentes classes apresentam características fenológicas distinguíveis. A seleção destas classes é fundamental, não só do ponto de vista agronómico, garantindo a monitorização do ciclo cultural, mas também para atestar a capacidade dos modelos DL em realizar a classificação fenológica nas condições descritas neste trabalho. A Tabela 7 indica as classes de cada cultura. Por exemplo, no caso da alface são definidas quatro classes: i) *coty*, BBCH 9-10, ii) *minus9*, BBCH 11-18, iii) *plus9*, BBCH 19 e iv) *ready*, BBCH 49.

¹² BBCH Monograph (<https://www.julius-kuehn.de/media/Veroeffentlichungen/bbch%20epaper%20en/page.pdf>). Último acesso em: 24/04/2021.

Tabela 7 – Classes definidas para cada cultura de acordo com a sua fenologia e ciclo cultural.

Cultura	Classes			
Acelga (BEAVV)				
	<i>coty</i>		<i>chard</i>	
Alface (LACSA)				
	<i>coty</i>	<i>minus9</i>	<i>plus9</i>	<i>ready</i>
Cenoura (DAUCS)				
	<i>coty</i>	<i>smallleaves</i>		<i>carrot</i>
Coentro (CORSA)				
	<i>coty</i>	<i>smallleaves</i>		<i>coriander</i>
Espinafre (SPQOL)				
	<i>spinach</i>		<i>big</i>	
Nabo (BRSRR)				
	<i>coty</i>	<i>smallleaves</i>		<i>turnip</i>
Rabanete (RAPSR)				
	<i>coty</i>	<i>smallleaves</i>	<i>bigleaves</i>	<i>root</i>
Rúcula (ERUVE)				
	<i>coty</i>	<i>minus9</i>		<i>plus9</i>

No repositório de imagens do PixelCropRobot constam as imagens de cada cultura nas duas resoluções referidas (3456 × 4608 px. e 864 × 1152 px.), as anotações das imagens 864 × 1152 px., nos formatos Pascal VOC (.xml) e YOLO (.txt) e ainda um conjunto de *scripts* (programas executáveis) em linguagem *Python* úteis na uniformização do repositório.

4.2. Modelos *Deep Learning* para deteção e classificação fenológica de plantas

4.2.1. *Augmentation* do repositório

A preparação do repositório para o treino dos modelos de deteção iniciou-se pela *augmentation*. Este processo engloba um conjunto de técnicas de processamento de imagem, responsáveis pela transformação das imagens. Assim, aumenta-se o tamanho e a variabilidade do repositório. A aplicação deste tipo de imagens, geradas por *augmentation*, no treino de modelos de deteção, melhora o desempenho destes, tornando-os mais robustos [115].

Pela aplicação das técnicas disponíveis no repositório *FruitDetect4AgRob*¹³, as imagens do nosso repositório sofreram as seguintes operações de *augmentation* (Figura 12):

- *scale* – aumento ou diminuição do tamanho;
- *angle* – rotação num ângulo específico;
- *translate* – translação;
- *flip* – inversão horizontal;
- *multiply* – aumento ou diminuição das condições de luminosidade;
- *blur* – desfoque gaussiano;
- *noise* – ruído gaussiano;
- combinações das anteriores.

¹³ *FruitDetect4AgRob* (<https://gitlab.inesctec.pt/agrob/fruitdetect4agrob>). Último acesso em: 17/07/2021.



Figura 12 – Técnicas de *augmentation* aplicadas ao repositório de imagens do PixelCropRobot.

4.2.2. Divisão do repositório

A técnica de validação cruzada permite avaliar a capacidade de generalização de um modelo DL, ou seja, permite avaliar o desempenho de um modelo num conjunto de dados não utilizados na sua construção.

Para tal, dividimos o total das imagens disponíveis, de forma aleatória, em três amostras: i) treino – imagens e respetivas anotações (exemplos) aplicadas ao treino do modelo, ii) validação – observações utilizadas para parametrizar o modelo durante o treino do mesmo e iii) teste – observações independentes daqueles utilizados durante o treino, que operacionalizam o modelo, processo denominado de inferência, sendo úteis na avaliação do desempenho do modelo [76]. São adicionadas às amostras de treino e validação as imagens provenientes da *augmentation* e respetivas anotações.

A Tabela 8 indica o número de anotações de referência existentes nas classes consideradas (Tabela 7) para cada uma das amostras (treino, validação e teste). É referido o número de imagens de cada amostra para as oito culturas em estudo.

Tabela 8 – Distribuição de referências pelas classes consideradas nas culturas em estudo nas amostras de treino, validação e teste. Apenas são referidas as imagens não transformadas por *augmentation*.

Cultura	Classe	Treino		Validação		Teste	
		ref	img	ref	img	ref	img
Acelga	<i>coty</i>	32	319	12	90	11	45
	<i>chard</i>	441		129		63	
Alface	<i>coty</i>	664	665	226	471	144	290
	<i>minus9</i>	2420		889		569	
	<i>plus9</i>	248		99		81	
	<i>ready</i>	348		166		111	
Cenoura	<i>coty</i>	486	321	157	106	131	106
	<i>smallleaves</i>	492		173		165	
	<i>carrot</i>	563		180		212	
Coentro	<i>coty</i>	331	163	188	64	224	94
	<i>smallleaves</i>	191		73		118	
	<i>coriander</i>	326		158		170	
Espinafre	<i>spinach</i>	122	162	37	54	41	54
	<i>big</i>	53		24		16	
Nabo	<i>coty</i>	101	158	66	93	116	62
	<i>smallleaves</i>	197		111		92	
	<i>turnip</i>	246		224		80	
Rabanete	<i>coty</i>	650	347	166	98	53	49
	<i>smallleaves</i>	664		190		52	
	<i>bigleaves</i>	218		57		39	
	<i>root</i>	163		47		34	
Rúcula	<i>coty</i>	172	188	40	62	69	62
	<i>minus9</i>	440		93		98	
	<i>plus9</i>	45		12		9	

ref: quantidade de referências, img: quantidade de imagens

4.2.3. Treino dos modelos SSD e YOLO

Como referido anteriormente, um modelo de deteção de objetos inclui um CNN, responsável pela extração de características, de forma automática, e um algoritmo de classificação (classificador) que prevê a posição dos objetos atribuindo-lhes a classe correspondente [77]. No âmbito deste trabalho serão aplicados para a deteção e classificação de plantas duas arquiteturas distintas de “modelos de passo único, nos quais as etapas de localização e classificação de objetos são realizadas em simultâneo. Os modelos de passo único são mais expeditos na deteção e classificação de objetos o que possibilita a sua aplicação em tempo real. Por se tratarem das arquiteturas mais populares e também por serem as aplicadas nas principais linhas de investigação do INESCTEC, são apresentadas as arquiteturas SSD (*Single Shot Multibox Detector*) [116] e YOLO (*You Only Look Once*) [117].

Na arquitetura SSD, primeiro é realizado o mapeamento das características da imagem dada como exemplo, através de um algoritmo de deteção (p. ex.: *VGG16*, *ResNet*, *MobileNet*). Depois, a partir do conjunto padrão de *bounding boxes* é prevista a posição dos objetos de interesse presentes no exemplo considerado, tendo por base o mapeamento prévio [118,119]. Por sua vez, os modelos YOLO (*You Only Look Once*) apresentados primeiramente por Redmon, Divvala et al. [117] enquadram a deteção de objetos como um problema de regressão. Desta forma, aplicam um único algoritmo à imagem completa. Esta imagem é dividida em regiões, a cada região é atribuído um conjunto de *bounding boxes* e associada uma probabilidade. Tipicamente estes modelos aplicam o algoritmo *Darknet* para a extração de características, disponível na plataforma desenvolvida por Redmon [120].

Existem várias plataformas que permitem o treino e aplicação de algoritmos DL para extrair características de imagens, nomeadamente: PyTorch¹⁴, Darknet¹⁵ e TensorFlow¹⁶. Tipicamente, estas plataformas dispõem de modelos previamente treinados (*transfer learning*) através de bases de dados de uso partilhado, como já mencionado anteriormente: Pascal VOC, ImageNet, MS COCO ou Open Images Dataset. Através de *transfer learning*, os modelos previamente treinados são aplicados a novos problemas, com diferentes conjuntos de dados. A transferência de conhecimento entre domínios equiparáveis permite não só economizar recursos (p. ex.: poder computacional ou tempo necessário ao treino), mas também colmatar a falta de

¹⁴ PyTorch (<https://pytorch.org/>). Último acesso em: 15/06/2021.

¹⁵ Darknet (<http://pjreddie.com/darknet/>). Último acesso em: 15/06/2021.

¹⁶ TensorFlow (<https://www.tensorflow.org/>). Último acesso em: 15/06/2021.

dados e a distribuição pouco balanceada destes, entre as classes definidas [82]. Kaya, Keceli et al. [121] comparam o desempenho de modelos DL treinados de origem e previamente treinados por *transfer learning* na classificação de folhas. Os autores concluíram que através de *transfer learning* o desempenho do modelo é otimizado.

Neste trabalho, o treino foi efetuado através da plataforma TensorFlow, caracterizada pela sua versatilidade. Da base de dados do TensorFlow, foram considerados três modelos SSD: SSD MobileNet v2, SSD Inception v2 e SSD ResNet 50 e ainda o modelo YOLO v4 da plataforma Darknet. Os modelos foram previamente treinados com a base de dados MS COCO. Os modelos SSD treinam por cerca de 50 000 *epochs* e acontece um processo de teste a cada 50 *epochs*. Estas sessões de teste permitem monitorizar a evolução do treino. O modelo YOLO v4 apresenta uma aprendizagem mais expedita, necessitando de menos *epochs*.

Nenhum dos modelos considerados é capaz de processar as imagens no seu tamanho original. Cada modelo diminui o tamanho das imagens das amostras de treino e validação num tamanho específico. Assim, foram treinados cinco modelos: i) SSD MobileNet v2 – 300 × 300 px., ii) SSD MobileNet v2 – 640 × 640 px., iii) SSD Inception v2 – 300 × 300 px., iv) SSD ResNet 50 – 640 × 640 px. e v) YOLO v4 – 416 × 416 px.

Os programas executáveis (*scripts*) necessários para treino e inferência utilizam TensorFlow r1.15.0 e são executados através do Google Colaboratory (Colab)¹⁷, com elevada capacidade de processamento. Embora seja variável ao longo das sessões, normalmente foram atribuídos NVIDIA Tesla T4 ou Tesla K80 com VRAM de 12 GB e um poder computacional entre 3,7 e 7,5.

A Figura 13 resume o processo de treino a que foram submetidos os modelos SSD e YOLO em estudo.

¹⁷ Google Colaboratory (<http://colab.research.google.com>). Último acesso em: 15/06/2021.

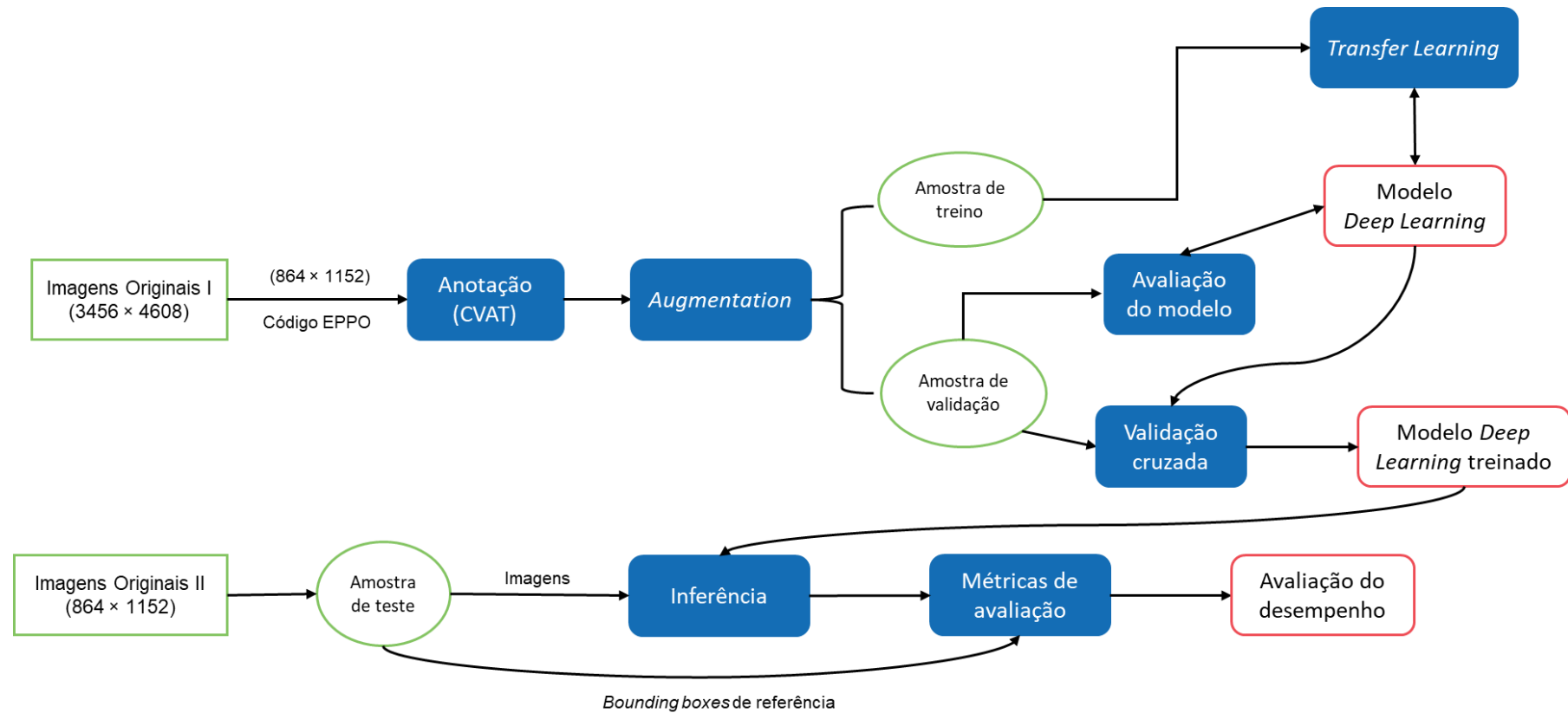


Figura 13 – Processo de treino e avaliação do desempenho dos modelos SSD e YOLO em estudo. Adaptado de: [122].

4.2.4. Avaliação do desempenho dos modelos *Deep Learning*

Como anteriormente referido, a inferência é realizada através do Colab. Deste processo resultam as previsões (realizadas pelo modelo), relacionadas com a ocorrência de plantas nas imagens sujeitas à inferência. É possível visualizar os resultados de duas formas: i) *bounding boxes* previstas nas imagens e ii) anotações das previsões correspondentes (nos formatos Pascal VOC ou YOLO).

Cada previsão realizada pelo modelo é representada pela *bounding box* e confiança, geralmente um valor entre 0 e 1, que mostra o grau de certeza associado à previsão. Para a avaliação do desempenho do modelo de deteção, são calculadas diferentes métricas. O que se pretende medir é a correspondência entre as *bounding boxes* previstas e as de referência [84].

Consideremos um objeto a ser detetado com uma *bounding box* de referência B_r e a deteção representada pela *bounding box* prevista B_p . A correspondência entre *bounding boxes* é avaliada através do parâmetro *Intersection Over Union* (IOU) – o quociente entre a área de interseção e a de união entre B_p e B_r , demonstrado pela Equação (1) e representado na Figura 14.

$$IOU = \frac{\text{área}(B_p \cap B_r)}{\text{área}(B_p \cup B_r)} = \frac{\text{área de interseção}}{\text{área de união}} \quad (1)$$

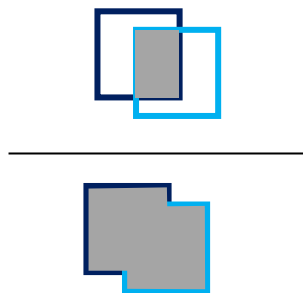


Figura 14 – Representação da métrica *Intersection Over Union*. Adaptado de: [114].

Definindo o limiar de IOU (no intervalo de 0 a 1), dispomos de métricas mais ou menos restritivas na avaliação das previsões. Caso se aproxime de 1 exige previsões praticamente perfeitas, por sua vez, quando o limiar se aproxima de 0 é mais flexível.

As métricas *precision* e *recall* permitem avaliar o modelo através da relação entre referências e previsões. Para calcular estas métricas cada previsão é classificada como:

- Verdadeiro Positivo (VP): previsão correta da referência correspondente;
 - Falso Positivo (FP): previsão incorreta de um objeto não referenciado ou previsão mal posicionada de um objeto referenciado;
- ou
- Falso Negativo (FN): referência não detetada.

A *precision* (Pr) representa a capacidade do modelo identificar apenas os objetos relevantes, ou seja, a percentagem de VP face ao total de previsões. Por sua vez, o *recall* (Rc) refere a capacidade do modelo encontrar todos os casos relevantes, ou seja, a percentagem de VP face ao total de referências.

Por exemplo, uma imagem dispõe de R referências e um modelo que produz D previsões, das quais C são corretas (com $C \leq R$), a *precision* é expressa através da Equação (2) e o *recall* pela Equação (3). São consideradas como corretas (Verdadeiros Positivos) todas as deteções com IOU superior ao limiar definido [114]. Tipicamente, aplica-se IOU igual a 0,5.

$$\text{precision (Pr)} = \frac{\sum_{n=1}^C VP_n}{\sum_{n=1}^C VP_n + \sum_{n=1}^{D-C} FP_n} = \frac{\sum_{n=1}^C VP_n}{\text{previsões}} \quad (2)$$

$$\text{recall (Rc)} = \frac{\sum_{n=1}^C VP_n}{\sum_{n=1}^C VP_n + \sum_{n=1}^{R-C} FN_n} = \frac{\sum_{n=1}^C VP_n}{\text{referências}} \quad (3)$$

Como referido anteriormente, um parâmetro sempre associado às previsões geradas pelo modelo é a confiança. Indica o nível de certeza do modelo face a determinada previsão. Este parâmetro ao ser considerado no cálculo das métricas *precision* e *recall*, permite que apenas as previsões com confiança superior à considerada sejam categorizadas como positivas. Assim, para cada valor de confiança é calculado os valores correspondentes de *precision* e *recall*.

De acordo com Padilla, Passos et al. [114], um modelo de deteção tem um bom desempenho quando, diminuindo a confiança, a *precision* mantém-se elevada e o *recall*

umenta. Ou seja, ao representar graficamente a relação entre *precision* e *recall*, uma elevada área sob a curva deverá indicar tanto elevada *precision* quanto *recall*.

Percebe-se pela Equação (3) que o *recall* tem tendência a diminuir com a evolução da confiança (devido ao decréscimo de VP), o mesmo não se verifica com a *precision*, com uma relação variável. Daí, a curva *precision* × *recall* apresenta um comportamento sinuoso, o que dificulta o cálculo da área sob a curva [114]. Para colmatar esta limitação, é calculada a métrica *Average Precision* (AP).

Para o cálculo da AP é necessária a ordenação do conjunto de confianças referentes às previsões do modelo ($\tau(k)$), a ordenação dos valores de *recall* ($R_r(k)$) e a interpolação dos pares *precision* × *recall* ($P_{r_{interp}}(R)$). Ao considerar o VOC Challenge [123] é aplicada a interpolação a todos os pontos e o parâmetro $R_r(k)$ é calculado de acordo com as Equações (4), (5) e (6):

$$R_r(0) = 1 \quad (4)$$

$$R_r(k) = R_c(\tau(k)), k = 1, 2, \dots, K \quad (5)$$

$$R_r(K + 1) = 0 \quad (6)$$

Assim, a *Average Precision* é calculada através da Equação (7):

$$AP = \sum_{k=0}^K (R_r(k) - R_r(k+1)) P_{r_{interp}}(R_r(k)) \quad (7)$$

Esta métrica é individual para cada classe definida para o treino do modelo. Ao dispor de uma base de dados com diversas classes, é calculada a *Mean Average Precision* (mAP), a média da *Average Precision* de todas as classes Equação (8).

$$mAP = \frac{1}{C} \sum_{i=1}^C AP_i, \quad (8)$$

i representa o índice da classe e C o número total de classes em estudo.

Otimizar a confiança é essencial para aprimorar o desempenho do modelo. Isto consegue-se através do cálculo da métrica F1, um valor de 0 a 1, resultado da primeira média harmónica entre a *precision* e o *recall*, como mostra Equação (9). Assim, pela combinação entre a *precision* e o *recall* compreende-se o comportamento do modelo. Ao assumir a confiança com maior valor de F1, o desempenho do modelo é otimizado no valor com melhor relação entre VP, FP e FN [122].

$$F1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (9)$$

Uma vez que a maioria das culturas apresentam mais que duas classes, é necessário recorrer a métodos de avaliação mais robustos e capazes de transmitir o desempenho do modelo para a classificação de plantas de acordo com a fase do ciclo cultural. Para tal, procedemos à construção de matrizes de confusão.

A matriz é construída tendo em conta duas dimensões principais – classe prevista e classe de referência. Através da comparação da classe prevista com a classe de referência, é considerada a previsão como correta ou incorreta [79,124]. À semelhança do referido anteriormente, cada previsão é classificada como:

- Verdadeiro Positivo (VP): classificação prevista de acordo com a classe de referência;
- Falso Positivo (FP): atribuição da classe errada a objetos referenciados ou classificação de objetos não referenciados;
- Falso Negativo (FN): referência não classificada.

O conjunto de classificações previstas que não englobam a classe em estudo são consideradas como Verdadeiro Negativo (VN).

Através desta informação é possível calcular as seguintes métricas para cada uma das classes [124]:

- *Exatidão*:

$$\text{exatidão} = \frac{VP + VN}{VP + FP + FN + VN} \quad (10)$$

- *Precision*:

$$\text{precision} = \frac{VP}{VP + FP} \quad (11)$$

- *Recall*:

$$\text{recall} = \frac{VP}{TP + FN} \quad (12)$$

- *F1*:

$$F1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (13)$$

Para comparar o desempenho dos modelos é importante considerar a exatidão média face ao conjunto das classes em estudo. Contudo, realizar a média simples dos valores de exatidão poderá não transmitir devidamente o valor dessa métrica. Em particular, quando a distribuição de objetos pelas classes consideradas é pouco balanceada. Nestes casos é sugerido por Brodersen, Ong et al. [125] o cálculo da métrica Exatidão Balanceada (EB), como indica a Equação (14):

$$EB = \frac{\sum_{i=1}^C \frac{VP_i}{R_i}}{C} \quad (14)$$

Em que i representa o índice da classe num total de C classes. A classe i apresenta VP_i previsões corretas, num total de R_i referências.

A Tabela 9 resume os parâmetros e as métricas calculadas para a avaliação do desempenho dos modelos SSD e YOLO treinados com as imagens do repositório do PixelCropRobot com o objetivo de detetar e classificar oito culturas hortícolas. A avaliação do desempenho foi realizada com o apoio dos repositórios *Object Detection Metrics*¹⁸ e *Review_object_detection_metrics*¹⁹.

¹⁸ Object Detection Metrics (<https://gitlab.inescotec.pt/agrob/Object-Detection-Metrics>). Último acesso em: 22/08/2021.

¹⁹ Review_object_detection_metrics (https://github.com/rafaelpadilla/review_object_detection_metrics). Último acesso em: 22/08/2021.

Tabela 9 – Parâmetros e métricas calculadas para a avaliação do desempenho dos modelos de detecção em estudo nas amostras de validação e teste considerando os problemas de detecção e classificação.

Parâmetro/Métrica	Detecção		Classificação
	Am. de validação	Am. de teste	Am. de teste
F1	X	X	X
VP, FP, FN	X		
<i>Precision</i>		X	X
<i>Recall</i>		X	X
<i>Average Precision</i>		X	
<i>Precision X Recall</i>		X	
VP, FP, FN, VN (Matriz de confusão)			X
Exatidão			X
Exatidão Balanceada			X

Am.: amostra

4.3. Técnicas de fotogrametria para reconstrução tridimensional de plantas

Para além da construção do repositório com imagens bidimensionais (2D), explorámos diversas técnicas de fotogrametria para a reconstrução tridimensional (3D), tanto de plantas individuais como de grupos de plantas.

O processo de reconstrução tridimensional inicia-se pela captação das imagens. Através do mesmo *smartphone* utilizado anteriormente (*Huawei Mate 10*) são gravados diversos vídeos, em alta definição total (*full HD*, 1920 × 1080), tanto a plantas individuais como a grupos de plantas.

No caso de plantas individuais a gravação do vídeo acontece como mostra a Figura 15. Foram selecionadas duas culturas com características bastantes distintas: rabanete em vaso, com a raiz proeminente e diversas folhas (Figura 15a) e coentro, numa fase avançada do ciclo vegetativo, onde se notam as inflorescências (Figura 15b).



Figura 15 – Método para gravação de vídeo para reconstrução tridimensional de plantas individuais. a) rabanete, b) coentro.

No primeiro caso a planta permaneceu numa posição fixa e ao deslocar o *smartphone* ao longo da sua estrutura foi possível captar todas as perspetivas da planta e a envolvente. No caso da planta de coentro, o *smartphone* foi colocado num suporte fixo e a planta numa plataforma giratória, permitindo assim obter informação de todas as

perspetivas da planta. Na Figura 16 é apresentado o método de gravação de vídeo para a reconstrução dos grupos de plantas.

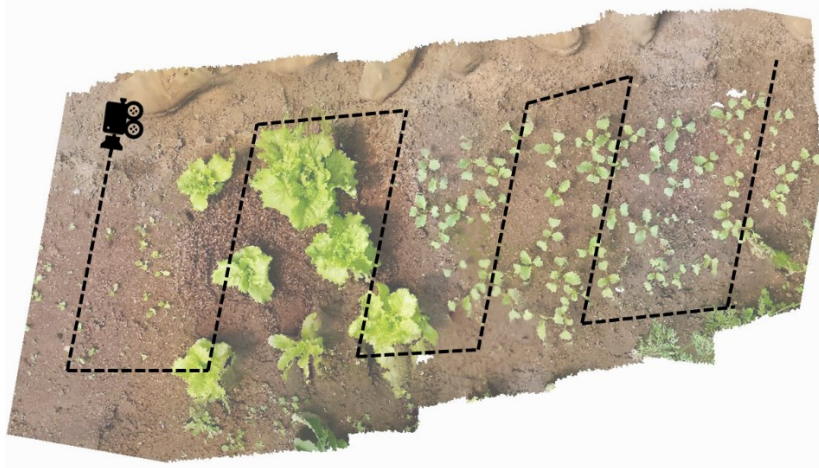


Figura 16 – Método para gravação de vídeo para a reconstrução tridimensional do grupo de plantas.

Em ambos os casos são extraídos os *frames* (imagens digitais fixas de um produto audiovisual, neste caso um vídeo) do vídeo correspondente, através do seguinte programa executável em linguagem *Python*:

```
1 import cv2
2 cap=cv2.VideoCapture('videopath.mp4')
3 i=0
4 while(cap.isOpened()):
5     ret, frame = cap.read()
6     if ret==False:
7         break
8     cv2.imwrite('frame'+str(i)+'.jpg',frame)
9     i+=1
10 cap.release()
11 cv2.destroyAllWindows()
```

Dispomos assim de um conjunto de imagens sequenciais, o processamento é distinto para cada um dos exemplos considerados e é descrito em seguida.

4.3.1. Planta individual

Numa primeira fase, propomos a reconstrução de uma planta de rabanete em vaso (Figura 17). Da extração dos *frames* do vídeo gravado foram obtidas 486 imagens.



Figura 17 – Exemplo da imagem de rabanete proposto para reconstrução tridimensional.

Recorreu-se ao *software* Agisoft Metashape²⁰ para o cálculo dos valores de SfM e MVS, portanto, a criação das nuvens de pontos. Após a introdução das imagens no *software*, estas são alinhadas tendo em conta a perspetiva de cada uma. Seguidamente, são geradas as nuvens dispersa e densa de pontos, representativas do rabanete. Pela transformação da nuvem densa de pontos numa rede de superfícies, por processos de *Multi-View Environment* (MVE) [102] é possível representar, tridimensionalmente, a planta.

As imagens da planta de rabanete foram submetidas no *software* sem nenhum processamento. Contudo, nem sempre todos os pontos são interessantes para incluir na reconstrução porque, por exemplo, não pertencem à estrutura da planta (p. ex.: solo, outras plantas ou outros objetos). Estes pontos influenciam negativamente o

²⁰ Agisoft Metashape, versão 1.7.4 (<https://www.agisoft.com/>). Último acesso em: 25/08/2021.

desempenho dos processos aplicados e, inclusive, a visualização correta dos resultados gerados pelo que são retirados da análise

Para ultrapassar este inconveniente e testar a versatilidade do *software* quando aplicado a situações mais complexas, propomos a reconstrução de uma planta de coentro (Figura 18). Como mencionado anteriormente, para a gravação do vídeo, o *smartphone* foi colocado num suporte fixo e a planta numa plataforma giratória. Foram selecionados 74 *frames* do total obtido.



Figura 18 – Imagem da planta de coentro antes do processamento no PlantCV.

Através dos métodos descritos na plataforma *PlantCV*²¹, é possível distinguir os pontos que pertencem à estrutura do coentro, do fundo e de outros objetos. Assim, definimos a estrutura da planta como região de interesse.

O processamento inicia-se com a conversão das imagens originais (RGB) no formato tons de cinzento (*grayscale*) tendo em conta o sistema colorimétrico HSV – definido pela tonalidade (H), saturação (S) e brilho (V) de cada *pixel* da imagem. É também considerado o sistema colorimétrico LAB, adaptação do sistema colorimétrico CIELAB, que transmite a luminosidade (L) e as cores primárias – verde e magenta (A), azul e amarelo (B) – de cada *pixel* da imagem. Com estas transformações compreende-se o contraste entre a planta, o fundo e outros objetos (que serão mais escuros ou mais

²¹ PlantCV - *Plant phenotyping using computer vision* (<https://plantcv.danforthcenter.org/>). Último acesso em: 14/06/2021.

claros que a planta). Pela visualização da conversão realizada para cada canal dos dois sistemas colorimétricos apresentados, seleciona-se aquele capaz de transmitir de forma mais consistente os pontos referentes à planta. Normalmente, a escolha recai sobre o canal A do sistema colorimétrico LAB (aquele que transmite a presença das cores verde e magenta).

Após esta seleção é imposto à imagem binária (em escala de cinzentos) um limiar colorimétrico, ou seja, pela definição de um determinado valor (entre 0 e 255, considerando 255 como branco) são extraídos da imagem os pixels de interesse (que assumem um valor acima do definido no limiar). Após este passo são aplicados processos de preenchimento ou dilatação com o objetivo de aprimorar a imagem.

Para a execução do processamento recorreu-se ao *Jupyter Notebook*²², uma ferramenta *online*, de acesso livre que permite a execução e partilha de código e muito mais. O programa executável aplicado às imagens de coentro foi adaptado da plataforma PlantCV²³ da seguinte forma:

```
1 from plantcv import plantcv as pcv
2 import matplotlib
3 class options:
4     def __init__(self):
5         self.image = "imagepath"
6         self.debug = "print"
7         self.writeimg= False
8         self.result = "results.json"
9         self.outdir = "./"
10 args = options()
11 pcv.params.debug = args.debug
12 img, path, filename = pcv.readimage(filename=args.image)
13 cs_plot=pcv.visualize.colorsspaces(rgb_img=img, original_img=False)
14 a=pcv.rgb2gray_lab(rgb_img=img, channel='a')
15 a_thresh=pcv.threshold.binary(gray_img=a, threshold=125, max_value=255,
16 object_type='dark')
17 a_fill = pcv.fill(bin_img=a_thresh, size=120)
18 a_fill_again = pcv.dilate(gray_img=a_fill, ksize=3, i=3)
19 a_fill_onemoretime = pcv.fill(bin_img=a_fill_again, size=500)
20 id_objects, obj_hierarchy=pcv.find_objects(img=img, mask=a_fill_onemoretime)
21 roi_custom, roi_hier_custom = pcv.roi.custom(img=img, vertices=[])
22 roi_objects, hierarchy, kept_mask, obj_area = pcv.roi_objects(img=img,
23 roi_contour=roi_custom, roi_hierarchy=roi_hier_custom, object_contour=id_objects,
24 obj_hierarchy=obj_hierarchy, roi_type='cutto')
25 masked2 = pcv.apply_mask(img=img, mask=kept_mask, mask_color='white')
```

²² Jupyter Notebook (<https://jupyter.org/>). Último acesso em: 15/08/2021.

²³ Analyze Color, PlantCV (https://plantcv.readthedocs.io/en/latest/analyze_color/). Último acesso em: 14/08/2021.

Após a definição da região de interesse de todas as imagens, estas são carregadas no *software* Agisoft Metashape e a criação das nuvens de pontos e do modelo tridimensional acontece de forma semelhante ao descrito anteriormente para a planta de rabanete.

4.3.2. Grupo de plantas

Para a reconstrução de vários grupos de plantas foi novamente aplicado o *software* Agisoft Metashape e o *software* OpenDroneMap (ODM)²⁴, um conjunto de ferramentas, de acesso livre²⁵, para processamento de imagens (tipicamente aéreas, de drone) em ortofotomapas e modelos tridimensionais. Os principais resultados gerados por esta ferramenta são: i) rede de superfícies tridimensional, ii) nuvem de pontos e iii) ortofotomapa. Os diferentes ficheiros gerados podem ser acedidos por ferramentas como: QGIS²⁶, CloudCompare²⁷ e MeshLab²⁸. O *software* foi executado através do Docker²⁹, com o seguinte comando:

```
docker run -ti --rm -v c:/Users/youruser/datasets:/datasets opendronemap/odm
--project-path /datasets project
```

Todas as ferramentas utilizadas são de acesso livre e de instalação expedita, sem necessidade de elevados conhecimentos em Programação ou outras áreas tecnológicas. As imagens e os programas executáveis são incluídos no repositório do PixelCropRobot³⁰.

²⁴ ODM (<https://www.opendronemap.org/odm/>). Último acesso em: 18/07/2021.

²⁵ OpenDroneMap Authors (<https://github.com/OpenDroneMap/ODM>). Último acesso em: 18/07/2021.

²⁶ QGIS (<https://www.qgis.org/en/site/>). Último acesso em: 30/08/2021.

²⁷ CloudCompare (<https://www.cloudcompare.org/>). Último acesso em: 28/08/2021.

²⁸ MeshLab (<https://www.meshlab.net/>). Último acesso em: 30/08/2021.

²⁹ Docker (<https://www.docker.com/>). Último acesso em: 18/07/2021.

³⁰ Rodrigues, Terra et al. PixelCropRobot dataset: images of vegetables crops in different phenological stages taken in greenhouses. doi:<https://doi.org/10.5281/zenodo.5574515>.

5. Resultados

5.1. Desempenho dos modelos SSD e YOLO

Serão apresentados os principais resultados da avaliação realizada aos modelos de deteção, com o objetivo de quantificar o seu desempenho. Estes resultados referem-se aos dois problemas em estudo: deteção (Figura 19a) e classificação fenológica de plantas (Figura 19c). A Figura 19b apresenta a anotação de acordo com a classe fenológica correspondente.

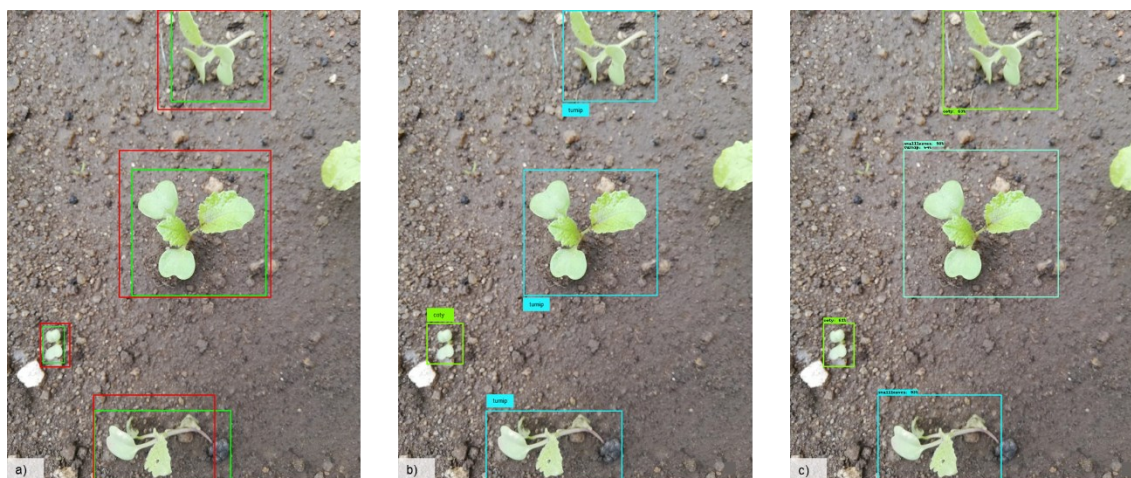


Figura 19 – Exemplificação do processo de deteção e classificação de plantas: a) deteção, verde: referências, vermelho: previsões do modelo. b) anotações de referência, c) classificação, cores diferentes significam classes diferentes.

5.1.1. Alface (LACSA)

Nesta secção são apresentados totais correspondentes aos modelos SSD e YOLO treinados para a alface. Remetemos para o Anexo II para consulta dos resultados relativos às restantes culturas (acelga, cenoura, coentro, espinafre, nabo, rabanete e rúcula).

Deteção

A análise do desempenho dos modelos DL inicia-se com a deteção, ou seja, pretende-se aferir a capacidade de cada modelo em detetar corretamente as plantas de alface. Como já referido anteriormente, através da otimização do F1 calculado na amostra de validação, é possível aprimorar o desempenho do modelo. A Tabela 10 refere a confiança com maior valor de F1 para cada modelo em estudo.

Tabela 10 – Confiança que otimiza o F1. Parâmetro calculado na amostra de validação sem *augmentation*.

Modelo*	Confiança $\geq$	F1
SSD MobileNet v2 (300 × 300)	92%	84%
SSD MobileNet v2 (640 × 640)	49%	86%
SSD Inception v2 (300 × 300)	44%	85%
SSD ResNet 50 (640 × 640)	51%	86%
YOLO v4 (416 × 416)	54%	85%

*a amostra de validação é constituída por 471 imagens e 1380 anotações de referência (Tabela 8).

De acordo com a Tabela 10, todos os modelos apresentam F1 superior a 80%, sendo o modelo SSD MobileNet v2 (640 × 640) aquele com valor superior (próximo de 86%). A Figura 20 apresenta a evolução do F1 com o aumento da confiança.

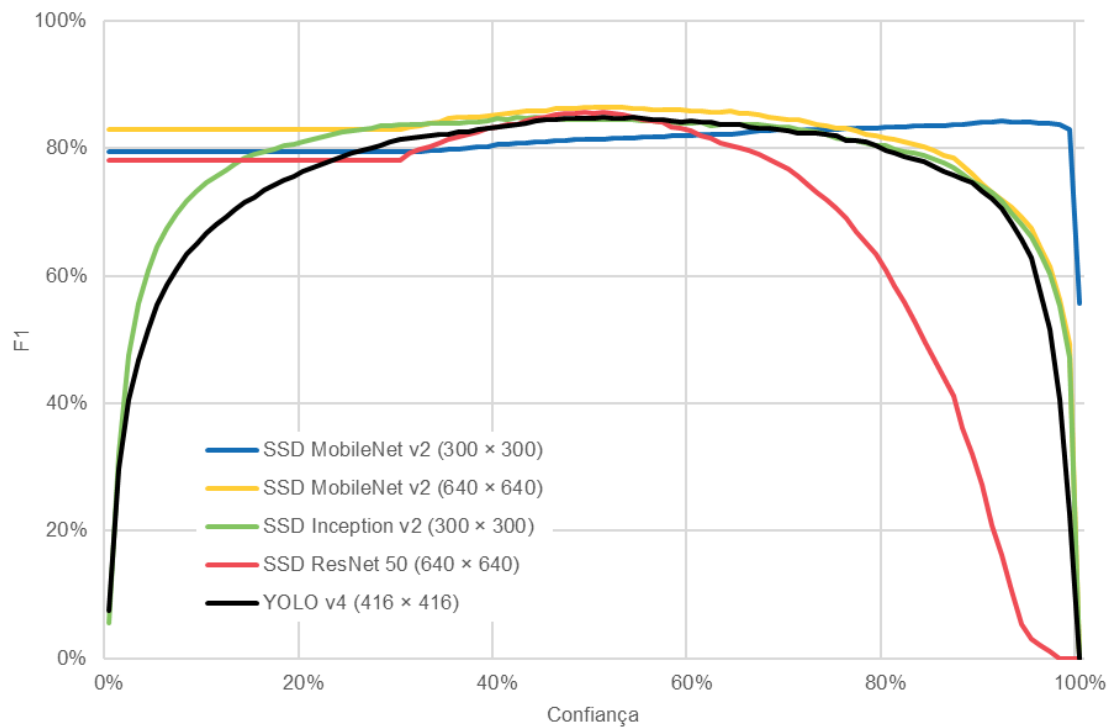


Figura 20 – Evolução do F1 com o aumento da confiança, na amostra de validação sem *augmentation*.

Tipicamente, modelos que apresentam curvas mais achatadas realizam previsões com maior confiança e baixos níveis de falsas deteções [122]. Neste caso, comparativamente com os restantes, ambos os modelos SSD MobileNet v2 e o modelo SSD ResNet 50 apresentam curvas mais achatadas, o que significa um baixo número

de falsas deteções. Este facto pode ser confirmado pela evolução de VP, FP e FN com o aumento da confiança, tal como apresentado na Figura 21.

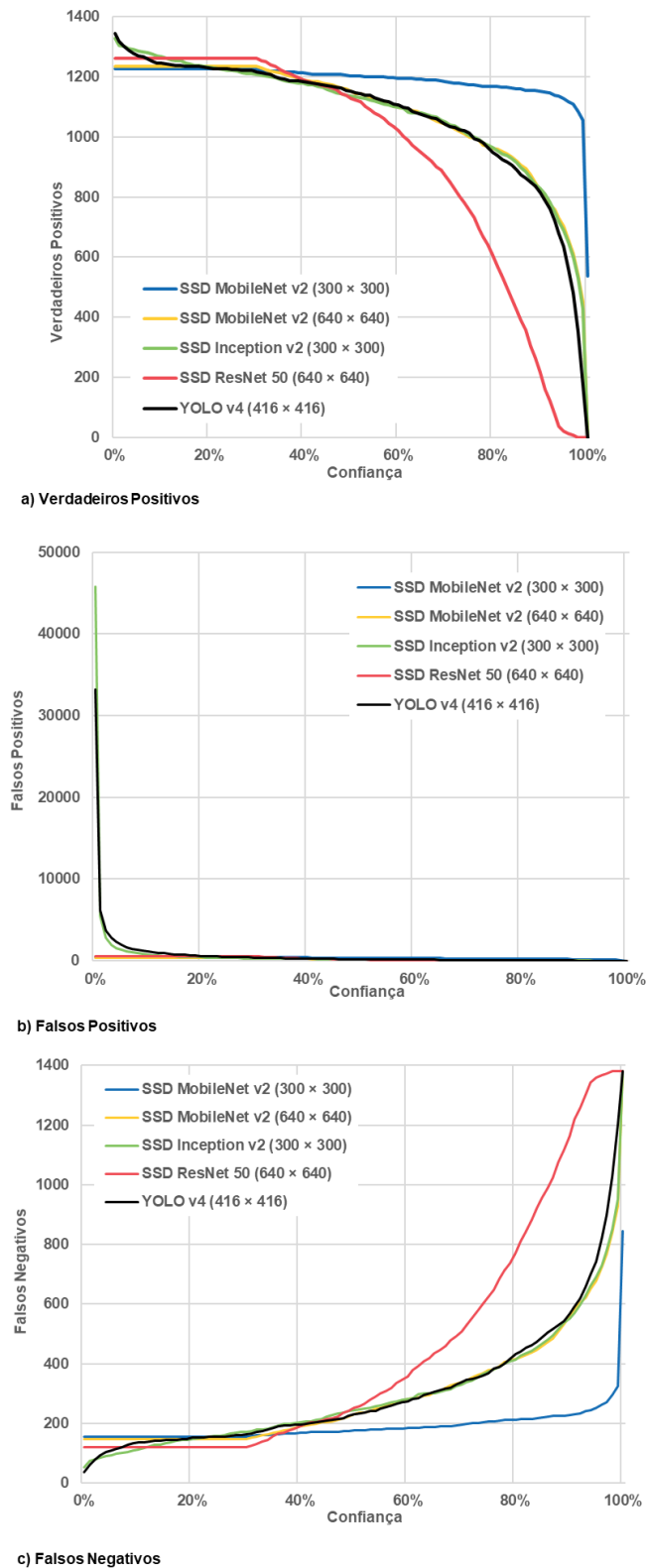


Figura 21 – Evolução do número de: a) Verdadeiros Positivos, b) Falsos Positivos e c) Falsos Negativos com o aumento da confiança, na amostra de validação sem *augmentation*.

Os modelos SSD Inception v2 e YOLO v4 produzem um número elevado de falsas detecções, que rapidamente diminui com a evolução da confiança (Figura 21). Por sua vez, os modelos SSD MobileNet v2 (300 × 300), SSD MobileNet v2 (640 × 640) e SSD ResNet 50 apresentam um valor de Falsos Positivos sempre inferior a 480, 363 e 584, respetivamente. Na confiança otimizada (valores da Tabela 10), todos os modelos são capazes de detetar corretamente cerca de 80% do total de referências (1380) da amostra de validação:

- SSD MobileNet v2 (300 × 300) – 1146;
- SSD MobileNet v2 (640 × 640) – 1159;
- SSD Inception v2 – 1165;
- SSD ResNet 50 – 1123;
- YOLO v4 – 1138.

A capacidade de generalização dos modelos treinados é avaliada através da amostra de teste, um conjunto de dados independente dos restantes, nunca aplicado ao modelo. Na Tabela 11 são apresentados os resultados das principais métricas definidas, considerando todas as previsões transmitidas pelo modelo (confiança igual a 0%) e as correspondentes à confiança otimizada.

Tabela 11 – Resultados de várias métricas para os modelos em estudo, considerando todas as detecções e a confiança otimizada, na amostra de teste sem *augmentation*.

Modelo*	Confiança ≥	AP	<i>Precision</i>	<i>Recall</i>	F1
SSD MobileNet v2 (300 × 300)	0%	84%	74%	88%	80%
SSD MobileNet v2 (640 × 640)	0%	86%	75%	88%	81%
SSD Inception v2 (300 × 300)	0%	85%	3%	94%	6%
SSD ResNet 50 (640 × 640)	0%	85%	69%	88%	77%
YOLO v4 (416 × 416)	0%	86%	4%	97%	7%
SSD MobileNet v2 (300 × 300)	92%	81%	87%	84%	85%
SSD MobileNet v2 (640 × 640)	49%	81%	88%	82%	85%
SSD Inception v2 (300 × 300)	44%	77%	85%	81%	83%
SSD ResNet 50 (640 × 640)	51%	73%	90%	74%	82%
YOLO v4 (416 × 416)	54%	78%	86%	81%	83%

AP – *Average Precision*

*A amostra de teste é constituída por 290 imagens e 905 anotações de referência

À exceção dos modelos SSD Inception v2 e YOLO v4 que apresentam uma *precision* próxima de 0% com um valor elevado de *recall* (94% e 97%, respetivamente), os restantes modelos assumem um bom compromisso entre *precision* e *recall*. Para confirmar este facto, a Figura 22 apresenta a curva *precision* × *recall* para a confiança 0% (Figura 22a) e para a confiança otimizada (Figura 22b).

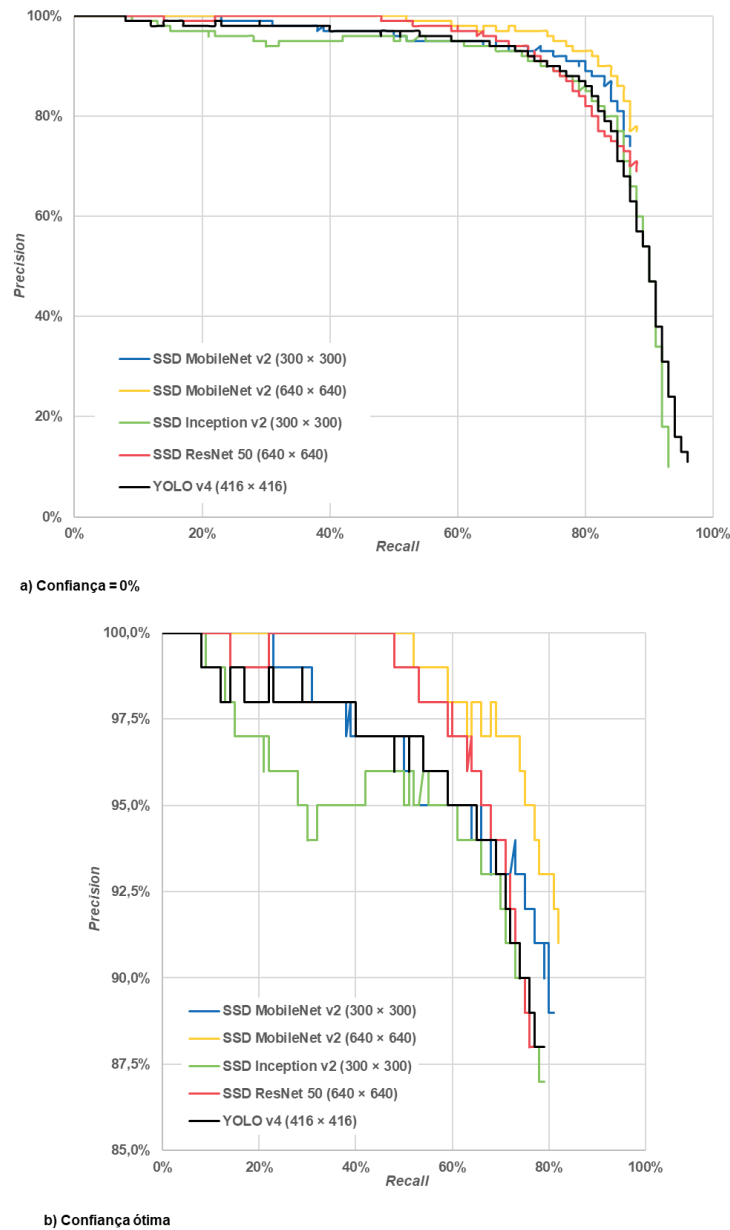


Figura 22 – Curvas *precision* × *recall* para a: a) confiança de 0% e b) confiança otimizada. Avaliação executada na amostra de teste sem *augmentation*.

Pela observação da Figura 22 e, considerando os resultados das métricas calculadas (Tabela 11), o modelo SSD MobileNet v2 é aquele com melhor desempenho na deteção de alface, apresentando valores de AP próximos de 81% (80,89% - SSD MobileNet v2 (300 × 300) e 80,86% - SSD MobileNet v2 (640 × 640)).

Classificação

Foi avaliada a capacidade de cada modelo classificar as plantas de alface detetadas de acordo com as classes fenológicas definidas (Tabela 7). Os resultados obtidos referem-se à confiança otimizada (Tabela 10) e podem ser observados na Tabela 12.

Tabela 12 – Matriz de confusão para os modelos SSD e YOLO em estudo na amostra de teste sem *augmentation*.

Conf. = 92%			SSD MobileNet v2 (300 × 300)				
Deteção			Referência			Não referenciada	
		<i>coty</i>	<i>minus9</i>	<i>plus9</i>	<i>ready</i>	solo	infestante
	<i>coty</i>	118	9	0	0	1	1
	<i>minus9</i>	1	529	35	30	1	14
	<i>plus9</i>	0	3	54	46	0	0
	<i>ready</i>	0	0	0	24	0	0
Conf. = 49%			SSD MobileNet v2 (640 × 640)				
Deteção			Referência			Não referenciada	
		<i>coty</i>	<i>minus9</i>	<i>plus9</i>	<i>ready</i>	solo	infestante
	<i>coty</i>	53	0	0	0	1	3
	<i>minus9</i>	17	519	7	2	3	27
	<i>plus9</i>	0	2	68	4	0	0
	<i>ready</i>	0	0	15	116	0	0
Conf. = 44%			SSD Inception v2 (300 × 300)				
Deteção			Referência			Não referenciada	
		<i>coty</i>	<i>minus9</i>	<i>plus9</i>	<i>ready</i>	solo	infestante
	<i>coty</i>	123	21	0	0	1	5
	<i>minus9</i>	0	476	14	2	0	11
	<i>plus9</i>	0	24	66	23	0	1
	<i>ready</i>	0	1	28	70	0	0
Conf. = 51%			SSD ResNet 50 (640 × 640)				
Deteção			Referência			Não referenciada	
		<i>coty</i>	<i>minus9</i>	<i>plus9</i>	<i>ready</i>	solo	infestante
	<i>coty</i>	61	13	0	0	0	0
	<i>minus9</i>	0	522	10	2	0	14
	<i>plus9</i>	0	9	57	26	0	0
	<i>ready</i>	0	0	7	28	0	0
Conf. = 65%			YOLO v4 (416 × 416)				
Deteção			Referência			Não referenciada	
		<i>coty</i>	<i>minus9</i>	<i>plus9</i>	<i>ready</i>	solo	infestante
	<i>coty</i>	138	44	0	0	4	9
	<i>minus9</i>	0	500	20	6	0	10
	<i>plus9</i>	0	6	11	2	0	0
	<i>ready</i>	0	0	25	30	0	0

Conf.: confiança que otimiza o valor de F1 na amostra de validação (Tabela 10).

Os modelos SSD MobileNet v2 (300 × 300) e SSD Inception v2 são aqueles que realizam um maior número de previsões (866 em ambos os casos). Contudo, o modelo SSD MobileNet v2 (640 × 640) apresenta uma melhor relação entre previsões corretas e total de previsões (cerca de 90% das 837 previsões realizadas são VP).

Para avaliar o desempenho de cada modelo na classificação de alface são calculadas as métricas exatidão, *precision*, *recall* e F1 (Tabela 13), relacionando os resultados transmitidos pela Tabela 12 e o total de referências de cada classe na amostra de teste.

Tabela 13 – Resultados das métricas exatidão, *precision* e *recall* para cada classe de alface para os modelos SSD e YOLO em estudo. Avaliação na amostra de teste sem *augmentation*.

<i>coty</i> n = 144	Exatidão	<i>Precision</i>	<i>Recall</i>	F1
SSD MobileNet v2 (300 × 300)	96%	91%	82%	86%
SSD MobileNet v2 (640 × 640)	90%	93%	37%	53%
SSD Inception v2 (300 × 300)	95%	82%	85%	84%
SSD ResNet 50 (640 × 640)	88%	82%	42%	56%
YOLO v4 (416 × 416)	92%	71%	96%	81%
<i>minus9</i> n = 569				
SSD MobileNet v2 (300 × 300)	87%	87%	93%	90%
SSD MobileNet v2 (640 × 640)	88%	90%	91%	91%
SSD Inception v2 (300 × 300)	87%	95%	84%	89%
SSD ResNet 50 (640 × 640)	91%	95%	92%	93%
YOLO v4 (416 × 416)	88%	93%	88%	90%
<i>plus9</i> n = 81				
SSD MobileNet v2 (300 × 300)	91%	52%	67%	59%
SSD MobileNet v2 (640 × 640)	98%	92%	84%	88%
SSD Inception v2 (300 × 300)	92%	58%	81%	68%
SSD ResNet 50 (640 × 640)	92%	62%	70%	66%
YOLO v4 (416 × 416)	91%	58%	14%	22%
<i>ready</i> n = 111				
SSD MobileNet v2 (300 × 300)	90%	100%	22%	36%
SSD MobileNet v2 (640 × 640)	98%	89%	100%	94%
SSD Inception v2 (300 × 300)	92%	71%	63%	67%
SSD ResNet 50 (640 × 640)	89%	80%	25%	38%
YOLO v4 (416 × 416)	88%	55%	27%	36%

n: número de referências na amostra de teste.

Os dados da Tabela 13 sugerem que ambos os modelos SSD MobileNet v2 desempenham melhor a classificação de alface, apresentando valores superiores de F1. Na maioria dos casos a classificação das plantas é feita de forma correta – elevada *precision* e são classificadas a maioria das plantas presentes nas imagens da amostra de teste – elevado *recall*. Para além disso, o modelo SSD MobileNet v2 (640 × 640) apresenta um valor superior de EB, próximo de 79%.

Imagens em escala de cinzentos (grayscale)

Como referido anteriormente, para avaliar a versatilidade dos modelos em estudo considerando imagens com diferentes características, as imagens de alface (RGB) foram convertidas no formato *grayscale*. Procedeu-se novamente à avaliação dos modelos, desta vez submetendo a amostra de teste com as imagens em escala de cinzentos. Os resultados desta avaliação estão resumidos na Tabela 14.

Tabela 14 – *Average Precision* e *Mean Average Precision* dos modelos em estudo. Avaliação executada na amostra de teste sem *augmentation*. Representação das imagens em escala de cinzentos. Resultados para a confiança 0%.

Modelo	<i>Average Precision</i>				mAP
	<i>coty</i>	<i>minus9</i>	<i>plus9</i>	<i>ready</i>	
SSD MobileNet v2 (300 × 300)	19%	20%	0%	0%	10%
SSD MobileNet v2 (640 × 640)	11%	49%	0%	0%	17%
SSD Inception v2 (300 × 300)	11%	28%	2%	0%	10%
SSD ResNet 50 (640 × 640)	31%	51%	7%	0%	23%
YOLO v4 (416 × 416)	63%	58%	21%	17%	40%

mAP: *Mean Average Precision*

Tendo em conta os resultados da Tabela 14, é provável que os algoritmos selecionados para a extração de características, o façam através de parâmetros colorimétricos, em particular a cor verde das plantas. Com o objetivo de otimizar estes resultados os modelos SSD e YOLO foram novamente treinados com as mesmas amostras, desta feita com imagens em escala de cinzentos. Os resultados deste processo de treino são apresentados em seguida. Na Tabela 15 é apresentada a confiança com maior valor de F1 para cada modelo em estudo

Tabela 15 – Valor de confiança que otimiza o valor de F1, na amostra de validação sem *augmentation*.

Modelo	Confiança ≥	F1
SSD MobileNet v2 (300 × 300)	39%	88%
SSD MobileNet v2 (640 × 640)	48%	87%
SSD Inception v2 (300 × 300)	49%	92%
SSD ResNet 50 (640 × 640)	53%	95%
YOLO v4 (416 × 416)	72%	92%

De acordo com a Tabela 15, o modelo SSD ResNet 50 (640 × 640) é aquele com maior valor de F1 (cerca de 95%). Para além disso, podemos afirmar que, modelos treinados com imagens *grayscale* apresentam melhores valores de F1 comparativamente àqueles treinados com imagens RGB (Tabela 10). A Figura 23 apresenta a evolução do F1 com o aumento da confiança.

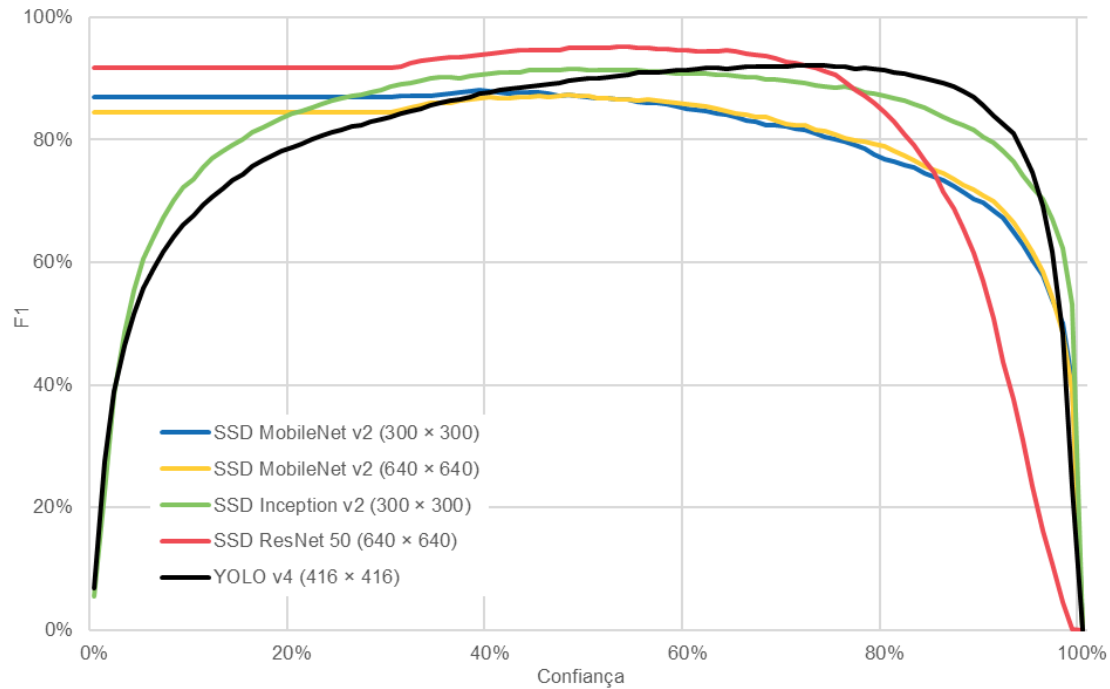


Figura 23 – Evolução do F1 com o aumento da confiança, na amostra de validação sem *augmentation*.

À semelhança do referido para as imagens RGB, também neste caso, os modelos SSD MobileNet v2 e o modelo SSD ResNet 50 apresentam curvas mais achatadas. A Figura 24 apresenta a evolução de VP, FP e FN com o aumento da confiança.

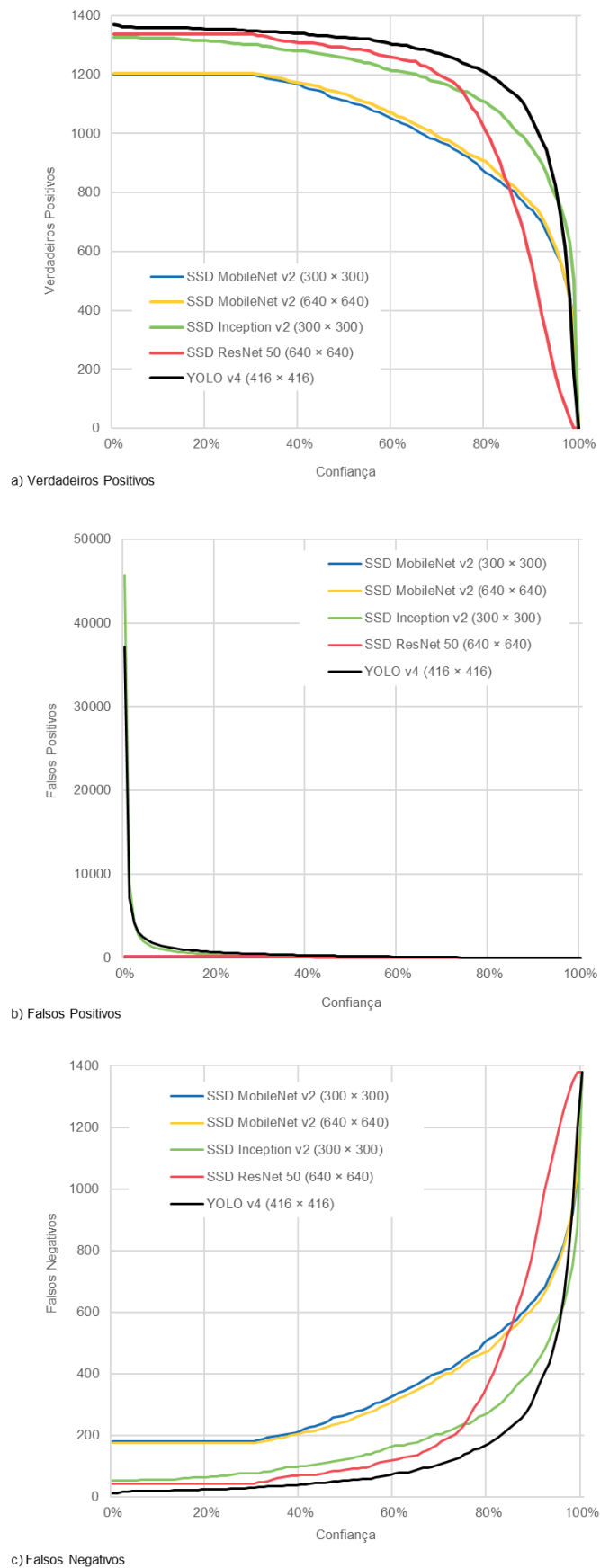


Figura 24 – Evolução do número de: a) Verdadeiros Positivos , b) Falsos Positivos e c) Falsos Negativos com a evolução da confiança, na amostra de validação sem *augmentation*.

Os modelos SSD MobileNet v2 (300 × 300), SSD MobileNet v2 (640 × 640) e SSD ResNet 50 apresentam um valor de FP sempre inferior a 178, 269 e 198, respetivamente. Ao considerar a confiança otimizada os modelos apresentam os seguintes valores de VP, considerando o total de 1380 referências:

- SSD MobileNet v2 (300 × 300) – 1171;
- SSD MobileNet v2 (640 × 640) – 1143;
- SSD Inception v2 – 1259;
- SSD ResNet 50 – 1285;
- YOLO v4 – 1262.

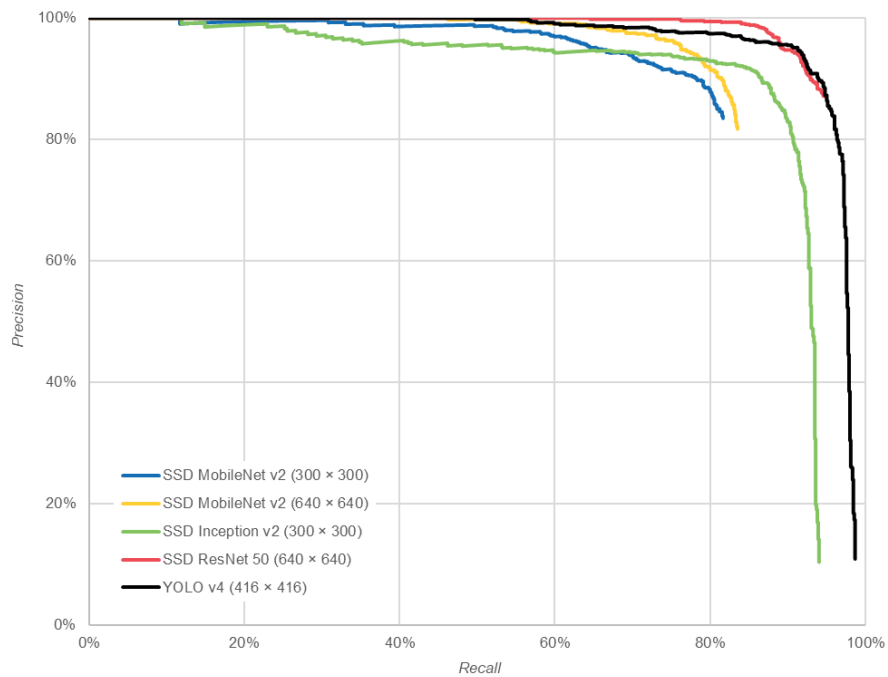
À exceção do modelo SSD MobileNet v2 (640 × 640), todos os modelos transmitem um número de VP superior, quando comparados com os valores relatados anteriormente para o treino com imagens RGB. Na Tabela 16 são apresentados os resultados das principais métricas calculadas na amostra de teste, considerando todas as previsões transmitidas pelo modelo e as correspondentes à confiança otimizada.

Tabela 16 – Resultados das métricas consideradas para os modelos em estudo. Cálculo executado para todas as deteções e na melhor confiança calculada, na amostra de teste sem augmentation.

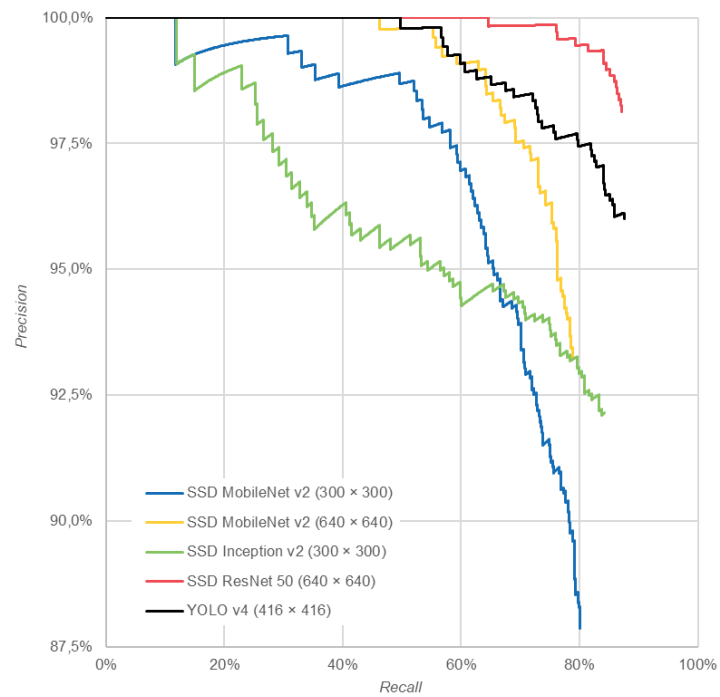
Modelo	Confiança ≥	AP	Precision	Recall	F1
SSD MobileNet v2 (300 × 300)	0%	80%	84%	82%	83%
SSD MobileNet v2 (640 × 640)	0%	82%	82%	84%	83%
SSD Inception v2 (300 × 300)	0%	89%	3%	94%	6%
SSD ResNet 50 (640 × 640)	0%	94%	87%	95%	91%
YOLO v4 (416 × 416)	0%	96%	4%	99%	8%
SSD MobileNet v2 (300 × 300)	39%	78%	88%	80%	84%
SSD MobileNet v2 (640 × 640)	48%	78%	93%	79%	86%
SSD Inception v2 (300 × 300)	49%	85%	88%	88%	88%
SSD ResNet 50 (640 × 640)	53%	88%	98%	88%	92%
YOLO v4 (416 × 416)	72%	91%	95%	91%	93%

AP: Average Precision

Tendo em conta os valores da Tabela 16, na confiança otimizada o modelo YOLO v4 é aquele que apresenta melhor desempenho a detetar alface em *grayscale*, com um valor de AP próximo de 91%. Surgem na Figura 25 as curvas *precision × recall* que ajudam a compreender estes resultados.



a) Confiança = 0%



b) Confiança ótima

Figura 25 – Curvas *precision × recall* para a: a) confiança de 0% e b) confiança ótima. Avaliação executada na amostra de teste sem *augmentation*.

A Tabela 17 apresenta a matriz de confusão para os modelos em estudo. Os resultados obtidos referem-se à confiança otimizada.

Tabela 17 – Matriz de confusão para os modelos SSD e YOLO em estudo para a confiança otimizada na amostra de teste sem *augmentation*.

Conf. = 39%			SSD MobileNet v2 (300 × 300)				
Deteção	Referência					Não referenciada	
		<i>coty</i>	<i>minus9</i>	<i>plus9</i>	<i>ready</i>	solo	infestante
	<i>coty</i>	94	3	0	0	11	2
	<i>minus9</i>	2	491	10	2	2	10
	<i>plus9</i>	0	0	66	0	0	0
	<i>ready</i>	0	0	3	124	0	0
Conf. = 48%			SSD MobileNet v2 (640 × 640)				
Deteção	Referência					Não referenciada	
		<i>coty</i>	<i>minus9</i>	<i>plus9</i>	<i>ready</i>	solo	infestante
	<i>coty</i>	48	2	0	0	0	2
	<i>minus9</i>	1	484	6	0	0	13
	<i>plus9</i>	0	1	69	1	0	1
	<i>ready</i>	0	0	16	110	0	0
Conf. = 49%			SSD Inception v2 (300 × 300)				
Deteção	Referência					Não referenciada	
		<i>coty</i>	<i>minus9</i>	<i>plus9</i>	<i>ready</i>	solo	infestante
	<i>coty</i>	122	0	0	0	0	5
	<i>minus9</i>	4	541	1	0	11	7
	<i>plus9</i>	0	2	71	1	0	0
	<i>ready</i>	0	0	9	115	0	0
Conf. = 53%			SSD ResNet 50 (640 × 640)				
Deteção	Referência					Não referenciada	
		<i>coty</i>	<i>minus9</i>	<i>plus9</i>	<i>ready</i>	solo	infestante
	<i>coty</i>	70	0	0	0	0	0
	<i>minus9</i>	1	546	1	0	0	1
	<i>plus9</i>	0	1	69	2	0	0
	<i>ready</i>	0	0	4	120	0	0
Conf. = 72%			YOLO v4 (416 × 416)				
Deteção	Referência					Não referenciada	
		<i>coty</i>	<i>minus9</i>	<i>plus9</i>	<i>ready</i>	solo	infestante
	<i>coty</i>	114	0	0	0	6	0
	<i>minus9</i>	0	550	0	0	0	4
	<i>plus9</i>	0	0	64	0	0	0
	<i>ready</i>	0	0	1	118	0	0

Conf. = confiança que otimiza o valor de F1 na amostra de validação (Tabela 15).

O modelo SSD Inception v2 realiza um total de 889 previsões, sendo o modelo com maior número. Contudo, o modelo YOLO v4 é aquele que apresenta uma melhor relação entre previsões corretas e total de previsões (839 previsões corretas de um total de 850). A Tabela 18 apresenta as métricas exatidão, *precision*, *recall* e F1 para cada classe considerada, calculadas na amostra de teste.

Tabela 18 – Resultados das métricas exatidão, *precision* e *recall* para cada classe de alface para os modelos SSD e YOLO em estudo. Avaliação na amostra de teste sem *augmentation*.

<i>coty</i> n = 144	Exatidão	<i>Precision</i>	<i>Recall</i>	F1
SSD MobileNet v2 (300 × 300)	92%	85%	65%	74%
SSD MobileNet v2 (640 × 640)	88%	92%	33%	49%
SSD Inception v2 (300 × 300)	97%	96%	85%	90%
SSD ResNet 50 (640 × 640)	92%	100%	49%	65%
YOLO v4 (416 × 416)	96%	95%	79%	86%
<i>minus9</i> n = 569				
SSD MobileNet v2 (300 × 300)	88%	95%	86%	90%
SSD MobileNet v2 (640 × 640)	87%	96%	85%	90%
SSD Inception v2 (300 × 300)	94%	96%	95%	95%
SSD ResNet 50 (640 × 640)	97%	99%	96%	98%
YOLO v4 (416 × 416)	97%	99%	97%	98%
<i>plus9</i> n = 81				
SSD MobileNet v2 (300 × 300)	98%	100%	81%	90%
SSD MobileNet v2 (640 × 640)	98%	96%	85%	90%
SSD Inception v2 (300 × 300)	99%	96%	88%	92%
SSD ResNet 50 (640 × 640)	98%	96%	85%	90%
YOLO v4 (416 × 416)	98%	100%	79%	88%
<i>ready</i> n = 111				
SSD MobileNet v2 (300 × 300)	100%	97%	100%	99%
SSD MobileNet v2 (640 × 640)	98%	87%	99%	93%
SSD Inception v2 (300 × 300)	99%	93%	100%	96%
SSD ResNet 50 (640 × 640)	100%	97%	100%	98%
YOLO v4 (416 × 416)	100%	99%	100%	100%

n: número de referências na amostra de teste.

Os dados da Tabela 18 sugerem que, na generalidade dos casos, todos os modelos têm um bom desempenho na classificação de alface a partir de imagens *grayscale*. O modelo SSD Inception v2 (300 × 300) apresenta um valor superior de EB, cerca de 92%.

Uma vez apresentados os resultados do desempenho dos modelos SSD e YOLO na deteção e classificação de alface, nas duas modalidades de imagem (RGB e *grayscale*), passamos agora à sua comparação. De forma geral, o desempenho dos modelos sujeitos a imagens *grayscale* é superior tanto na deteção, com um valor de AP próximo de 91% no modelo YOLO v4 (nas imagens RGB a AP apresenta o maior valor no modelo SSD MobileNet v2 (300 × 300) – 81%), como na classificação, onde o modelo SSD Inception v2 (300 × 300) se destaca dos restantes com um valor de EB próximo de 92% (nas imagens RGB a EB apresenta o maior valor no modelo SSD MobileNet v2 (640 × 640) – 79%). Na classificação independente de cada classe, os modelos treinados com imagens *grayscale* apresentam melhor desempenho, refletido nos valores superiores tanto de Exatidão como de F1. Esta diferença é particularmente notória nas classes *plus9* e *ready*. O modelo SSD MobileNet v2 (640 × 640) apresenta um desempenho semelhante, nas quatro classes consideradas, para ambas as modalidades de imagem (RGB e *grayscale*).

Como referido anteriormente, a análise completa dos resultados para as restantes culturas encontra-se no Anexo II. De seguida é apresentada uma análise geral ao desempenho dos modelos SSD e YOLO considerados para a deteção e classificação fenológica das oito culturas hortícolas em estudo.

5.1.2. Resultados gerais das culturas estudadas

A Tabela 19 resume as métricas calculadas na amostra de teste para a confiança otimizada. Assim é possível avaliar e comparar o desempenho dos diferentes modelos nas operações de detecção e classificação fenológica das oito culturas hortícolas em estudo (acelga, alface, cenoura, coentro, espinafre, nabo, rabanete e rúcula).

Tabela 19 – Resultados das métricas calculadas na amostra de teste para a confiança otimizada. Métricas calculadas para a detecção: *Average Precision*, *Precision*, *Recall* e *F1*. Métricas calculadas para a classificação: Exatidão Balanceada e *F1 Médio*. Resultados gerais para as oito culturas (acelga, alface, cenoura, coentro, espinafre, nabo, rabanete e rúcula).

Modelo	Confiança ≥	AP	Precision	Recall	F1	EB	F1 Médio
SSDMobileNetv2_300_300_acelga	33%	81%	86%	81%	83%	75%	77%
SSDMobileNetv2_300_300_alface	92%	81%	87%	84%	85%	66%	71%
SSDMobileNetv2_300_300_alface_grayscale	39%	78%	88%	80%	84%	83%	88%
SSDMobileNetv2_300_300_cenoura	30%	56%	82%	60%	69%	55%	64%
SSDMobileNetv2_300_300_coentro	37%	66%	82%	70%	76%	72%	78%
SSDMobileNetv2_300_300_espinafre	37%	98%	98%	98%	98%	100%	100%
SSDMobileNetv2_300_300_nabo	32%	63%	76%	68%	72%	81%	85%
SSDMobileNetv2_300_300_rabanete	34%	86%	90%	88%	89%	93%	96%
SSDMobileNetv2_300_300_rúcula	31%	90%	91%	91%	91%	98%	99%
SSDMobileNetv2_300_300_média	s.s.	78%	87%	80%	83%	80%	84%
SSDMobileNetv2_640_640_acelga	46%	82%	94%	82%	88%	61%	69%
SSDMobileNetv2_640_640_alface	49%	81%	88%	82%	85%	79%	88%
SSDMobileNetv2_640_640_alface_grayscale	48%	78%	93%	79%	86%	76%	81%
SSDMobileNetv2_640_640_cenoura	30%	70%	78%	74%	76%	71%	75%
SSDMobileNetv2_640_640_coentro	39%	68%	80%	72%	76%	74%	78%
SSDMobileNetv2_640_640_espinafre	65%	100%	98%	100%	99%	100%	99%
SSDMobileNetv2_640_640_nabo	36%	63%	71%	69%	70%	83%	85%
SSDMobileNetv2_640_640_rabanete	37%	83%	93%	84%	89%	85%	91%
SSDMobileNetv2_640_640_rúcula	39%	89%	91%	89%	90%	94%	95%
SSDMobileNetv2_640_640_média	s.s.	79%	88%	81%	84%	80%	85%
SSDIceptionv2_300_300_acelga	14%	82%	88%	82%	85%	66%	75%
SSDIceptionv2_300_300_alface	44%	81%	88%	82%	85%	78%	79%
SSDIceptionv2_300_300_alface_grayscale	49%	85%	88%	88%	88%	92%	93%
SSDIceptionv2_300_300_cenoura	21%	59%	81%	62%	70%	62%	71%
SSDIceptionv2_300_300_coentro	13%	68%	77%	72%	74%	72%	77%
SSDIceptionv2_300_300_espinafre	32%	95%	92%	96%	94%	98%	97%
SSDIceptionv2_300_300_nabo	24%	59%	83%	61%	71%	66%	72%
SSDIceptionv2_300_300_rabanete	24%	87%	88%	89%	88%	93%	95%
SSDIceptionv2_300_300_rúcula	39%	89%	91%	89%	90%	100%	100%
SSDIceptionv2_300_300_média	s.s.	78%	86%	80%	83%	81%	84%
SSDResNet50_640_640_acelga	52%	84%	94%	85%	89%	68%	77%
SSDResNet50_640_640_alface	51%	73%	90%	74%	82%	57%	79%
SSDResNet50_640_640_alface_grayscale	53%	88%	97%	88%	92%	82%	88%
SSDResNet50_640_640_cenoura	49%	72%	88%	75%	81%	72%	78%
SSDResNet50_640_640_coentro	47%	75%	83%	79%	81%	79%	83%
SSDResNet50_640_640_espinafre	41%	100%	100%	100%	100%	99%	99%
SSDResNet50_640_640_nabo	57%	66%	87%	68%	76%	68%	74%
SSDResNet50_640_640_rabanete	55%	94%	95%	94%	95%	96%	98%
SSDResNet50_640_640_rúcula	60%	96%	99%	96%	97%	98%	99%
SSDResNet50_640_640_média	s.s.	83%	93%	84%	88%	80%	86%
YOLOv4_416_416_acelga	65%	94%	93%	95%	94%	98%	99%
YOLOv4_416_416_alface	54%	78%	86%	81%	83%	56%	67%
YOLOv4_416_416_alface_grayscale	72%	91%	95%	91%	93%	89%	93%
YOLOv4_416_416_cenoura	59%	83%	88%	86%	87%	84%	86%
YOLOv4_416_416_coentro	50%	79%	75%	89%	81%	86%	86%
YOLOv4_416_416_espinafre	32%	100%	97%	100%	98%	100%	99%
YOLOv4_416_416_nabo	54%	87%	90%	89%	89%	90%	91%
YOLOv4_416_416_rabanete	66%	92%	92%	93%	93%	93%	95%
YOLOv4_416_416_rúcula	39%	98%	91%	99%	95%	100%	99%
YOLOv4_416_416_média	s.s.	89%	90%	91%	90%	88%	91%
Média Geral	s.s.	81%	88%	83%	86%	82%	86%

AP: *Average Precision*, EB: Exatidão Balanceada, s.s.: sem significado

No cômputo geral, para o problema de detecção: a AP média é próxima de 81% e o F1 aproxima-se de 86%, já no que diz respeito à classificação, a EB apresenta um valor médio próximo de 82% e o F1 próximo de 86%.

O modelo YOLO v4 (416 × 416) é aquele que apresenta melhor desempenho, tanto na detecção como na classificação. Isto justifica-se pelo comportamento do modelo: ao enquadrar a detecção de objetos como um problema de regressão, permite a detecção e classificação numa operação simultânea, de forma mais eficaz. Pela comparação dos modelos SSD, o modelo SSD ResNet 50 (640 × 640) é aquele que apresenta melhor desempenho, à exceção da detecção de alface e, principalmente, da classificação de alface, espinafre e nabo. As diferenças podem ser justificadas pela arquitetura dos modelos, em particular do classificador ou pelas características das imagens utilizadas para o treino dos modelos.

Os modelos treinados com as imagens de espinafre apresentam melhor desempenho em comparação com as restantes culturas. Na maioria das imagens desta cultura consta apenas uma planta, esta particularidade pode ajudar a justificar estes resultados. Pelo contrário, os modelos treinados com imagens de cenoura são aqueles que apresentam pior desempenho. Neste caso, tendo em conta as imagens do repositório, os resultados poderão ser justificados pelas semelhanças entre as plantas de cenoura e as infestantes ou outros elementos presentes nas imagens. No que se refere à classificação, provavelmente as classes definidas não apresentam características suficientemente distinguíveis, o que favorece a previsão errada das plantas de cenoura.

5.2. Reconstrução tridimensional

5.2.1. Planta individual

Na reconstrução tridimensional da planta de rabanete são gerados os seguintes resultados: i) nuvem dispersa de pontos (Figura 26b), ii) nuvem densa de pontos (Figura 26c) e iii) modelo tridimensional (Figura 26d). Todas estas representações foram acedidas pela ferramenta MeshLab³¹.

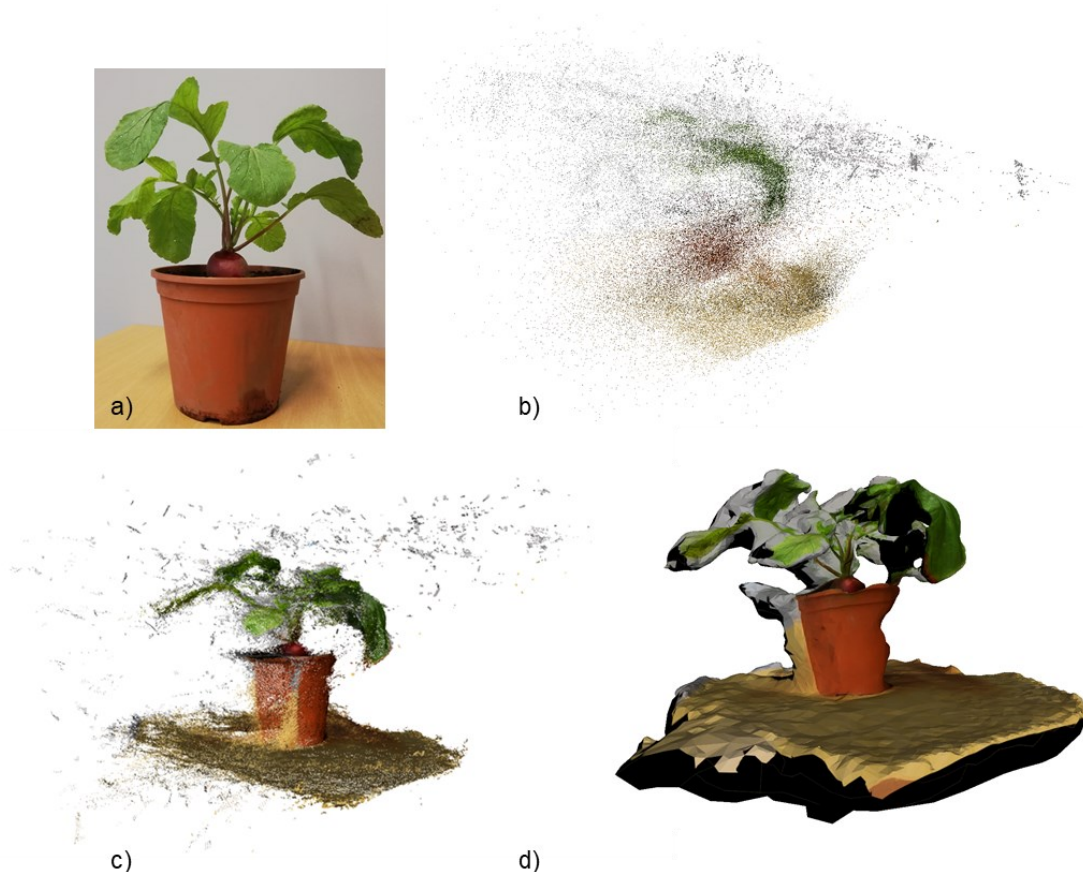


Figura 26 – Reconstrução tridimensional de rabanete. a) imagem original, b) nuvem dispersa de pontos, c) nuvem densa de pontos, d) modelo tridimensional.

No caso do coentro, as imagens submetidas para reconstrução foram processadas, sendo extraída a região de interesse de cada uma delas (Figura 27b). À semelhança do exemplo anterior são geradas as nuvens de pontos dispersa (Figura 27c) e densa (Figura 27d), e o modelo tridimensional (Figura 27e).

³¹ MeshLab (<https://www.meshlab.net/>). Último acesso em: 30/08/2021.

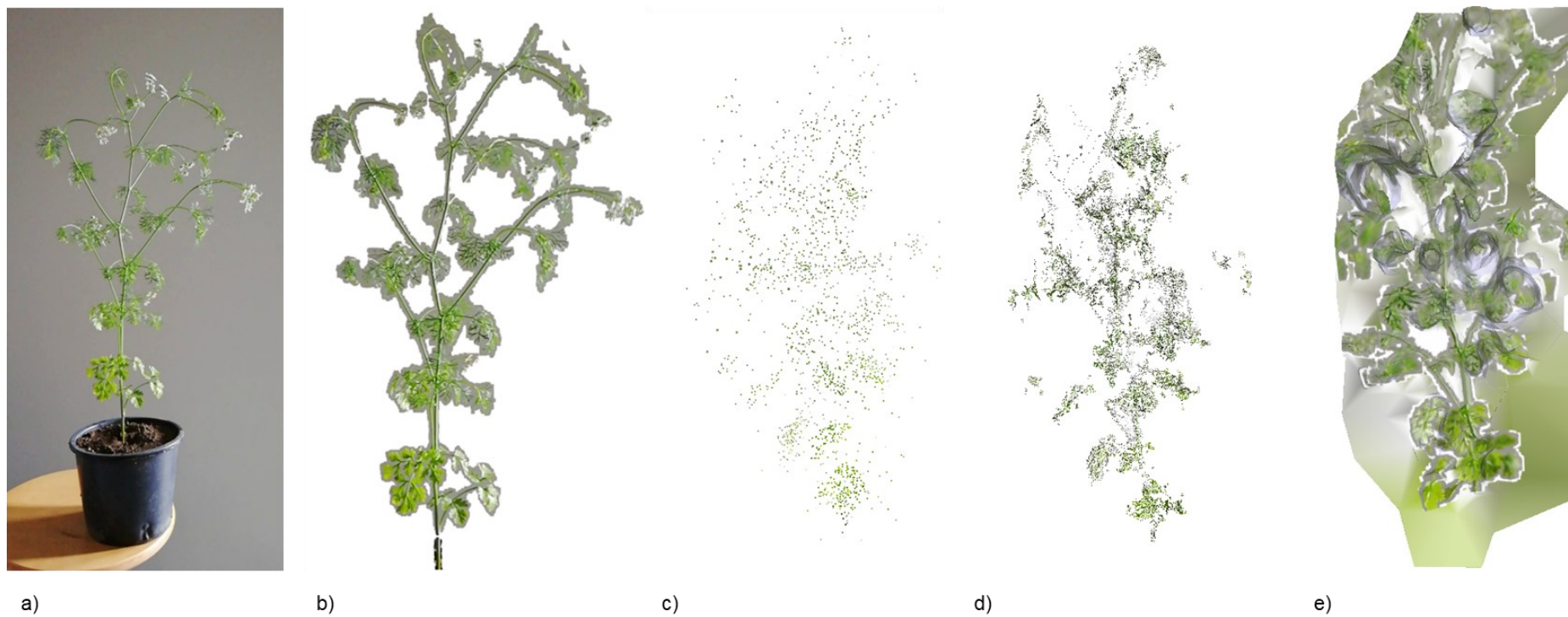


Figura 27 – Reconstrução tridimensional de coentro. a) imagem original, b) região de interesse, c) nuvem dispersa de pontos, d) nuvem densa de pontos, e) modelo tridimensional.

Para ambos os casos, é possível reconstruir devidamente as plantas consideradas. Em particular, pela visualização da nuvem de pontos densa (Figura 26c e Figura 27d) compreende-se a arquitetura principal das plantas e os restantes elementos presentes no ambiente. Parece-nos possível que a reconstrução tridimensional de plantas individuais através de técnicas de fotogrametria e *software* como o Agisoft Metashape partindo de imagens RGB bidimensionais permite a parametrização do crescimento vegetal. Como referido anteriormente, esta informação é importante para suportar as operações culturais e apoiar na tomada de decisão.

5.2.2. Grupo de plantas

No que diz respeito à reconstrução completa do talhão através do ODM, apresentamos três exemplos distintos: i) reconstrução do talhão com plantas de pequena dimensão, numa fase inicial do ciclo cultural (Figura 28), ii) reconstrução do talhão com plantas numa fase mais avançada do ciclo cultural (Figura 29) e iii) reconstrução dos talhões completos (Figura 31).

No primeiro exemplo (reconstrução do talhão com plantas de pequena dimensão, numa fase inicial do ciclo cultural), propomos a reconstrução de um talhão com plantas de alface e rabanete. Através de um conjunto de 164 imagens é possível criar as nuvens de pontos: dispersa (Figura 28a) e densa (Figura 28b) e o modelo tridimensional (Figura 28c).

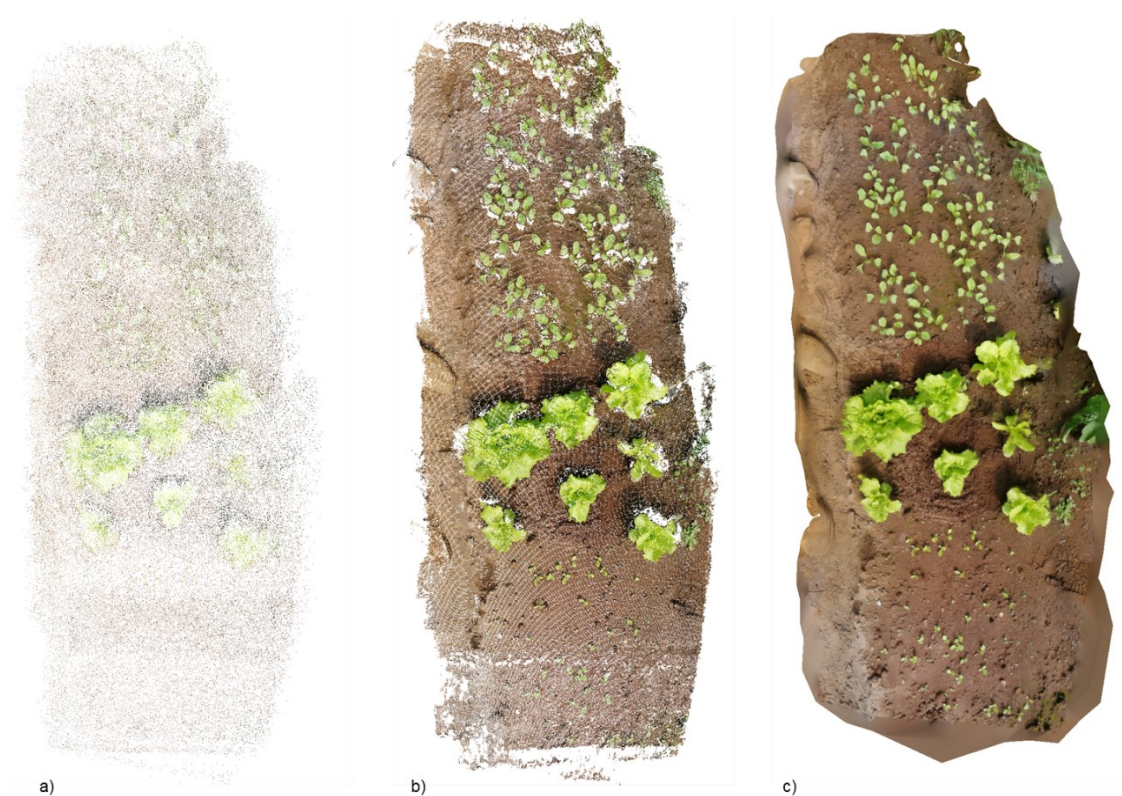


Figura 28 – Reconstrução do talhão na fase inicial do ciclo cultural. a) nuvem de pontos dispersa, b) nuvem de pontos densa, c) modelo tridimensional.

No segundo exemplo é proposta a reconstrução de um talhão com diversas culturas: rabanete, acelga, coentro, cenoura, nabo, alface e rúcula. À semelhança do demonstrado anteriormente, foram produzidas as nuvens dispersa (Figura 29a) e densa (Figura 29b) de pontos e o modelo tridimensional (Figura 29c).



Figura 29 – Reconstrução do talhão numa fase avançada do ciclo cultural. a) nuvem de pontos dispersa, b) nuvem de pontos densa, c) modelo tridimensional, d) pormenor do modelo tridimensional.

Na Figura 29d é possível observar um pormenor do modelo tridimensional do talhão. Com isto, é possível confirmar a capacidade do modelo em representar com elevada qualidade a altura das plantas e o relevo do solo.

Por último, propomos a reconstrução dos dois talhões anteriormente apresentados e onde foram instaladas as culturas que constam no repositório de imagens do PixelCropRobot. As imagens foram captadas 42 dias após a sementeira (Figura 30).



Figura 30 – Talhões 42 dias após a sementeira. a) sementeira individual de cada cultura, b) sementeira a lanço da mistura das sementes de todas as culturas.

A Figura 31 apresenta os resultados obtidos para esta reconstrução: Figura 31a: nuvem dispersa de pontos, Figura 31b: nuvem densa de pontos e Figura 31c: modelo tridimensional.



Figura 31 – Reconstrução tridimensional dos talhões. a) nuvem de pontos dispersa, b) nuvem de pontos densa, c) modelo tridimensional

Com os três exemplos descritos e os resultados correspondentes, podemos afirmar que a reconstrução tridimensional do talhão é bastante versátil. Através de técnicas de fotogrametria e de *software* como o ODM é possível reconstruir um talhão agrícola com diferentes características e posteriormente avaliar essas mesmas características (p. ex.: altura das plantas, área de solo ocupada por plantas, profundidade do solo).

6. Discussão

6.1. PixelCropRobot

É descrito na secção 3 o PixelCropRobot, um robô cartesiano de três eixos móveis, de baixo custo (cerca de 623€, detalhes no Anexo I). Do ponto de vista estrutural, a plataforma robótica desenvolvida cumpre a maioria dos objetivos a que se propôs. Nomeadamente, foi desenvolvida uma solução completamente *open-source*, construída através de componentes *off-the-shelf*, de baixo custo, leve, dirigível com um sistema de localização apoiado em códigos QR.

O sistema de localização baseado em códigos QR mostrou-se eficaz em condições laboratoriais. Em ambiente agrícola é particularmente importante o local de instalação da tira com os códigos QR. As condições de luminosidade e a oclusão dos códigos por elementos próprios da produção agrícola (p. ex.: solo, plantas) podem dificultar a deteção e descodificação destes.

Em relação à velocidade máxima apresentada para cada eixo (7,7 cm/s no eixo X, 2 cm/s no eixo Y e 0,4 cm/s no eixo Z), estas não deverão constituir um fator limitante à operabilidade do robô. É avançado por Terra [6] que o PixelCropRobot não estará preparado para movimentações constantes, devido aos erros associados à orientação do robô. O Sistema Universal de Ferramentas (SUF) percorre uma distância sobre a ponte de cerca de 1,2 m, portanto o talhão agrícola monitorizado pelo PixelCropRobot não deverá ter uma largura superior a este valor. O limite de altura do eixo Z (do SUF) está dependente dos manipuladores a desenvolver, em particular da forma como será estabelecida a ligação entre o SUF e o manipulador. Por tudo o que foi referido podemos atribuir um TRL 3 e uma classificação 2 de acordo com Ribeiro, Andujar et al. [31] para mobilidade/localização: semi dependente, dirigível e com sensibilidade parcial à envolvente e 2 para manipulação/atuação: semi dependente, desenvolve ações fixas com sensibilidade parcial face à envolvente.

Futuramente, a captação de imagens será responsabilidade do PixelCropRobot, aumenta-se a robustez do repositório de imagens, tanto em quantidade como em diversidade, podendo ser estudadas outras culturas. Para além disso, ao equipar o robô com sensores óticos e outros componentes tecnológicos, diversas operações de monitorização poderão ser levadas a cabo por esta plataforma. Certamente que o

PixelCropRobot contribuirá para uma agricultura evoluída, economicamente eficiente e ecologicamente sustentável.

6.2. Detecção e classificação fenológica de plantas

De acordo com os resultados (Tabela 19 e Anexo II) verificou-se que os cinco modelos considerados desempenham com sucesso as operações propostas, detecção e classificação fenológica, nas oito culturas hortícolas consideradas. Em relação à detecção a AP média é próxima de 81% e o F1 aproxima-se de 86%, já no que diz respeito à classificação, a EB apresenta um valor médio próximo de 82% e o F1 de 86%.

De forma geral, o modelo YOLO v4 (416 × 416) é aquele que apresenta melhor desempenho, tanto na detecção como na classificação das plantas, quando comparado com os modelos SSD. Nesta arquitetura, ao comparar o modelo SSD ResNet 50 (640 × 640) com os restantes modelos SSD, na generalidade dos casos este apresenta melhor desempenho, contudo, assume um menor desempenho na detecção de alface e, principalmente, na classificação de alface, espinafre e nabo. Estas diferenças podem ser justificadas pela arquitetura dos modelos, em particular do classificador ou pelas características das imagens utilizadas para o treino dos modelos. Este facto é mencionado por Chechliński, SiemiąTkowska et al. [86], os autores sugerem que a utilização de imagens de qualidade reduzida e condições de luminosidade diversas são fatores que prejudicam o desempenho dos modelos de detecção.

O modelo SSD MobileNet v2 (300 × 300) quando comparado com o modelo SSD MobileNet v2 (640 × 640) apresenta um desempenho inferior na detecção de acelga e espinafre, na classificação de alface e nabo e com desempenho inferior em ambas as operações no caso da cenoura e do coentro. O modelo SSD MobileNet v2 (640 × 640) apresenta um desempenho inferior na detecção e classificação de alface *grayscale* e rabanete, e um pior desempenho a classificar acelga e rúcula. O modelo YOLO v4 (416 × 416) assume um desempenho inferior na detecção e classificação de alface, o que vai ao encontro do verificado por Zheng, Kong et al. [85], onde o modelo YOLO v3 apresenta um desempenho inferior ao modelo SSD considerado pelos autores. Por outro lado, não só para esta cultura, mas no geral das categorias consideradas pelos autores, o modelo *ResNet 50* (mais especificamente *RetNet ResNet 50*) apresenta um desempenho superior relativamente aos restantes modelos considerados, à semelhança do verificado neste trabalho.

Ao considerar as características das imagens e, portanto, das culturas em estudo, podemos afirmar que os modelos treinados com imagens de espinafre apresentam melhor desempenho em comparação com as restantes culturas. Na maioria das imagens desta cultura consta apenas uma planta, esta particularidade pode ajudar a justificar estes resultados. Pelo contrário, os modelos treinados com imagens de cenoura são aqueles que apresentam pior desempenho, de uma forma geral. Neste caso, tendo em conta as imagens do repositório, os resultados poderão ser justificados pelas semelhanças entre as plantas de cenoura e as infestantes ou outros elementos presentes nas imagens. No que se refere à classificação, provavelmente as classes definidas não apresentam características suficientemente distinguíveis, o que favorece a previsão errada das plantas de cenoura. Para apoiar nesta discussão, os modelos foram treinados com imagens *grayscale* de alface, os quais mostraram um desempenho superior quando comparado com o apresentado pelos modelos treinados com imagens RGB. Podemos sugerir que as imagens *grayscale* ao conterem menos informação, nomeadamente colorimétrica, impõe aos modelos uma maior capacidade de extração de características e assim aumentam a sua robustez na deteção de plantas.

Podemos agora discutir os resultados relativos à classificação, em particular das classes iniciais, ou seja, a classe *coty* (apenas o espinafre não apresenta esta classe). A Exatidão média para a classificação desta classe é: SSD MobileNet v2 (300 × 300), cerca de 88%, SSD MobileNet v2 (640 × 640), 89%, SSD Inception v2 (300 × 300), 89%, SSD ResNet 50 (640 × 640), 90% e YOLO v4 (416 × 416), 95%. Uma vez que se tratam de plantas com um tamanho reduzido, uma das formas para otimizar estes resultados será através da diminuição da resolução das imagens, à semelhança do realizado por Samiei, Rasti et al. [96] que aplicam nos seus modelos de classificação fenológica imagens de 89 × 89 (px.). O elevado desempenho demonstrado por todos os modelos, a classificar corretamente a fase inicial do ciclo cultural permitirá uma monitorização consistente do crescimento das plantas, o que pode favorecer a antecipação da tomada de decisões agronómicas.

6.3. Reconstrução tridimensional

Fomos capazes, através de técnicas de fotogrametria, principalmente o método SfM, apresentar modelos que representam, tridimensionalmente, tanto plantas individuais como grupos de plantas. Ao aplicar estas técnicas a diferentes escalas espaciais compreendemos a versatilidade das mesmas.

No que diz respeito às plantas individuais fomos capazes de reconstruir duas plantas com características distintas, em particular, o coentro que, no exemplo apresentado apresenta uma arquitetura complexa. Pela comparação visual dos modelos gerados no âmbito deste trabalho com os apresentados por Liu, Acosta-Gamboa et al. [104], podemos afirmar que os nossos, embora apresentem algumas falhas, foram capazes de reconstruir plantas mais complexas que a utilizada pelos autores (*Arabidopsis thaliana*).

No caso da reconstrução de grupos de plantas, foram apresentados três exemplos distintos que refletem três fases distintas do ciclo cultural de culturas hortícolas. Tanto as nuvens de pontos como os modelos tridimensionais apresentados apresentam qualidade suficiente para serem utilizados com o objetivo de monitorizar o crescimento vegetal, através da extração de características e cálculo de parâmetros que quantifiquem o desenvolvimento da planta.

Tanto os modelos anteriormente referidos para deteção e classificação automática de plantas quanto as técnicas de fotogrametria responsáveis pela reconstrução tridimensional são passíveis de ser aplicadas no PixelCropRobot com várias aplicações na monitorização do crescimento vegetal.

7. Conclusão

No âmbito deste trabalho foi desenvolvido o PixelCropRobot, um robô cartesiano de três eixos. A locomoção do robô é realizada através de rodas e o sistema de localização baseia-se em códigos QR. A maioria dos objetivos a que nos propusemos aquando da construção do robô foram cumpridos. Para comprovar a capacidade do robô em operar em ambiente real seria necessário testá-lo num campo agrícola.

Foram captadas, processadas e anotadas 4123 imagens de oito culturas hortícolas que constam no repositório de imagens do PixelCropRobot. Procedeu-se à gravação de vídeos tanto de plantas individuais como de grupos de plantas.

Esta informação foi processada através de modelos *Deep Learning* para deteção e classificação fenológica de plantas. No que diz respeito à deteção a *Average Precision* média é próxima de 81% e o F1 aproximou-se de 86%, referente à classificação, a Exatidão Balanceada apresenta um valor médio próximo de 82% e o F1 próximo de 86%. O modelo YOLO v4 (416 × 416) é aquele que apresenta melhor desempenho, tanto na deteção como na classificação das plantas, podendo assim ser considerado o modelo mais versátil, tendo em conta os modelos estudados. A deteção precisa das culturas possibilita o desenvolvimento de métodos mais eficazes de controlo de infestantes, por sua vez, a classificação correta da fase do ciclo cultural suporta melhores práticas culturais (p. ex: rega, pulverização, fertilização), adaptadas ao desenvolvimento da cultura, nomeadamente no que diz respeito à aplicação localizada e variável de fatores de produção.

Através de técnicas de fotogrametria foram realizados diversos processos de reconstrução tridimensional a plantas individuais e grupos de plantas. Através dos resultados gerados (nuvens de pontos e modelos de superfícies) será, futuramente, possível avaliar o arranjo geométrico e características relacionadas, parametrizando o crescimento vegetal a diferentes níveis espaciais. As informações recolhidas e processadas podem ser aplicadas tanto para apoiar as decisões agronómicas como para suportar operações culturais.

O rigor da deteção e classificação obtido para as várias culturas estudadas, aliado ao baixo custo da plataforma robótica desenvolvida e os sensores (RGB) testados, constituem indicadores da sua utilidade para a fenotipagem de culturas para melhorar a eficiência de diferentes práticas agrícolas. Estes fatores justificam a continuidade dos trabalhos para desenvolvimento do PixelCropRobot no sentido de elevar o TRL do seu protótipo aqui apresentado.

8. Referências Bibliográficas

ISO 8373:2012. (2012). *Robots and robotic devices – Vocabulary*. Geneva, Switzerland.

1. Oliveira, L.F.; Silva, M. and Moreira, A. Agricultural Robotics: A State of the Art Survey. In Proceedings of the International Conference on Climbing and Walking Robots and the Support Technologies for Mobile Machines, Moscow, Russian Federation, 24-26 August, 2020; pp. 279-286. doi:<https://doi.org/10.13180/clawar.2020.24-26.08.44>.
2. Bechar, A. and Vigneault, C. Agricultural robots for field operations: Concepts and components. *Biosys. Eng.* **2016**, *149*, 94-111. doi:<https://doi.org/10.1016/j.biosystemseng.2016.06.014>.
3. Mavridou, E.; Vrochidou, E.; Papakostas, G.A.; Pachidis, T. and Kaburlasos, V.G. Machine Vision Systems in Precision Agriculture for Crop Farming. *Journal of Imaging* **2019**, *5*, 89. doi:<https://doi.org/10.3390/jimaging5120089>.
4. Sharma, A.; Jain, A.; Gupta, P. and Chowdary, V. Machine Learning Applications for Precision Agriculture: A Comprehensive Review. *IEEE Access* **2021**, *9*, 4843-4873. doi:<https://dx.doi.org/10.1109/access.2020.3048415>.
5. Narvaez, F.Y.; Reina, G.; Torres-Torriti, M.; Kantor, G. and Cheein, F.A. A Survey of Ranging and Imaging Techniques for Precision Agriculture Phenotyping. *IEEE/ASME Transactions on Mechatronics* **2017**, *22*, 2428-2439. doi:<http://dx.doi.org/10.1109/TMECH.2017.2760866>.
6. Terra, F. Robô de baixo custo para mini hortas. Masters Thesis, Engineering Faculty, Porto University, Repositório Aberto da Universidade do Porto, 2021. Available online:<https://hdl.handle.net/10216/134951> (accessed on 16/07/2021)
7. Terra, F.; Rodrigues, L.; Magalhães, S.; Santos, F.; Moura, P. and Cunha, M. PixelCropRobot, a cartesian multitask platform for microfarms automation. In Proceedings of the International Symposium of Asian Control Association on Intelligent Robotics and Industrial Automation (IRIA), Goa, India, 20-22 September, 2021; pp. 382-387. doi:<https://doi.org/10.1109/IRIA53009.2021.9588786>.
8. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Networks* **2015**, *61*, 85-117. doi:<https://doi.org/10.1016/j.neunet.2014.09.003>.
9. Paulus, S. Measuring crops in 3D: using geometry for plant phenotyping. *Plant Methods* **2019**, *15*. doi:<https://doi.org/10.1186/s13007-019-0490-0>.
10. Gentry, M. Local Heat, Local Food: Integrating Vertical Hydroponic Farming with District Heating in Sweden. *Energy* **2019**, *174*, 191-197. doi:<https://dx.doi.org/10.1016/j.energy.2019.02.119>.
11. Stein, A. and Santini, F. The sustainability of “local” food: a review for policy-makers. *Review of Agricultural, Food and Environmental Studies* **2021**. doi:<https://dx.doi.org/10.1007/s41130-021-00148-w>.
12. Reyes-García, V.; Aceituno, L.; Vila, S.; Calvet-Mir, L.; Garnatje, T.; Jesch, A.; Lastra, J.J.; Parada, M.; Rigat, M.; Vallès, J. and Pardo-De-Santayana, M. Home Gardens in Three Mountain Regions of the Iberian Peninsula: Description, Motivation for Gardening, and Gross Financial Benefits. *J. Sustainable Agric.* **2012**, *36*, 249-270. doi:<https://dx.doi.org/10.1080/10440046.2011.627987>.
13. Odebode, S. Assessment of home gardening as a potential source of household income in Akinyele local government area of Oyo State. *Nigerian Journal Horticultural Science* **2006**, *11*, 47-

55. Available online: <https://www.ajol.info/index.php/njhs/article/view/3398> (accessed on 09/03/2021).
14. Kumar, Y. and Rufus, E. Smart Kitchen Garden using “BloThrough” at a Low Cost. In Proceedings of the Fourteenth International Conference on Information Processing (ICINPRO), Bangalore, India, 21-23 December, 2018; pp. 1-3. doi:<https://doi.org/10.1109/ICINPRO43533.2018.9096844>.
15. International Federation of Robotics. Available online: <https://ifr.org/> (accessed on 22/10/2020).
16. Gonzalez-de-Santos, P.; Fernández, R.; Sepúlveda, D.; Navas, E.; Emmi, L. and Armada, M. Field Robots for Intelligent Farms-Inhering Features from Industry. *Agronomy* **2020**, *10*, 1638. doi:<https://doi.org/10.3390/agronomy10111638>.
17. Fountas, S.; Mylonas, N.; Malounas, I.; Rodias, E.; Hellmann Santos, C. and Pekkeriet, E. Agricultural Robotics for Field Operations. *Sensors* **2020**, *20*, 2672. doi:<https://doi.org/10.3390/s20092672>.
18. Haibo, L.; Shuliang, D.; Zunmin, L. and Chuijie, Y. Study and Experiment on a Wheat Precision Seeding Robot. *Journal of Robotics* **2015**, *2015*, 696301. doi:<https://doi.org/10.1155/2015/696301>.
19. Lee, W.; Slaughter, D. and Giles, D. Robotic Weed Control System for Tomatoes. *Precision Agriculture* **1999**, *1*, 95-113. doi:<https://doi.org/10.1023/A:1009977903204>.
20. Reiser, D.; Sehsah, E.; Bumann, O.; Morhard, J. and Griepentrog, H. Development of an Autonomous Electric Robot Implement for Intra-Row Weeding in Vineyards. *Agriculture* **2019**, *9*, 18. doi:<https://doi.org/10.3390/agriculture9010018>.
21. Ecorobotix, Technology for Environment. Available online: <https://www.ecorobotix.com/en/> (accessed on 14/11/2020).
22. Schor, N.; Berman, S.; Dombrovsky, A.; Elad, Y.; Ignat, T. and Bechar, A. Development of a robotic detection system for greenhouse pepper plant diseases. *Precision Agriculture* **2017**, *18*, 394-409. doi:<https://doi.org/10.1007/s11119-017-9503-z>.
23. Sultan, M.M.; Zaman, Q.U.; Esau, T.J.; Price, G.W. and Prithiviraj, B. Development of an artificial cloud lighting condition system using machine vision for strawberry powdery mildew disease detection. *Comput. Electron. Agric.* **2019**, *158*, 219-225. doi:<https://doi.org/10.1016/j.compag.2019.02.007>.
24. Vidoni, R.; Gallo, R.; Ristorto, G.; Carabin, G.; Mazzetto, F.; Scalera, L. and Gasparetto, A. ByeLab: An Agricultural Mobile Robot Prototype for Proximal Sensing and Precision Farming. In Proceedings of the ASME International Mechanical Engineering Congress and Exposition, Tampa, Florida, USA, 3-9 November, 2017. doi:<https://doi.org/10.1115/IMECE2017-71216>.
25. Virlet, N.; Sabermanesh, K.; Sadeghi-Tehran, P. and Hawkesford, M.J. Field Scanalyzer: An automated robotic field phenotyping platform for detailed crop monitoring. *Funct. Plant Biol.* **2017**, *44*, 143-153. doi:<https://doi.org/10.1071/FP16163>.
26. Hayashi, S.; Ganno, K.; Ishii, Y. and Tanaka, I. Robotic Harvesting System for Eggplants. *Japan Agricultural Research Quarterly* **2002**, *36*, 163-168. doi:<https://doi.org/10.6090/jarq.36.163>.
27. Rowley, J. Developing flexible automation for mushroom harvesting (*Agaricus bisporus*) : innovation report. Tese de Doutorado, Universidade de Warwick, University of Warwick Institutional Repository, 2009. Available online: http://wrap.warwick.ac.uk/36893/1/WRAP_THESIS_Rowley_2009.pdf (accessed on 15/05/2021)

28. Green, O.; Schmidt, T.; Pietrzkowski, R.; Jensen, K.; Larsen, M. and Jørgensen, R. Commercial autonomous agricultural platform: Kongsilde Robotti. In Proceedings of the Second International Conference on Robotics, Associated High-Technologies and Equipment for Agriculture and Forestry, Madrid, Spain, 21-23 May, 2014; pp. 351-356. Available online: <http://hdl.handle.net/10261/130548> (accessed on 29/02/2021).
29. Grimstad, L. and From, P. The Thorvald II Agricultural Robotic System. *Robotics* **2017**, 6, 24. doi:<https://doi.org/10.3390/robotics6040024>.
30. Jasiński, M.; Mączak, J.; Szulim, P. and Radkowski, S. Autonomous Agricultural Robot – Testing of the Vision System for Plants/Weed Classification. In Proceedings of the Automation, Warsaw, Poland, 21-23 March, 2018; pp. 473-482. doi:https://doi.org/10.1007/978-3-319-77179-3_44.
31. Ribeiro, A.; Andujar, D.; Karouta, J.; Michalidis, A.; Vieri, M.; Lombardo, S.; de Pascale, V.; da Silva, J.; Correia, M.; Coimbra, J. and Tomé, N. Robotics for Future Agriculture - Advanced robotics for precision agriculture, State-of-art and future trends. **2019**, Available online: <http://sparkle-project.eu/wp-content/uploads/2019/09/Robotics-for-future-agriculture.pdf> (accessed on 08/02/2021).
32. Comissão Europeia. Technology Readiness Levels (TRL). **2014**, Available online: https://ec.europa.eu/research/participants/data/ref/h2020/wp/2014_2015/annexes/h2020-wp1415-annex-g-trl_en.pdf (accessed on 18/06/2021).
33. Gil, L.; Andrade, M.H. and Costa, M.d.C. Os TRL (*Technology Readiness Levels*) como ferramenta na avaliação tecnológica. *Revista Ingenium* N° 139 **2014**, 94-96. Available online: https://repositorio.lneg.pt/bitstream/10400.9/2771/1/Os%20TRL%20%28TECHNOLOGY%20READINESS%20LEVELS%29%20COMO%20FERRAMENTA%20NA%20AVALIA%c3%87%c3%83O%20TECNOL%c3%93GICA_LGil.pdf (accessed on 18/06/2021).
34. Bogue, R. Robots poised to transform agriculture. *International Journal of Robotics Research and Application* **2021**, 48, 637-642. doi:<https://doi.org/10.1108/IR-05-2021-0097>.
35. Lowenberg-Deboer, J.; Huang, I.Y.; Grigoriadis, V. and Blackmore, S. Economics of robots and automation in field crop production. *Precision Agriculture* **2020**, 21, 278-299. doi:<https://dx.doi.org/10.1007/s11119-019-09667-5>.
36. Bechar, A. and Vigneault, C. Agricultural robots for field operations. Part 2: Operations and systems. *Biosys. Eng.* **2017**, 153, 110-128. doi:<https://dx.doi.org/10.1016/j.biosystemseng.2016.11.004>.
37. Iron Ox. Available online: <https://ironox.com/> (accessed on 15/10/2020).
38. Sky Greens. Available online: <https://www.skygreens.com/> (accessed on 15/10/2020).
39. Agbotic. Agbotic, Building a smarter, brighter farm future with roots and robots. Available online: <https://www.agbotic.com/> (accessed on 23/10/2020).
40. Easton Robotics. Easton Robotics, Agricultural Robotics for Small Farms. Available online: <https://www.eastonrobotics.com/> (accessed on 24/11/2020).
41. Naïo Technologies. Autonomous Oz weeding robot. Available online: <https://www.naio-technologies.com/en/oz/> (accessed on 25/10/2020).
42. Tertill The Weeding Robot. Available online: <https://tertill.com/> (accessed on 14/4/2021).
43. Wu, X.; Aravecchia, S.; Lottes, P.; Stachniss, C. and Pradalier, C. Robotic weed control using automated weed and crop classification. *Journal of Field Robotics* **2020**, 37, 322-340. doi:<https://doi.org/10.1002/rob.21938>.

44. FarmBot. FarmBot - Open Source CNC Farming. Available online: <https://farm.bot/> (accessed on 22/07/2020).
45. OpenFarm - Farm and garden through knowledge sharing. Available online: <https://openfarm.cc/en> (accessed on 20/10/2020).
46. Aronson, R. FarmBot - Humanity's open-source automated precision farming machine. **2013**, Available online: <https://farm.bot/pages/whitepaper> (accessed on 29/10/2020).
47. ROMI - Robotics for Microfarms. Available online: <https://romi-project.eu/> (accessed on 18/10/2020).
48. Colliaux, D. and Hanappe, P. LettuceThink : A open and versatile robotic platform for weeding and crop monitoring on microfarms. *International Journal of Sustainable Agricultural Management and Informatics* **2017**, 1, 171-182. Available online: <https://csl.sony.fr/wp-content/themes/sony/uploads/pdf/colliaux-17a.pdf> (accessed on 1/11/2020).
49. Axbom, S. and Ralsgård, L. Design of an Autonomous Weeding Vehicle - Used in the Agricultural Industry. Tese de Mestrado, Lund University, Lund University Libraries, 2018. Available online: <https://lup.lub.lu.se/luur/download?func=downloadFile&recordId=8968729&fileId=8968730> (accessed on 24/11/2020)
50. Oliveira, L.F.P.; Moreira, A.P. and Silva, M.F. Advances in Agriculture Robotics: A State-of-the-Art Review and Challenges Ahead. *Robotics* **2021**, 10, 52. doi: <https://doi.org/10.3390/robotics10020052>.
51. Lenain, R.; Peyrache, J.; Savary, A. and Séverac, G. Agricultural robotics: part of the new deal? - FIRA 2020 conclusions: With 27 agricultural robot information sheets. **2021**, doi: <https://doi.org/10.35690/978-2-7592-3382-3>.
52. Khan, N.; Medlock, G.; Graves, S. and Anwar, S. GPS Guided Autonomous Navigation of a Small Agricultural Robot with Automated Fertilizing System. *SAE Technical Paper* **2018**. doi: <https://doi.org/10.4271/2018-01-0031>.
53. Durmuş, H.; Güneş, E.O.; Kırıcı, M. and Üstündağ, B.B. The design of general purpose autonomous agricultural mobile-robot: "AGROBOT". In Proceedings of the Fourth International Conference on Agro-Geoinformatics (Agro-geoinformatics), Istanbul, Turkey, 20-24 July, 2015; pp. 49-53. doi: <https://doi.org/10.1109/Agro-Geoinformatics.2015.7248088>.
54. Sukkarieh, S. Mobile on-farm digital technology for smallholder farmers. In Proceedings of the Crawford Fund 2017 Annual Conference, Canberra, Australia, 7-8 August, 2017; pp. 92-100. doi: <https://doi.org/10.22004/ag.econ.266635>.
55. FarmBot Eden. Available online: <https://farmboteden.wordpress.com/> (accessed on 15/11/2020).
56. Aronson, R. FarmBot Genesis - Systems training slides. **2020**, Available online: <https://farm.bot/pages/training> (accessed on 29/10/2020).
57. GreenBot, a versatile all-rounder. Available online: <https://precisionmakers.com/en/greenbot> (accessed on 15/06/2021).
58. Clearpath Robotics Inc. Jackal, Unmanned Ground Vehicle. Available online: <https://clearpathrobotics.com/jackal-small-unmanned-ground-vehicle/> (accessed on 26/07/2021).
59. dCentralized Systems. Land Care Robot. Available online: <https://directedmachines.com/LCR24Z.html> (accessed on 15/06/2021).

60. de Araújo Batista, A.V.; Albiero, D.; de Araújo Viana, T.V.; de Almeida Monteiro, L.; Chioderoli, C.A.; de Sousa, I.R.S. and Azevedo, B.M. Multifunctional Robot at low cost for small farms. *Ciência Rural* **2017**, 47. doi:<https://doi.org/10.1590/0103-8478cr20170080>.
61. Xue, J.; Fan, B.; Zhang, X. and Feng, Y. An Agricultural Robot for Multipurpose Operations in a Greenhouse. In Proceedings of the International Conference on Mechanical and Mechatronics Engineering (ICMME), Bangkok, Thailand, 26-27 March, 2017. doi:<https://doi.org/10.12783/dtetr/icmme2017/9073>.
62. Cortés, C.; Jaimes, C. and Moreno, G. RobotGreen: A Teleoperated Agricultural Robot for Structured Environments. *Journal of Engineering Science and Technology Review* **2018**, 11, 145-155. doi:<https://doi.org/10.25103/jestr.116.18>.
63. Dhondt, S.; Wuyts, N. and Inzé, D. Cell to whole-plant phenotyping: the best is yet to come. *Trends Plant Sci.* **2013**, 18, 428-439. doi:<https://doi.org/10.1016/j.tplants.2013.04.008>.
64. Mora, M.; Avila, F.; Carrasco-Benavides, M.; Maldonado, G.; Olguín-Cáceres, J. and Fuentes, S. Automated computation of leaf area index from fruit trees using improved image processing algorithms applied to canopy cover digital photographs. *Comput. Electron. Agric.* **2016**, 123, 195–202. doi:<https://doi.org/10.1016/j.compag.2016.02.011>.
65. Gongal, A.; Silwal, A.; Amatya, S.; Karkee, M.; Zhang, Q. and Lewis, K. Apple crop-load estimation with over-the-row machine vision system. *Comput. Electron. Agric.* **2016**, 120, 26–35. doi:<https://doi.org/10.1016/j.compag.2015.10.022>.
66. Karkee, M.; Adhikari, B.; Amatya, S. and Zhang, Q. Identification of pruning branches in tall spindle apple trees for automated pruning. *Comput. Electron. Agric.* **2014**, 103, 127-135. doi:<https://doi.org/10.1016/j.compag.2014.02.013>.
67. Du, L.; Gong, W.; Shi, S.; Yang, J.; Sun, J.; Zhu, B. and Song, S. Estimation of rice leaf nitrogen contents based on hyperspectral LIDAR. *International Journal of Applied Earth Observation and Geoinformation* **2016**, 44, 136-143. doi:<https://doi.org/10.1016/j.jag.2015.08.008>.
68. Li, L.; Zhang, Q. and Huang, D. A Review of Imaging Techniques for Plant Phenotyping. *Sensors* **2014**, 14, 20078-20111. doi:<https://doi.org/10.3390/s141120078>.
69. Johnson, S. *Stephen Johnson on Digital Photography*; O'Reilly Media, Incorporated: USA, 2006.
70. Fahlgren, N.; Gehan, M.A. and Baxter, I. Lights, camera, action: high-throughput plant phenotyping is ready for a close-up. *Curr. Opin. Plant Biol.* **2015**, 24, 93-99. doi:<https://doi.org/10.1016/j.pbi.2015.02.006>.
71. Li, Z.; Guo, R.; Li, M.; Chen, Y. and Li, G. A review of computer vision technologies for plant phenotyping. *Comput. Electron. Agric.* **2020**, 176, 105672. doi:<https://doi.org/10.1016/j.compag.2020.105672>.
72. Kamilaris, A. and Prenafeta-Boldú, F.X. Deep learning in agriculture: A survey. *Comput. Electron. Agric.* **2018**, 147, 70-90. doi:<https://doi.org/10.1016/j.compag.2018.02.016>.
73. Hartley, R. and Zisserman, A. *Multiple View Geometry in Computer Vision*, 2^a ed.; Cambridge University Press: Cambridge, 2004.
74. Helm, J.M.; Swiergosz, A.M.; Haeberle, H.S.; Karnuta, J.M.; Schaffer, J.L.; Krebs, V.E.; Spitzer, A.I. and Ramkumar, P.N. Machine Learning and Artificial Intelligence: Definitions, Applications, and Future Directions. *Current Reviews in Musculoskeletal Medicine* **2020**, 13, 69-76. doi:<https://dx.doi.org/10.1007/s12178-020-09600-8>.

75. IBM. What is computer vision? Available online: <https://www.ibm.com/topics/computer-vision> (accessed on 17/04/2021).
76. Mohri, M.; Rostamizadeh, A. and Talwalkar, A. *Foundations of Machine Learning*, 2^a ed.; MIT press: 2018.
77. Agarwal, S.; du Terrail, J.O. and Jurie, F. Recent Advances in Object Detection in the Age of Deep Convolutional Neural Networks. *arXiv* **2019**. arXiv:1809.03193v2.
78. Choi, R.Y.; Coyner, A.S.; Kalpathy-Cramer, J.; Chiang, M.F. and Campbell, J.P. Introduction to Machine Learning, Neural Networks, and Deep Learning. *Translational Vision Science & Technology* **2020**, 9, 14. doi:<https://doi.org/10.1167/tvst.9.2.14>.
79. Benos, L.; Tagarakis, A.C.; Dolias, G.; Berruto, R.; Kateris, D. and Bochtis, D. Machine Learning in Agriculture: A Comprehensive Updated Review. *Sensors* **2021**, 21, 3758. doi:<https://dx.doi.org/10.3390/s21113758>.
80. Wäldchen, J.; Rzanny, M.; Seeland, M. and Mäder, P. Automated plant species identification - Trends and future directions. *PLoS Comp. Biol.* **2018**, 14. doi:<https://dx.doi.org/10.1371/journal.pcbi.1005993>.
81. Kamilaris, A. and Prenafeta-Boldú, F.X. A review of the use of convolutional neural networks in agriculture. *The Journal of Agricultural Science* **2018**, 156, 312-322. doi:<https://doi.org/10.1017/S0021859618000436>.
82. Oquab, M.; Bottou, L.; Laptev, I. and Sivic, J. Learning and Transferring Mid-level Image Representations Using Convolutional Neural Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23-28 June, 2014. doi:<https://dx.doi.org/10.1109/cvpr.2014.222>.
83. Saleem, M.H.; Potgieter, J. and Arif, K.M. Automation in Agriculture by Machine and Deep Learning Techniques: A Review of Recent Developments. *Precision Agriculture* **2021**, 22, 2053–2091. doi:<https://doi.org/10.1007/s11119-021-09806-x>.
84. Hasan, A.M.; Sohel, F.; Diepeveen, D.; Laga, H. and Jones, M.G.K. A survey of deep learning techniques for weed detection from images. *Comput. Electron. Agric.* **2021**, 184, 106067. doi:<https://dx.doi.org/10.1016/j.compag.2021.106067>.
85. Zheng, Y.; Kong, J.; Jin, X.; Wang, X.; Su, T. and Zuo, M. CropDeep: The Crop Vision Dataset for Deep-Learning-Based Classification and Detection in Precision Agriculture. *Sensors* **2019**, 19, 1058. doi:<https://doi.org/10.3390/s19051058>.
86. Chechliński, Ł.; Siemiątkowska, B. and Majewski, M. A System for Weeds and Crops Identification Based on Convolutional Neural Network. In Proceedings of the Automation, Warsaw, Poland, 21-23 March, 2018; pp. 193-202. doi:https://doi.org/10.1007/978-3-319-77179-3_18.
87. Chavan, T.R. and Nandedkar, A.V. AgroAVNET for crops and weeds classification: A step forward in automatic farming. *Comput. Electron. Agric.* **2018**, 154, 361-372. doi:<https://doi.org/10.1016/j.compag.2018.09.021>.
88. Giselsson, T.; Jørgensen, R.; Jensen, P.; Dyrmann, M. and Midtiby, H. A Public Image Database for Benchmark of Plant Seedling Classification Algorithms. *arXiv* **2017**. arXiv:1711.05458.
89. Dyrmann, M.; Karstoft, H. and Midtiby, H.S. Plant species classification using deep convolutional neural network. *Biosys. Eng.* **2016**, 151, 72-80. doi:<https://doi.org/10.1016/j.biosystemseng.2016.08.024>.

90. Knoll, F.J.; Czymbek, V.; Poczihoski, S.; Holtorf, T. and Hussmann, S. Improving efficiency of organic farming by using a deep learning classification approach. *Comput. Electron. Agric.* **2018**, *153*, 347-356. doi:<https://doi.org/10.1016/j.compag.2018.08.032>.
91. Czymbek, V.; Harders, L.O.; Knoll, F.J. and Hussmann, S. Vision-Based Deep Learning Approach for Real-Time Detection of Weeds in Organic Farming. In Proceedings of the IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Auckland, New Zealand, 20-23 May, 2019; pp. 1-5. doi:<https://doi.org/10.1109/I2MTC.2019.8826921>.
92. Milioto, A.; Lottes, P. and Stachniss, C. Real-time Semantic Segmentation of Crop and Weed for Precision Agriculture Robots Leveraging Background Knowledge in CNNs. *arXiv* **2018**. arXiv:1709.06764v2.
93. Lottes, P.; Behley, J.; Milioto, A. and Stachniss, C. Fully Convolutional Networks With Sequential Information for Robust Crop and Weed Detection in Precision Farming. *IEEE Robotics and Automation Letters* **2018**, *3*, 2870-2877. doi:<https://doi.org/10.1109/LRA.2018.2846289>.
94. Chaaro, L. and Antón, L.M. Crop and weed detection using image processing and deep learning techniques. Bachelor Degree, School of Engineering Science, University of Skövde, DiVa Portal, 2020. Available online:<https://www.diva-portal.org/smash/get/diva2:1447357/FULLTEXT01.pdf> (accessed on 13/02/2021)
95. Yalcin, H. Phenology recognition using deep learning: DeepPheno. In Proceedings of the 26th Signal Processing and Communications Applications Conference (SIU), Izmir, Turkey, 2-5 May, 2018; pp. 1-4. doi:<https://doi.org/10.1109/SIU.2018.8404165>.
96. Samiei, S.; Rasti, P.; Ly Vu, J.; Buitink, J. and Rousseau, D. Deep learning-based detection of seedling development. *Plant Methods* **2020**, *16*, 103. doi:<https://doi.org/10.1186/s13007-020-00647-9>.
97. Osorio, K.; Puerto, A.; Pedraza, C.; Jamaica, D. and Rodríguez, L. A Deep Learning Approach for Weed Detection in Lettuce Crops Using Multispectral Images. *AgriEngineering* **2020**, *2*, 471-488. doi:<https://doi.org/10.3390/agriengineering2030032>.
98. Paproki, A.; Sirault, X.; Berry, S.; Furbank, R. and Fripp, J. A novel mesh processing based technique for 3D plant analysis. *BMC Plant Biol.* **2012**, *12*, 63. doi:<https://doi.org/10.1186/1471-2229-12-63>.
99. van der Heijden, G.; Song, Y.; Horgan, G.; Polder, G.; Dieleman, A.; Bink, M.; Palloix, A.; van Eeuwijk, F. and Glasbey, C. SPICY: towards automated phenotyping of large pepper plants in the greenhouse. *Funct. Plant Biol.* **2012**, *39*, 870-877. doi:<https://doi.org/10.1071/FP12019>.
100. Schönberger, J. and Frahm, J.M. Structure-from-Motion Revisited. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27-30 June, 2016. doi:<https://doi.org/10.1109/CVPR.2016.445>.
101. Mathews, A.J. and Jensen, J.L.R. Visualizing and Quantifying Vineyard Canopy LAI Using an Unmanned Aerial Vehicle (UAV) Collected High Density Structure from Motion Point Cloud. *Remote Sensing* **2013**, *5*, 2164-2183. doi:<https://doi.org/10.3390/rs5052164>.
102. Fuhrmann, S.; Langguth, F. and Goesele, M. MVE – A Multi-View Reconstruction Environment. In Proceedings of the Eurographics Workshop on Graphics and Cultural Heritage, Darmstadt, Germany, 6-8 October, 2014; pp. 11-18. doi:<http://dx.doi.org/10.2312/gch.20141299>.
103. Wintz, T.; Lahlou, A. and Hanappe, P. Results of trained models. *ROMI - Robotics for Microfarms* **2020**. Available online:<https://media.romi-project.eu/documents/ROMI-D6.3-Results-of-trained-models.pdf> (accessed on 29/03/2021).

104. Liu, S.; Acosta-Gamboa, L.M.; Huang, X. and Lorence, A. Novel Low Cost 3D Surface Model Reconstruction System for Plant Phenotyping. *Journal of Imaging* **2017**, *3*, 39. doi:<https://doi.org/doi:10.3390/jimaging3030039>.
105. Gehan, M.A.; Fahlgren, N.; Abbasi, A.; Berry, J.C.; Callen, S.T.; Chavez, L.; Doust, A.N.; Feldman, M.J.; Gilbert, K.B.; Hodge, J.G.; Hoyer, J.S.; Lin, A.; Liu, S.; Lizárraga, C.; Lorence, A.; Miller, M.; Platon, E.; Tessman, M. and Sax, T. PlantCV v2: Image analysis software for high-throughput plant phenotyping. *PeerJ* **2017**, *5*. doi:<https://doi.org/10.7717/peerj.4088>.
106. Yang, Z. and Han, Y. A Low-Cost 3D Phenotype Measurement Method of Leafy Vegetables Using Video Recordings from Smartphones. *Sensors* **2020**, *20*. doi:<https://dx.doi.org/10.3390/s20216068>.
107. Martinez-Guanter, J.; Ribeiro, Á.; Peteinatos, G.G.; Pérez-Ruiz, M.; Gerhards, R.; Bengochea-Guevara, J.M.; Machleb, J. and Andújar, D. Low-Cost Three-Dimensional Modeling of Crop Plants. *Sensors* **2019**, *19*. doi:<https://dx.doi.org/10.3390/s19132883>.
108. Jay, S.; Rabatel, G.; Hadoux, X.; Moura, D. and Gorretta, N. In-field crop row phenotyping from 3D modeling performed using Structure from Motion. *Comput. Electron. Agric.* **2015**, *110*, 70-77. doi:<https://dx.doi.org/10.1016/j.compag.2014.09.021>.
109. Atali, G.; Garip, Z.; Ozkan, S.S. and Karayel, D. Path Planning of Mobile Robots Based on QR Code. *Academic Perspective Procedia* **2018**, *1*, 31-38. doi:<https://doi.org/10.33793/acperpro.01.01.9>.
110. Lu, Y. and Young, S. A survey of public datasets for computer vision tasks in precision agriculture. *Comput. Electron. Agric.* **2020**, *178*. doi:<https://doi.org/10.1016/j.compag.2020.105760>.
111. Lameski, P.; Zdravevski, E.; Trajkovik, V. and Kulakov, A. Weed Detection Dataset with RGB Images Taken Under Variable Light Conditions. In Proceedings of the International Conference on ICT Innovations, Skopje, Macedonia, 18-23 September, 2017; pp. 112-119. doi:https://dx.doi.org/10.1007/978-3-319-67597-8_11.
112. Jiang, H.; Zhang, C.; Qiao, Y.; Zhang, Z.; Zhang, W. and Song, C. CNN feature based graph convolutional network for weed and crop recognition in smart farming. *Comput. Electron. Agric.* **2020**, *174*. doi:<https://doi.org/10.1016/j.compag.2020.105450>.
113. Sudars, K.; Jasko, J.; Namatevs, I.; Ozola, L. and Badaukis, N. Dataset of annotated food crops and weed images for robotic computer vision control. *Data in Brief* **2020**, *31*, 105833. doi:<https://doi.org/10.1016/j.dib.2020.105833>.
114. Padilla, R.; Passos, W.L.; Dias, T.L.B.; Netto, S.L. and da Silva, E.A.B. A Comparative Analysis of Object Detection Metrics with a Companion Open-Source Toolkit. *Electronics* **2021**, *10*, 279. doi:<https://doi.org/10.3390/electronics10030279>.
115. Shorten, C. and Khoshgoftaar, T.M. A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data* **2019**, *6*. doi:<https://dx.doi.org/10.1186/s40537-019-0197-0>.
116. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.-Y. and Berg, A.C. SSD: Single Shot MultiBox Detector. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 11-14 October, 2016; pp. 21-37. doi:https://doi.org/10.1007/978-3-319-46448-0_2.
117. Redmon, J.; Divvala, S.; Girshick, R. and Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27-30 June, 2016; pp. 779-788. doi:<https://doi.org/10.1109/CVPR.2016.91>.

118. Zhao, Z.; Zheng, P.; Xu, S. and Wu, X. Object Detection With Deep Learning: A Review. *IEEE Transactions on Neural Networks and Learning Systems* **2019**, *30*, 3212-3232. doi:<https://doi.org/10.1109/TNNLS.2018.2876865>.
119. Liu, G.; Nouaze, J.C.; Touko Mbouembe, P.L. and Kim, J.H. YOLO-Tomato: A Robust Algorithm for Tomato Detection Based on YOLOv3. *Sensors* **2020**, *20*, 2145. doi:<https://dx.doi.org/10.3390/s20072145>.
120. Redmon, J.C. Darknet: Open Source Neural Networks in C. Available online: <http://pjreddie.com/darknet/> (accessed on 15/06/2021).
121. Kaya, A.; Keceli, A.S.; Catal, C.; Yalic, H.Y.; Temucin, H. and Tekinerdogan, B. Analysis of transfer learning for deep neural network based plant classification models. *Comput. Electron. Agric.* **2019**, *158*, 20-29. doi:<https://dx.doi.org/10.1016/j.compag.2019.01.041>.
122. Magalhães, S.A.; Castro, L.; Moreira, G.; dos Santos, F.N.; Cunha, M.; Dias, J. and Moreira, A.P. Evaluating the Single-Shot MultiBox Detector and YOLO Deep Learning Models for the Detection of Tomatoes in a Greenhouse. *Sensors* **2021**, *21*, 3569. doi:<https://doi.org/10.3390/s21103569>.
123. Everingham, M.; Eslami, S.M.A.; Van Gool, L.; Williams, C.K.I.; Winn, J. and Zisserman, A. The Pascal Visual Object Classes Challenge: A Retrospective. *International Journal of Computer Vision* **2015**, *111*, 98-136. doi:<https://doi.org/10.1007/s11263-014-0733-5>.
124. Grandini, M.; Bagli, E. and Visani, G. Metrics for multi-class classification: an overview. *arXiv* **2020**. arXiv:2008.05756.
125. Brodersen, K.H.; Ong, C.S.; Stephan, K.E. and Buhmann, J.M. The Balanced Accuracy and Its Posterior Distribution. In Proceedings of the 20th International Conference on Pattern Recognition, Istanbul, Turkey, 23-26 August, 2010; pp. 3121-3124. doi:<https://doi.org/10.1109/ICPR.2010.764>.

Anexo I – Lista de materiais do PixelCropRobot

Tabela I i – Lista de materiais do PixelCropRobot.

Fornecedor	Descrição	Quantidade	Preço Unitário	Sub Total
RatRig	Aluminium Spacer 6 mm	6	0,23 €	1,38 €
	Belt Clamp	2	0,42 €	0,84 €
	Cable Carrier (30*15 mm) 1 m	1	12,14 €	12,14 €
	Cable Carrier (30*15 mm) 1,5 m	1	17,71 €	17,71 €
	Corner Bracket 90 degrees	16	0,65 €	10,40 €
	Eccentric Spacer 6 mm	6	1,50 €	9,00 €
	Flexible Coupling 5*8 mm	1	4,15 €	4,15 €
	Hex Locking Nut M5	12	0,15 €	1,80 €
	Joining Plate 90 degrees	2	2,58 €	5,16 €
	Joining Plate T	4	2,58 €	10,32 €
	Lead Screw TR8*8 1 m	1	29,99 €	29,99 €
	Nema 17 Stepper Motor	2	11,99 €	23,98 €
	Nema 17 Stepper Motor Cable	2	2,85 €	5,70 €
	Nut Block for Lead Screw	1	4,79 €	4,79 €
	Screw M5 8 mm	4 (x25)	2,02 €	8,08 €
	Screw M5 12 mm	8	0,07 €	0,56 €
	Screw M5 30 mm	12	0,13 €	1,56 €
	Timing Belt 2GT 6 mm	1	2,49 €	2,49 €
	Timing Pulley 2GT	1	1,64 €	1,64 €
	T-Nut M5	85	0,20 €	17,00 €
	V-Slot 2020 1 m	1	7,50 €	7,50 €
Mauser	V-Slot 2040 0,5 m	2	5,15 €	10,30 €
	V-Slot 2040 1 m	2	10,30 €	20,60 €
	V-Slot 2060 1,5 m	1	19,37 €	19,37 €
	V-wheels	12	3,90 €	46,80 €
	Controlador DRV8825	2	3,90 €	7,80 €
	Velleman VMA400 (4 relés)	1	10,40 €	10,40 €
	Screw M3 10 mm	22	0,05 €	1,10 €
PTRobotics	Nut M2	4	0,10 €	0,40 €
	Raspberry Pi 4 Model B	1	44,90 €	44,90 €
	Raspberry Pi Camera Module v2	1	30,49 €	30,49 €
	Microswitch R 27 mm	2	1,08 €	2,16 €
Ferragens do Lena	Screw M2 10 mm	4	0,06 €	0,24 €
	SK8 Shaft Support	4	2,21 €	8,84 €
	Linear Shaft Rod 8 mm	1	5,29 €	5,29 €
	KFL08 Flange Bearing 8 mm	4	2,95 €	11,80 €
AndyMark	Screw M6 35 mm	8	0,06 €	0,48 €
	Screw M5 20 mm	6	0,07 €	0,42 €
	Screw M4 50 mm	24	0,05 €	1,20 €
	Screw M3 15 mm	8	0,01 €	0,08 €
	Nut M6	8	0,01 €	0,08 €
	Nut M5	12	0,01 €	0,12 €
	Nut M4	24	0,01 €	0,24 €
Cytron	HiGrip Wheels	2	10,08 €	20,16 €
RS Components	Power Window Motor	2	21,01 €	42,02 €
INESC TEC	12V Battery	1	19,85 €	19,85 €
	Cross-slide Plate	1	20,00 €	20,00 €
	Ligações elétricas	1	10,00 €	10,00 €
	AgloT	1	75,00 €	75,00 €
	Driver PCB	1	16,14 €	16,14 €

Tabela I ii – Lista de materiais do PixelCropRobot (continuação).

Impressão 3D	Motor_wheel_connector	2	20,16 €
	Motor_enclosure	2	
	Motor_enclosure_cover	2	
	Free_wheel_enclosure	2	
	Free_wheel_enclosure_cover	2	
	Free_wheel_fixation_interior	2	
	Free_wheel_fixation_exterior	2	
	StepperZ_base	1	
	StepperY_enclosure	1	
	Microswitch_support	1	
	Belt_end	1	
	Vertical_support	3	
	Horizontal_support	5	
	Camera_support	1	
Total			622,63 €

Anexo II – Desempenho dos modelos SSD e YOLO

Acelga (BEAVV)

Tabela II i – Confiância que otimiza o F1. Parâmetro calculado na amostra de validação sem *augmentation*.

Modelo	Confiância $\geq$	F1
SSD MobileNet v2 (300 × 300)	33%	94%
SSD MobileNet v2 (640 × 640)	46%	92%
SSD Inception v2 (300 × 300)	14%	94%
SSD ResNet 50 (640 × 640)	52%	94%
YOLO v4 (416 × 416)	65%	95%

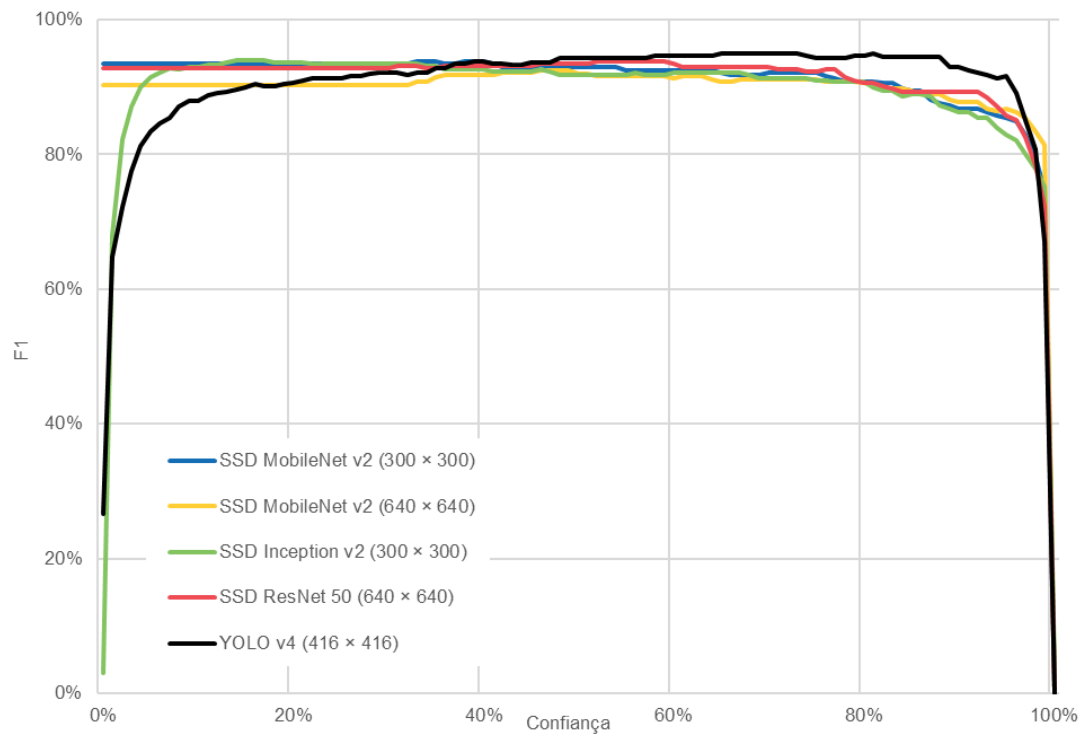


Figura II i – Evolução do F1 com o aumento da confiança, na amostra de validação sem *augmentation*.

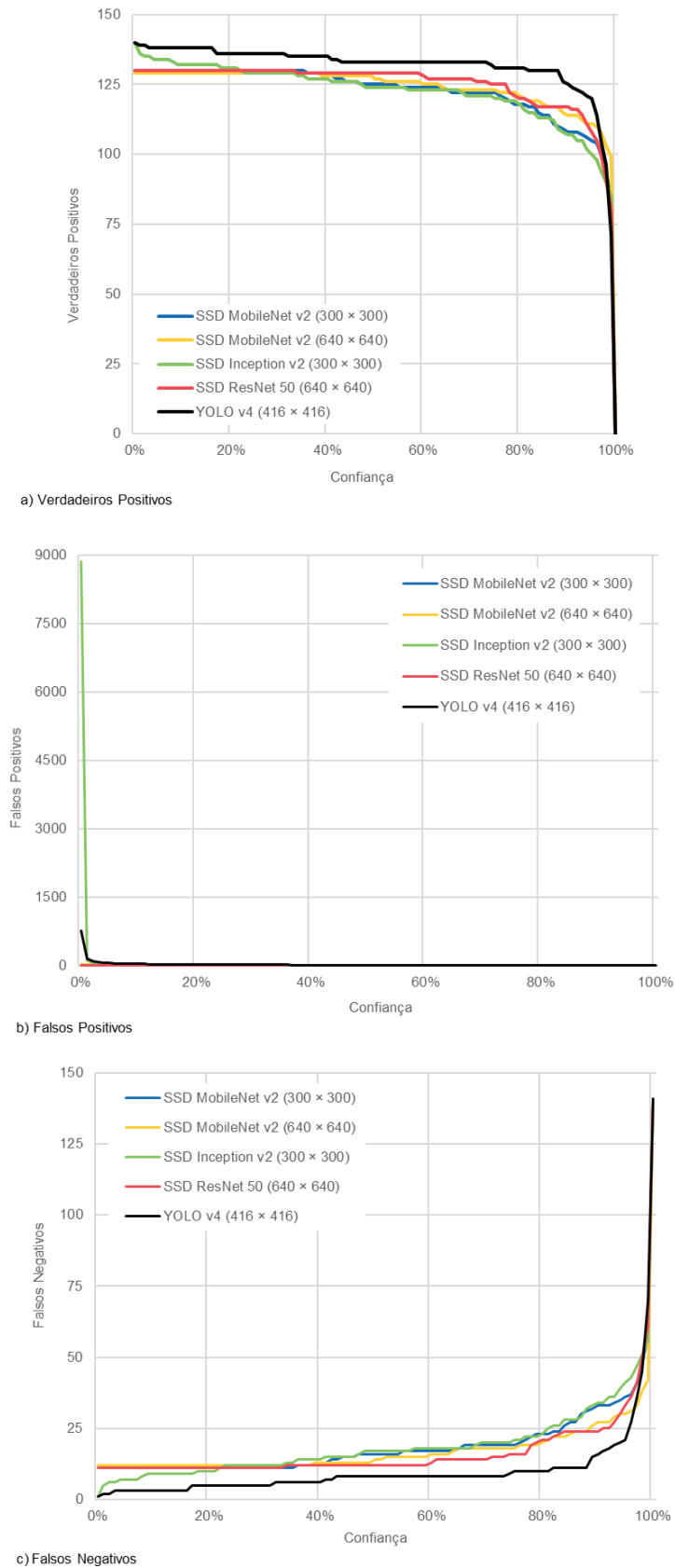
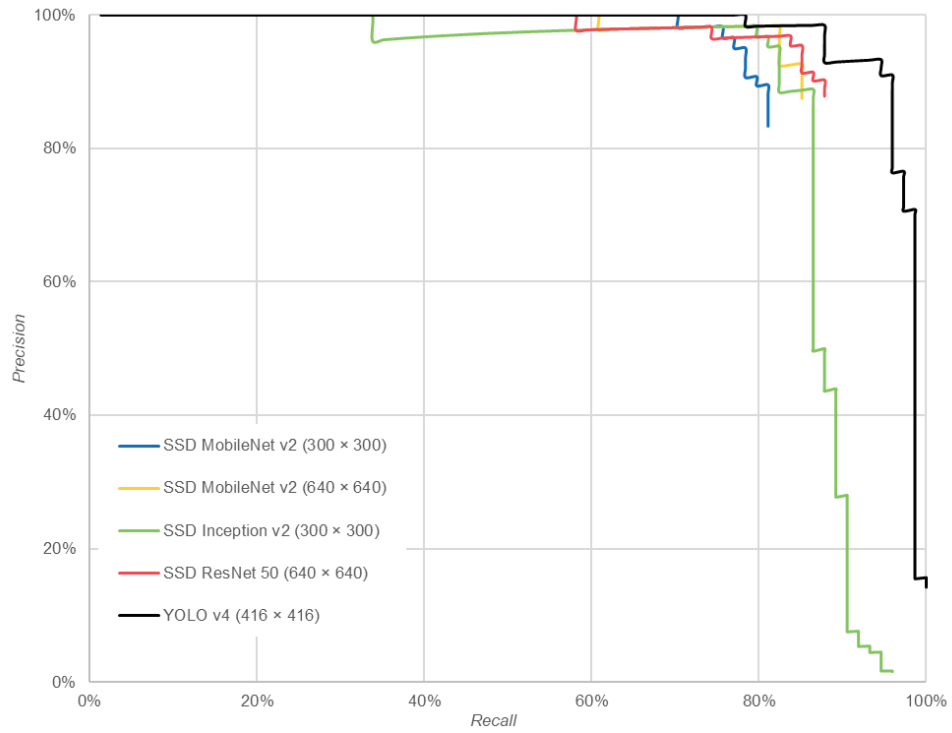


Figura II ii – Evolução do número de: a) Verdadeiros Positivos , b) Falsos Positivos e c) Falsos Negativos com a evolução da confiança, na amostra de validação sem *augmentation*.

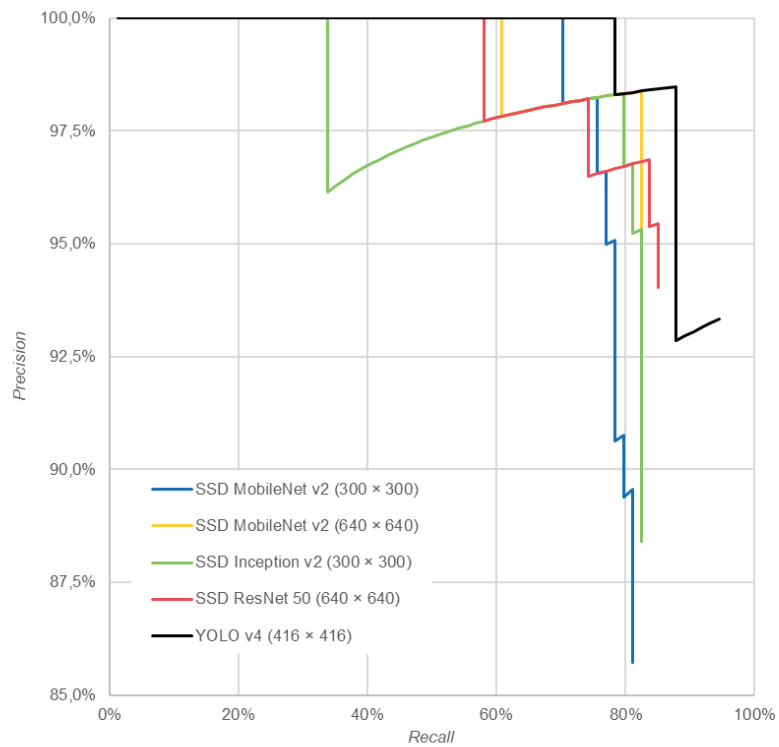
Tabela II ii – Resultados das métricas consideradas para os modelos em estudo. Cálculo executado para todas as deteções e na melhor confiança calculada, na amostra de teste sem *augmentation*.

Modelo	Confiança $\geq$	AP	<i>Precision</i>	<i>Recall</i>	F1
SSD MobileNet v2 (300 × 300)	0%	81%	83%	81%	82%
SSD MobileNet v2 (640 × 640)	0%	85%	88%	85%	86%
SSD Inception v2 (300 × 300)	0%	87%	2%	96%	3%
SSD ResNet 50 (640 × 640)	0%	87%	88%	88%	88%
YOLO v4 (416 × 416)	0%	97%	14%	100%	25%
SSD MobileNet v2 (300 × 300)	33%	81%	86%	81%	83%
SSD MobileNet v2 (640 × 640)	46%	82%	94%	82%	88%
SSD Inception v2 (300 × 300)	14%	82%	88%	82%	85%
SSD ResNet 50 (640 × 640)	52%	84%	94%	85%	89%
YOLO v4 (416 × 416)	65%	94%	93%	95%	94%

AP: *Average Precision*



a) Confiância = 0%



b) Confiância ótima

Figura II iii – Curvas *precision × recall* para a: a) confiança de 0% e b) confiança ótima. Avaliação executada na amostra de teste sem *augmentation*.

Tabela II iii – Matriz de confusão para os modelos SSD e YOLO em estudo para a confiança otimizada na amostra de teste sem *augmentation*.

Conf. = 33%		SSD MobileNet v2 (300 × 300)			
Deteção		Referência		Não referenciada	
		<i>coty</i>	<i>chard</i>	solo	infestante
	<i>coty</i>	6	3	1	0
	<i>chard</i>	0	60	0	0
Conf. = 46%		SSD MobileNet v2 (640 × 640)			
Deteção		Referência		Não referenciada	
		<i>coty</i>	<i>chard</i>	solo	infestante
	<i>coty</i>	3	0	0	0
	<i>chard</i>	2	60	0	0
Conf. = 14%		SSD Inception v2 (300 × 300)			
Deteção		Referência		Não referenciada	
		<i>coty</i>	<i>chard</i>	solo	infestante
	<i>coty</i>	4	0	0	0
	<i>chard</i>	0	60	1	0
Conf. = 52%		SSD ResNet 50 (640 × 640)			
Deteção		Referência		Não referenciada	
		<i>coty</i>	<i>chard</i>	solo	infestante
	<i>coty</i>	4	0	0	0
	<i>chard</i>	0	63	0	0
Conf. = 65%		YOLO v4 (416 × 416)			
Deteção		Referência		Não referenciada	
		<i>coty</i>	<i>chard</i>	solo	infestante
	<i>coty</i>	11	0	0	0
	<i>chard</i>	0	61	0	0

Conf.: confiança que otimiza o valor de F1 na amostra de validação.

Tabela II iv – Resultados das métricas exatidão, *precision* e *recall* para cada classe de alface para os modelos SSD e YOLO em estudo. Avaliação na amostra de teste sem *augmentation*.

<i>coty</i> n = 11	Exatidão	<i>Precision</i>	<i>Recall</i>	F1
MobileNet v2 (300 × 300)	88%	60%	55%	57%
MobileNet v2 (640 × 640)	89%	100%	27%	43%
Inception v2 (300 × 300)	90%	100%	36%	53%
ResNet 50 (640 × 640)	91%	100%	36%	53%
YOLO v4 (416 × 416)	100%	100%	100%	100%

<i>chard</i> n = 63	Exatidão	<i>Precision</i>	<i>Recall</i>	F1
MobileNet v2 (300 × 300)	96%	100%	95%	98%
MobileNet v2 (640 × 640)	93%	97%	95%	96%
Inception v2 (300 × 300)	94%	98%	95%	97%
ResNet 50 (640 × 640)	100%	100%	100%	100%
YOLO v4 (416 × 416)	97%	100%	97%	98%

n: número de referências na amostra de teste.

Cenoura (DAUCS)

Tabela II v – Confiância que otimiza o F1. Parâmetro calculado na amostra de validação sem *augmentation*.

Modelo	Confiância $\geq$	F1
SSD MobileNet v2 (300 × 300)	30%	64%
SSD MobileNet v2 (640 × 640)	30%	70%
SSD Inception v2 (300 × 300)	21%	64%
SSD ResNet 50 (640 × 640)	49%	80%
YOLO v4 (416 × 416)	59%	85%

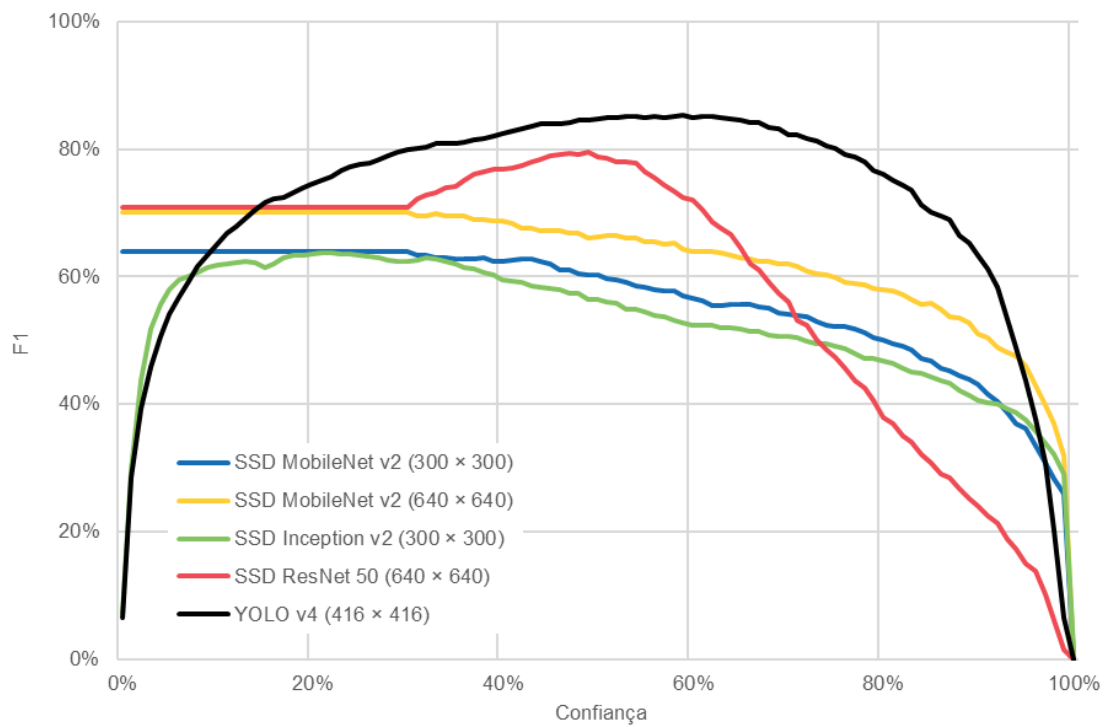


Figura II iv – Evolução do F1 com o aumento da confiança, na amostra de validação sem *augmentation*.

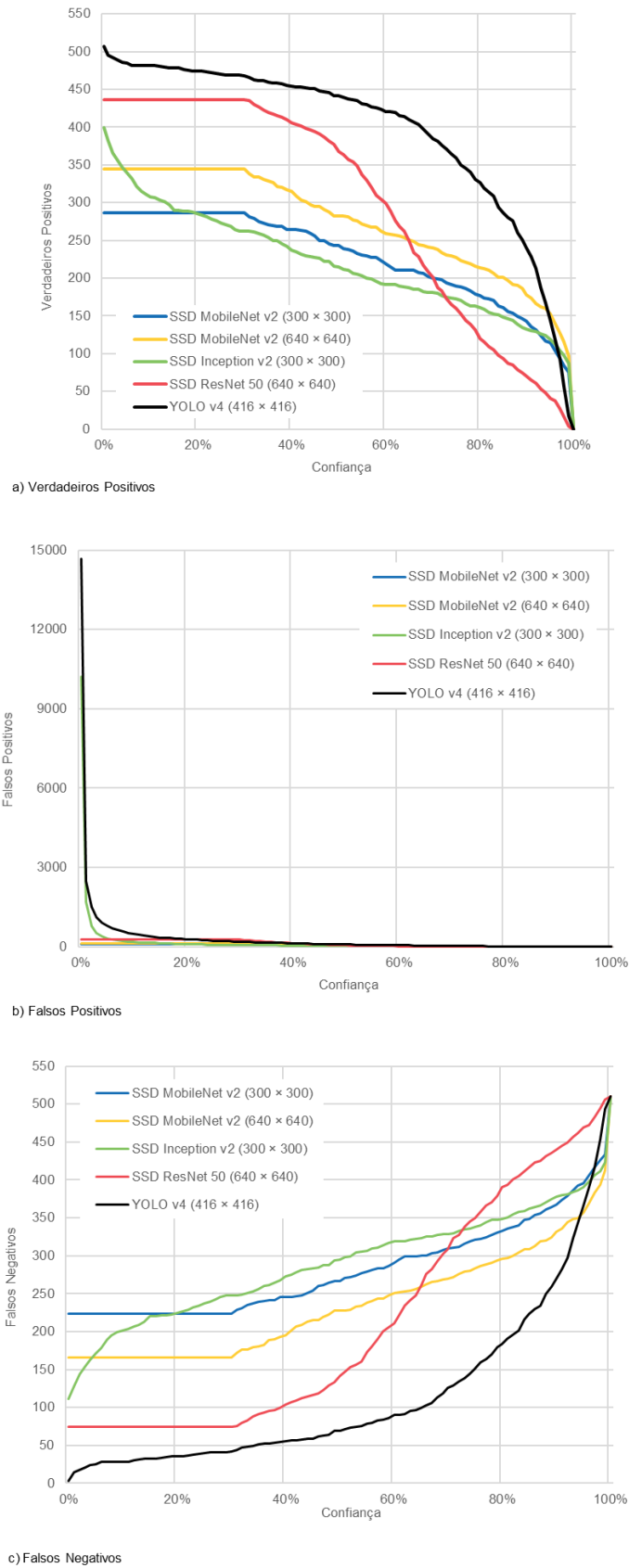
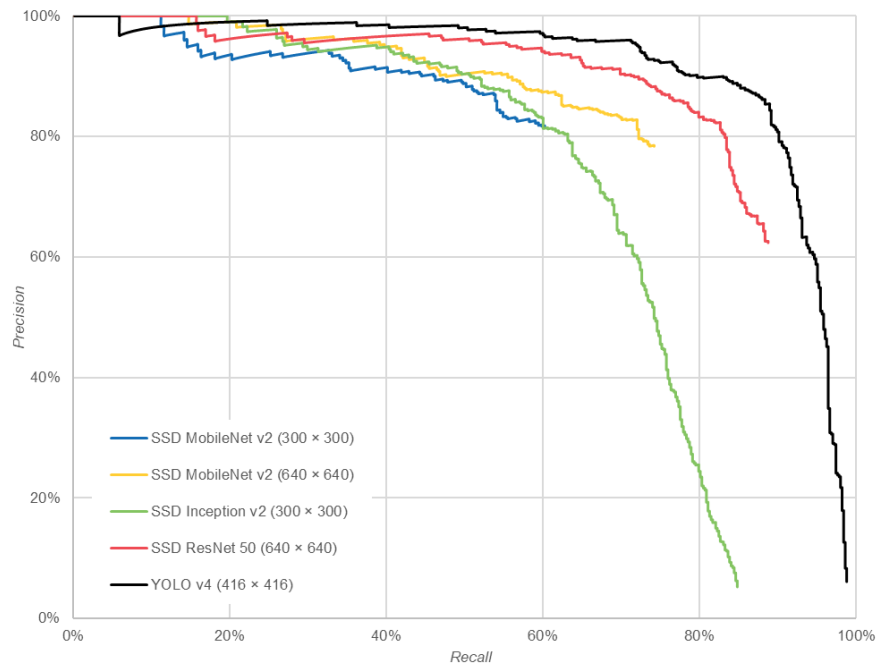


Figura II v – Evolução do número de: a) Verdadeiros Positivos, b) Falsos Positivos e c) Falsos Negativos com a evolução da confiança, na amostra de validação sem *augmentation*.

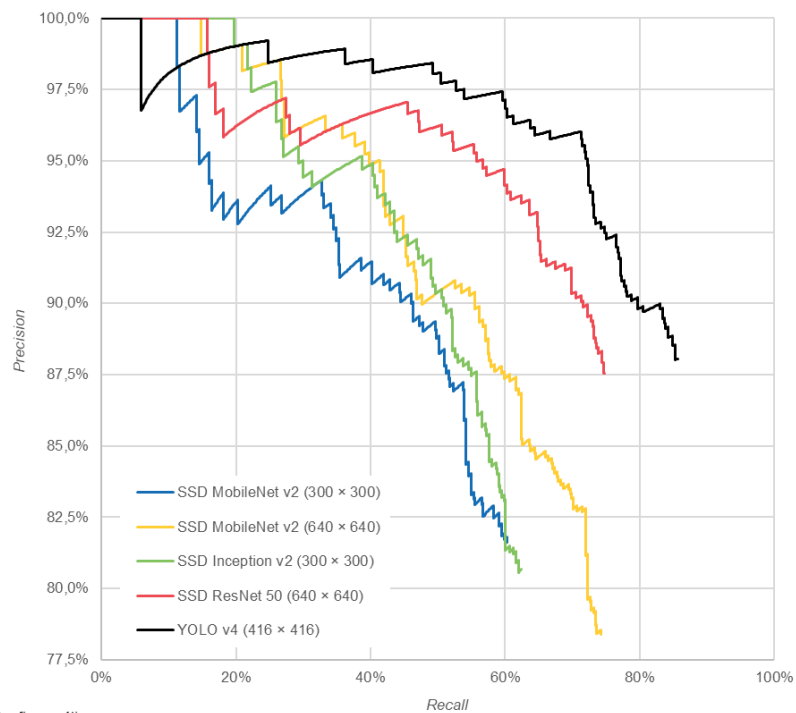
Tabela II vi – Resultados das métricas consideradas para os modelos em estudo. Cálculo executado na amostra de teste sem *augmentation* para todas as deteções e na melhor confiança calculada.

Modelo	Confiança $\geq$	AP	<i>Precision</i>	<i>Recall</i>	F1
SSD MobileNet v2 (300 × 300)	0%	56%	82%	60%	69%
SSD MobileNet v2 (640 × 640)	0%	70%	78%	74%	76%
SSD Inception v2 (300 × 300)	0%	70%	4%	85%	8%
SSD ResNet 50 (640 × 640)	0%	83%	62%	89%	73%
YOLO v4 (416 × 416)	0%	92%	3%	99%	6%
SSD MobileNet v2 (300 × 300)	30%	56%	82%	60%	69%
SSD MobileNet v2 (640 × 640)	30%	70%	78%	74%	76%
SSD Inception v2 (300 × 300)	21%	59%	81%	62%	70%
SSD ResNet 50 (640 × 640)	49%	72%	88%	75%	81%
YOLO v4 (416 × 416)	59%	83%	88%	86%	87%

AP: *Average Precision*



a) Confiança = 0%



b) Confiança ótima

Figura II vi – Curvas *precision × recall* para a: a) confiança de 0% e b) confiança ótima. Avaliação executada na amostra de teste sem *augmentation*.

Tabela II vii – Matriz de confusão para os modelos SSD e YOLO em estudo na amostra de teste sem *augmentation*, para a confiança otimizada.

Conf. = 30%		SSD MobileNet v2 (300 × 300)				
Deteção		Referência			Não referenciada	
		<i>coty</i>	<i>smallleaves</i>	<i>carrot</i>	solo	infestante
	<i>coty</i>	44	12	0	0	4
	<i>smallleaves</i>	17	104	33	1	2
	<i>carrot</i>	0	0	147	0	0
Conf. = 30%		SSD MobileNet v2 (640 × 640)				
Deteção		Referência			Não referenciada	
		<i>coty</i>	<i>smallleaves</i>	<i>carrot</i>	solo	infestante
	<i>coty</i>	67	26	1	0	8
	<i>smallleaves</i>	15	109	10	2	4
	<i>carrot</i>	0	8	206	3	0
Conf. = 21%		SSD Inception v2 (300 × 300)				
Deteção		Referência			Não referenciada	
		<i>coty</i>	<i>smallleaves</i>	<i>carrot</i>	solo	infestante
	<i>coty</i>	70	27	0	2	2
	<i>smallleaves</i>	9	81	4	0	0
	<i>carrot</i>	0	4	176	0	1
Conf. = 49%		SSD ResNet 50 (640 × 640)				
Deteção		Referência			Não referenciada	
		<i>coty</i>	<i>smallleaves</i>	<i>carrot</i>	solo	infestante
	<i>coty</i>	76	20	0	0	2
	<i>smallleaves</i>	15	114	9	0	1
	<i>carrot</i>	0	0	189	0	1
Conf. = 59%		YOLO v4 (416 × 416)				
Deteção		Referência			Não referenciada	
		<i>coty</i>	<i>smallleaves</i>	<i>carrot</i>	solo	infestante
	<i>coty</i>	103	6	0	1	6
	<i>smallleaves</i>	15	139	22	0	1
	<i>carrot</i>	0	0	189	0	1

Conf.: confiança que otimiza o valor de F1 na amostra de validação.

Tabela II viii – Resultados das métricas exatidão, *precision* e *recall* para cada classe de alface para os modelos SSD e YOLO em estudo. Avaliação na amostra de teste sem *augmentation*.

<i>coty</i> n = 131	Exatidão	<i>Precision</i>	<i>Recall</i>	F1
MobileNet v2 (300 × 300)	76%	73%	34%	46%
MobileNet v2 (640 × 640)	81%	66%	51%	58%
Inception v2 (300 × 300)	79%	69%	53%	60%
ResNet 50 (640 × 640)	84%	78%	58%	66%
YOLO v4 (416 × 416)	92%	89%	79%	83%

<i>smallleaves</i> n = 165	Exatidão	<i>Precision</i>	<i>Recall</i>	F1
MobileNet v2 (300 × 300)	72%	66%	63%	65%
MobileNet v2 (640 × 640)	82%	78%	66%	71%
Inception v2 (300 × 300)	77%	86%	49%	63%
ResNet 50 (640 × 640)	83%	82%	69%	75%
YOLO v4 (416 × 416)	87%	79%	84%	81%

<i>carrot</i> n = 212	Exatidão	<i>Precision</i>	<i>Recall</i>	F1
MobileNet v2 (300 × 300)	84%	100%	69%	82%
MobileNet v2 (640 × 640)	96%	95%	97%	96%
Inception v2 (300 × 300)	90%	97%	83%	90%
ResNet 50 (640 × 640)	95%	99%	89%	94%
YOLO v4 (416 × 416)	95%	99%	89%	94%

n: número de referências na amostra de teste.

Coentro (CORSA)

Tabela II ix – Confiança que otimiza o F1. Parâmetro calculado na amostra de validação sem *augmentation*.

Modelo	Confiança $\geq$	F1
SSD MobileNet v2 (300 × 300)	37%	76%
SSD MobileNet v2 (640 × 640)	39%	77%
SSD Inception v2 (300 × 300)	13%	76%
SSD ResNet 50 (640 × 640)	47%	82%
YOLO v4 (416 × 416)	50%	81%

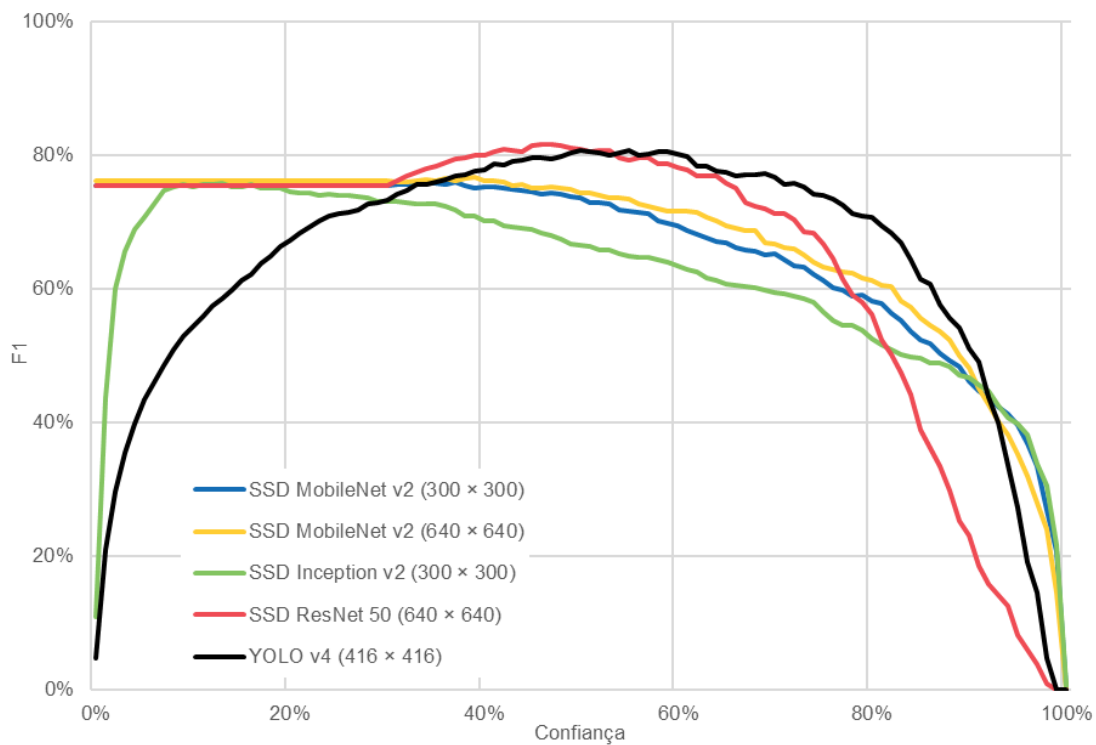


Figura II vii – Evolução do F1 com o aumento da confiança, na amostra de validação sem *augmentation*.

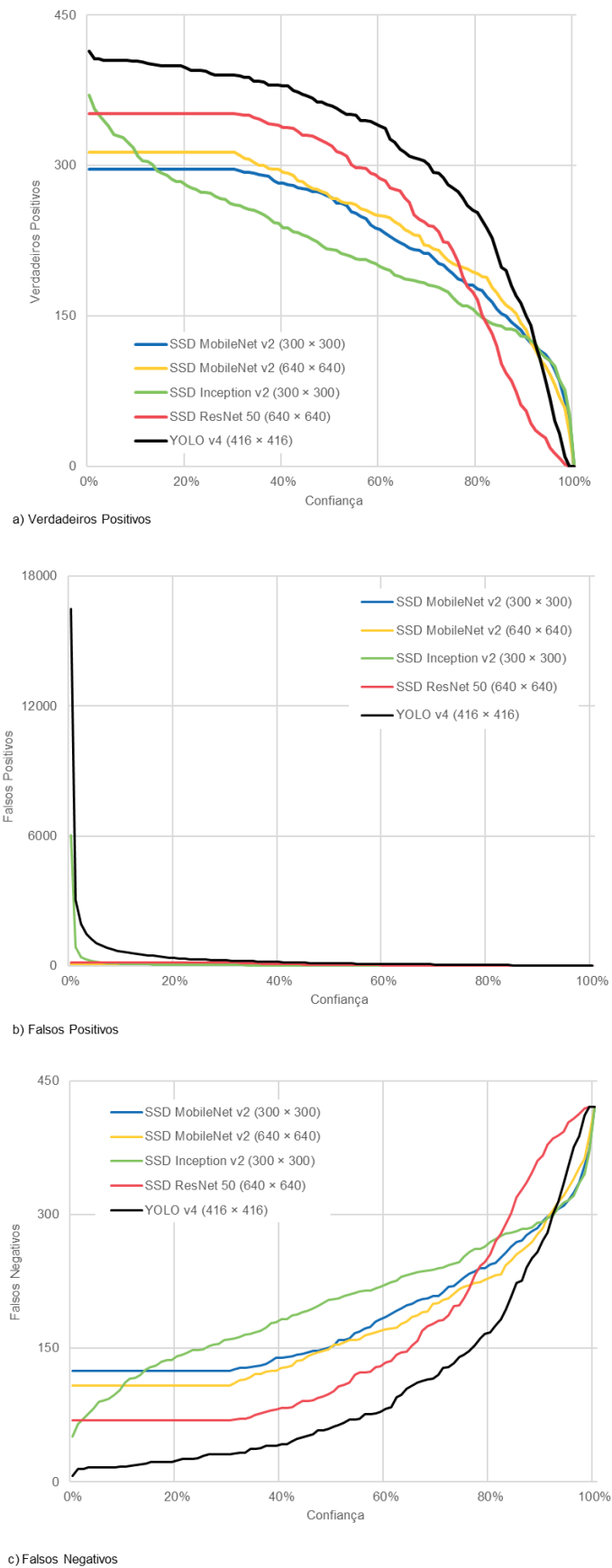
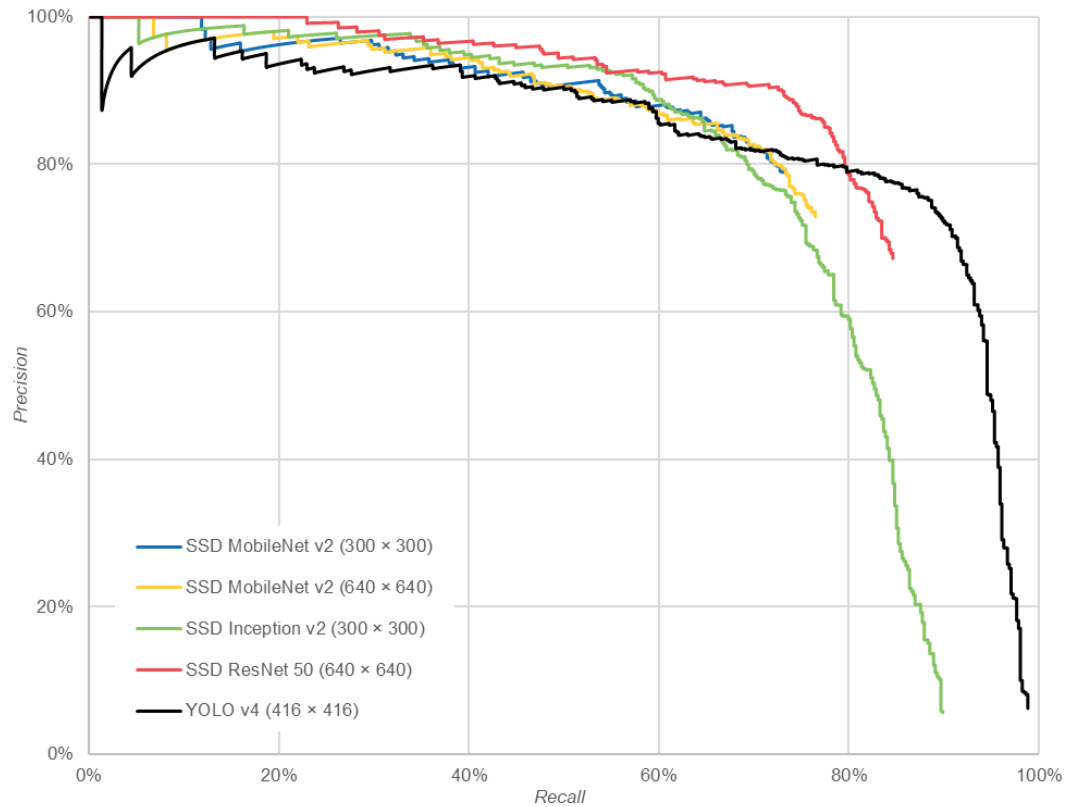


Figura II viii – Evolução do número de: a) Verdadeiros Positivos , b) Falsos Positivos e c) Falsos Negativos com a evolução da confiança, na amostra de validação sem *augmentation*.

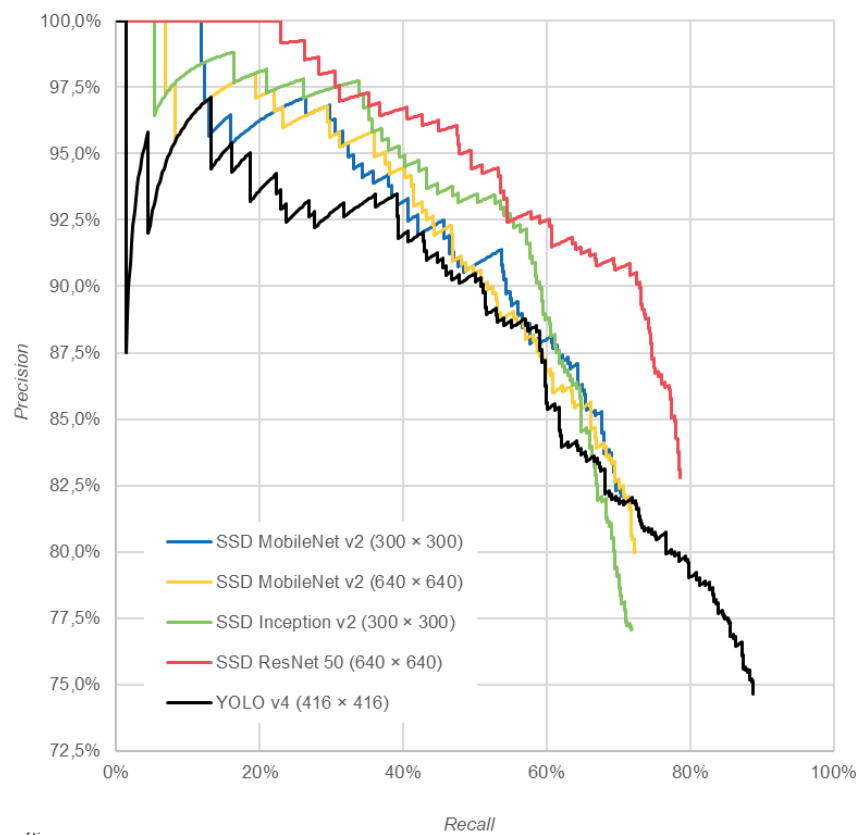
Tabela II x – Resultados das métricas consideradas para os modelos em estudo. Cálculo executado na amostra de teste sem *augmentation* para todas as deteções e na melhor confiança calculada.

Modelo	Confiança $\geq$	AP	<i>Precision</i>	<i>Recall</i>	F1
SSD MobileNet v2 (300 × 300)	0%	68%	79%	73%	76%
SSD MobileNet v2 (640 × 640)	0%	71%	73%	76%	75%
SSD Inception v2 (300 × 300)	0%	77%	5%	90%	9%
SSD ResNet 50 (640 × 640)	0%	80%	67%	85%	75%
YOLO v4 (416 × 416)	0%	84%	2%	99%	4%
SSD MobileNet v2 (300 × 300)	37%	66%	82%	70%	76%
SSD MobileNet v2 (640 × 640)	39%	68%	80%	72%	76%
SSD Inception v2 (300 × 300)	13%	68%	77%	72%	74%
SSD ResNet 50 (640 × 640)	47%	75%	83%	79%	81%
YOLO v4 (416 × 416)	50%	79%	75%	89%	81%

AP: *Average Precision*



a) Confiança = 0%



b) Confiança ótima

Figura II ix – Curvas *precision × recall* para a: a) confiança de 0% e b) confiança ótima. Avaliação executada na amostra de teste sem *augmentation*.

Tabela II xi – Matriz de confusão para os modelos SSD e YOLO em estudo para a confiança otimizada na amostra de teste sem *augmentation*.

Conf. = 37%		SSD MobileNet v2 (300 × 300)				
Deteção		Referência			Não referenciada	
		<i>coty</i>	<i>smallleaves</i>	<i>coriander</i>	solo	infestante
	<i>coty</i>	76	4	0	1	1
	<i>smallleaves</i>	1	87	9	0	0
	<i>coriander</i>	0	9	67	0	0
Conf. = 39%		SSD MobileNet v2 (640 × 640)				
Deteção		Referência			Não referenciada	
		<i>coty</i>	<i>smallleaves</i>	<i>coriander</i>	solo	infestante
	<i>coty</i>	93	14	0	0	2
	<i>smallleaves</i>	6	69	2	0	1
	<i>coriander</i>	0	8	75	0	0
Conf. = 13%		SSD Inception v2 (300 × 300)				
Deteção		Referência			Não referenciada	
		<i>coty</i>	<i>smallleaves</i>	<i>coriander</i>	solo	infestante
	<i>coty</i>	38	1	0	0	0
	<i>smallleaves</i>	13	71	4	0	1
	<i>coriander</i>	0	10	71	0	0
Conf. = 47%		SSD ResNet 50 (640 × 640)				
Deteção		Referência			Não referenciada	
		<i>coty</i>	<i>smallleaves</i>	<i>coriander</i>	solo	infestante
	<i>coty</i>	48	4	0	0	1
	<i>smallleaves</i>	11	79	9	0	1
	<i>coriander</i>	0	9	62	0	0
Conf. = 50%		YOLO v4 (416 × 416)				
Deteção		Referência			Não referenciada	
		<i>coty</i>	<i>smallleaves</i>	<i>coriander</i>	solo	infestante
	<i>coty</i>	104	1	0	0	0
	<i>smallleaves</i>	5	84	2	0	0
	<i>coriander</i>	0	15	72	0	0

Conf.: confiança que otimiza o valor de F1 na amostra de validação.

Tabela II xii – Resultados das métricas exatidão, *precision* e *recall* para cada classe de alface para os modelos SSD e YOLO em estudo. Avaliação na amostra de teste sem *augmentation*.

<i>coty</i> n = 224	Exatidão	<i>Precision</i>	<i>Recall</i>	F1
MobileNet v2 (300 × 300)	78%	81%	67%	73%
MobileNet v2 (640 × 640)	82%	86%	71%	77%
Inception v2 (300 × 300)	82%	80%	82%	81%
ResNet 50 (640 × 640)	87%	88%	82%	85%
YOLO v4 (416 × 416)	86%	83%	84%	84%
<i>smallleaves</i> n = 118	Exatidão	<i>Precision</i>	<i>Recall</i>	F1
MobileNet v2 (300 × 300)	87%	82%	64%	72%
MobileNet v2 (640 × 640)	86%	71%	75%	73%
Inception v2 (300 × 300)	85%	74%	60%	66%
ResNet 50 (640 × 640)	89%	82%	72%	77%
YOLO v4 (416 × 416)	90%	79%	77%	78%
<i>coriander</i> n = 170	Exatidão	<i>Precision</i>	<i>Recall</i>	F1
MobileNet v2 (300 × 300)	91%	92%	84%	87%
MobileNet v2 (640 × 640)	89%	91%	75%	83%
Inception v2 (300 × 300)	90%	94%	75%	84%
ResNet 50 (640 × 640)	91%	89%	85%	87%
YOLO v4 (416 × 416)	98%	96%	96%	96%

n: número de referências na amostra de teste.

Espinafre (SPQOL)

Tabela II xiii – Confiança que otimiza o F1. Parâmetro calculado na amostra de validação sem *augmentation*.

Modelo	Confiança ≥	F1
SSD MobileNet v2 (300 × 300)	37%	98%
SSD MobileNet v2 (640 × 640)	65%	96%
SSD Inception v2 (300 × 300)	32%	93%
SSD ResNet 50 (640 × 640)	41%	98%
YOLO v4 (416 × 416)	32%	96%

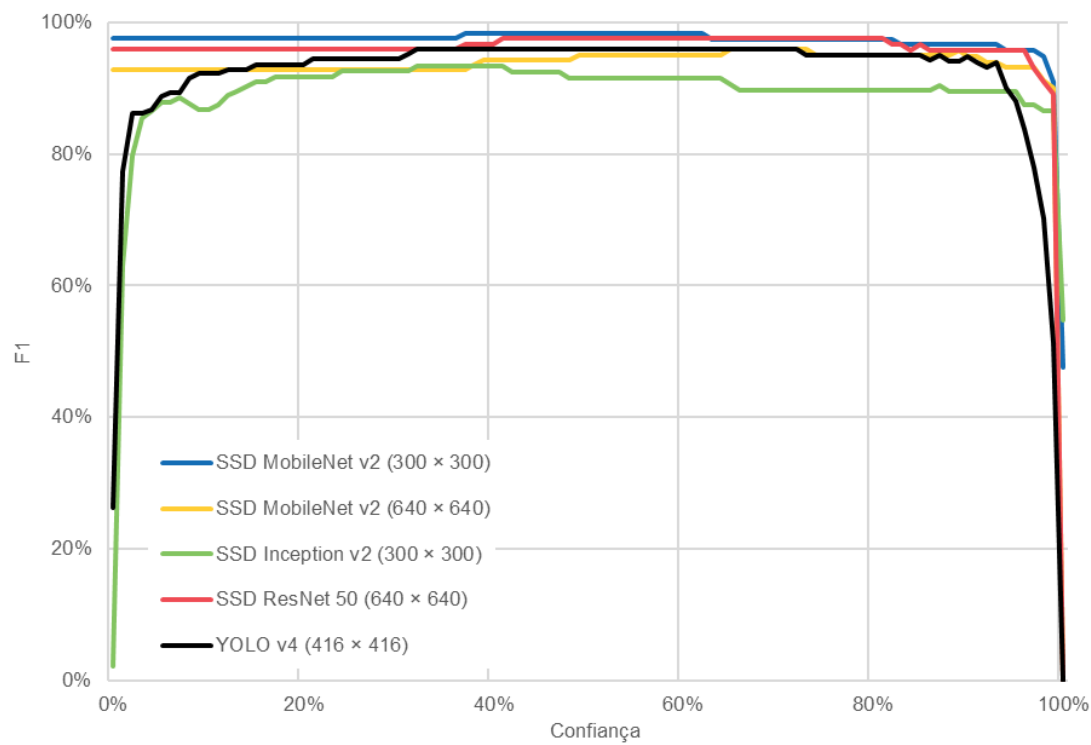


Figura II x – Evolução do F1 com o aumento da confiança, na amostra de validação sem *augmentation*.

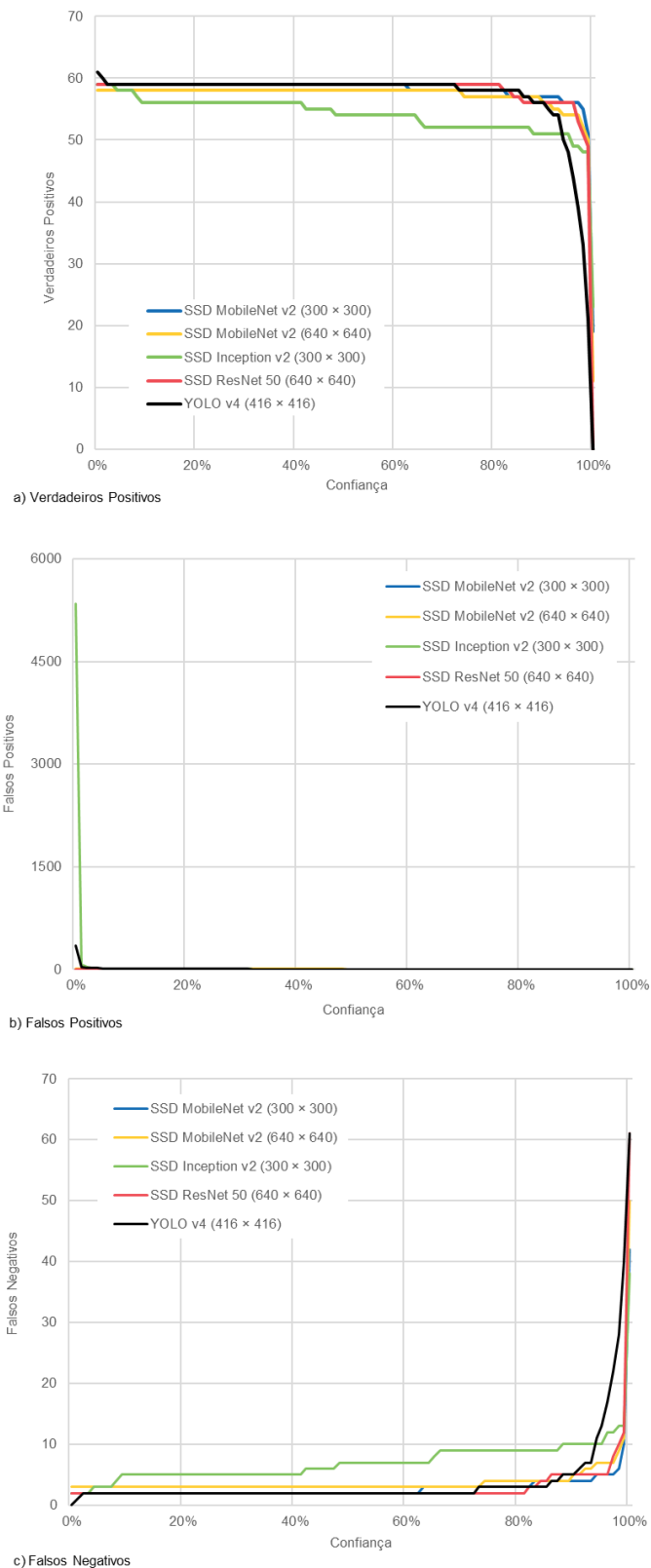
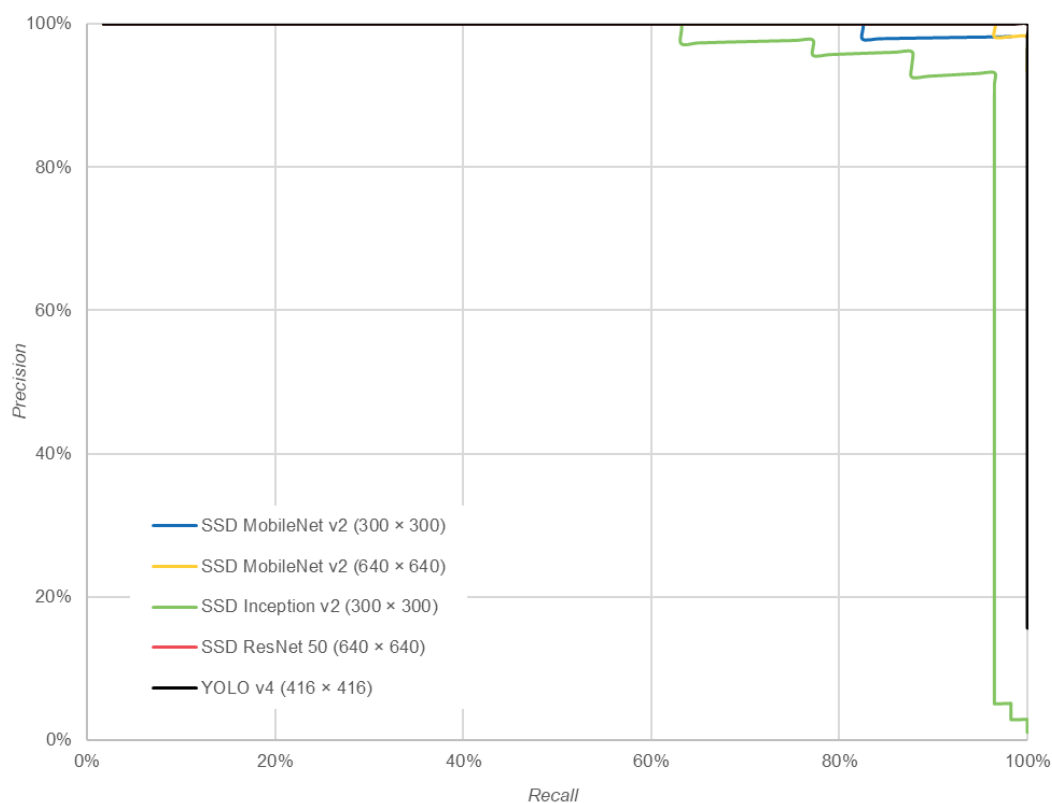


Figura II xi – Evolução do número de: a) Verdadeiros Positivos, b) Falsos Positivos e c) Falsos Negativos com a evolução da confiança, na amostra de validação sem *augmentation*.

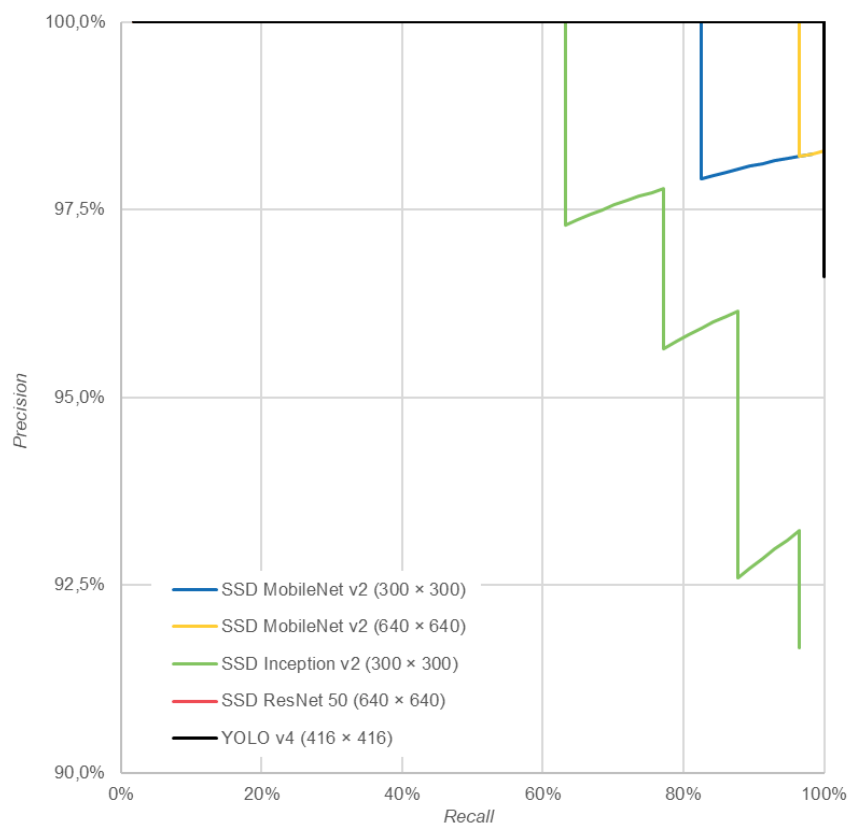
Tabela II xiv – Resultados das métricas consideradas para os modelos em estudo. Cálculo executado na amostra de teste sem *augmentation* para todas as detecções e na melhor confiança calculada.

Modelo	Confiança $\geq$	AP	<i>Precision</i>	<i>Recall</i>	F1
SSD MobileNet v2 (300 × 300)	0%	98%	98%	98%	98%
SSD MobileNet v2 (640 × 640)	0%	100%	93%	100%	97%
SSD Inception v2 (300 × 300)	0%	95%	1%	100%	2%
SSD ResNet 50 (640 × 640)	0%	100%	98%	100%	99%
YOLO v4 (416 × 416)	0%	100%	16%	100%	27%
SSD MobileNet v2 (300 × 300)	37%	98%	98%	98%	98%
SSD MobileNet v2 (640 × 640)	65%	100%	98%	100%	99%
SSD Inception v2 (300 × 300)	32%	95%	92%	96%	94%
SSD ResNet 50 (640 × 640)	41%	100%	100%	100%	100%
YOLO v4 (416 × 416)	32%	100%	97%	100%	98%

AP: *Average Precision*



a) Confiança = 0%



b) Confiança ótima

Figura II xii – Curvas *precision × recall* para a: a) confiança de 0% e b) confiança ótima. Avaliação executada na amostra de teste sem *augmentation*.

Tabela II xv – Matriz de confusão para os modelos SSD e YOLO em estudo na amostra de teste sem *augmentation*, para a confiança otimizada.

Conf. = 37%		SSD MobileNet v2 (300 × 300)			
Deteção		Referência		Não referenciada	
		<i>spinach</i>	<i>big</i>	solo	infestante
	<i>spinach</i>	41	0	0	0
	<i>big</i>	0	16	0	0
Conf. = 65%		SSD MobileNet v2 (640 × 640)			
Deteção		Referência		Não referenciada	
		<i>spinach</i>	<i>big</i>	solo	infestante
	<i>spinach</i>	41	1	0	0
	<i>big</i>	0	16	0	0
Conf. = 32%		SSD Inception v2 (300 × 300)			
Deteção		Referência		Não referenciada	
		<i>spinach</i>	<i>big</i>	solo	infestante
	<i>spinach</i>	39	2	0	1
	<i>big</i>	0	16	0	0
Conf. = 41%		SSD ResNet 50 (640 × 640)			
Deteção		Referência		Não referenciada	
		<i>spinach</i>	<i>big</i>	solo	infestante
	<i>spinach</i>	40	0	0	0
	<i>big</i>	0	16	0	0
Conf. = 32%		YOLO v4 (416 × 416)			
Deteção		Referência		Não referenciada	
		<i>spinach</i>	<i>big</i>	solo	infestante
	<i>spinach</i>	41	0	0	1
	<i>big</i>	0	16	0	0

Conf. = confiança que otimiza o valor de F1 na amostra de validação.

Tabela II xvi – Resultados das métricas exatidão, *precision* e *recall* para cada classe de alface para os modelos SSD e YOLO em estudo. Avaliação na amostra de teste sem *augmentation*.

<i>spinach</i> n = 41	Exatidão	<i>Precision</i>	<i>Recall</i>	F1
MobileNet v2 (300 × 300)	100%	100%	100%	100%
MobileNet v2 (640 × 640)	98%	98%	100%	99%
Inception v2 (300 × 300)	92%	93%	95%	94%
ResNet 50 (640 × 640)	98%	100%	98%	99%
YOLO v4 (416 × 416)	98%	98%	100%	99%

<i>big</i> n = 16	Exatidão	<i>Precision</i>	<i>Recall</i>	F1
MobileNet v2 (300 × 300)	100%	100%	100%	100%
MobileNet v2 (640 × 640)	100%	100%	100%	100%
Inception v2 (300 × 300)	100%	100%	100%	100%
ResNet 50 (640 × 640)	100%	100%	100%	100%
YOLO v4 (416 × 416)	100%	100%	100%	100%

n: número de referências na amostra de teste.

Nabo (BRSRR)

Tabela II xvii – Confiança que otimiza o F1. Parâmetro calculado na amostra de validação sem *augmentation*.

Modelo	Confiança ≥	F1
SSD MobileNet v2 (300 × 300)	32%	74%
SSD MobileNet v2 (640 × 640)	36%	71%
SSD Inception v2 (300 × 300)	24%	70%
SSD ResNet 50 (640 × 640)	57%	77%
YOLO v4 (416 × 416)	54%	85%

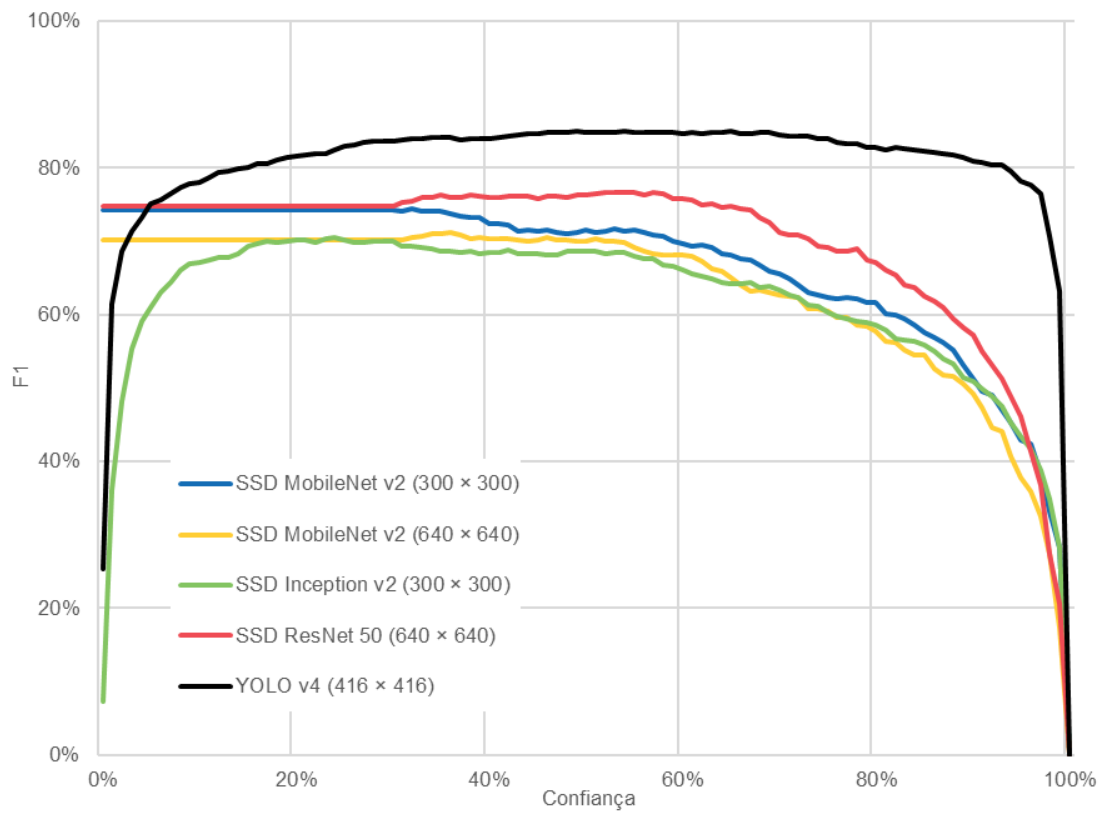


Figura II xiii – Evolução do F1 com o aumento da confiança, na amostra de validação sem *augmentation*.

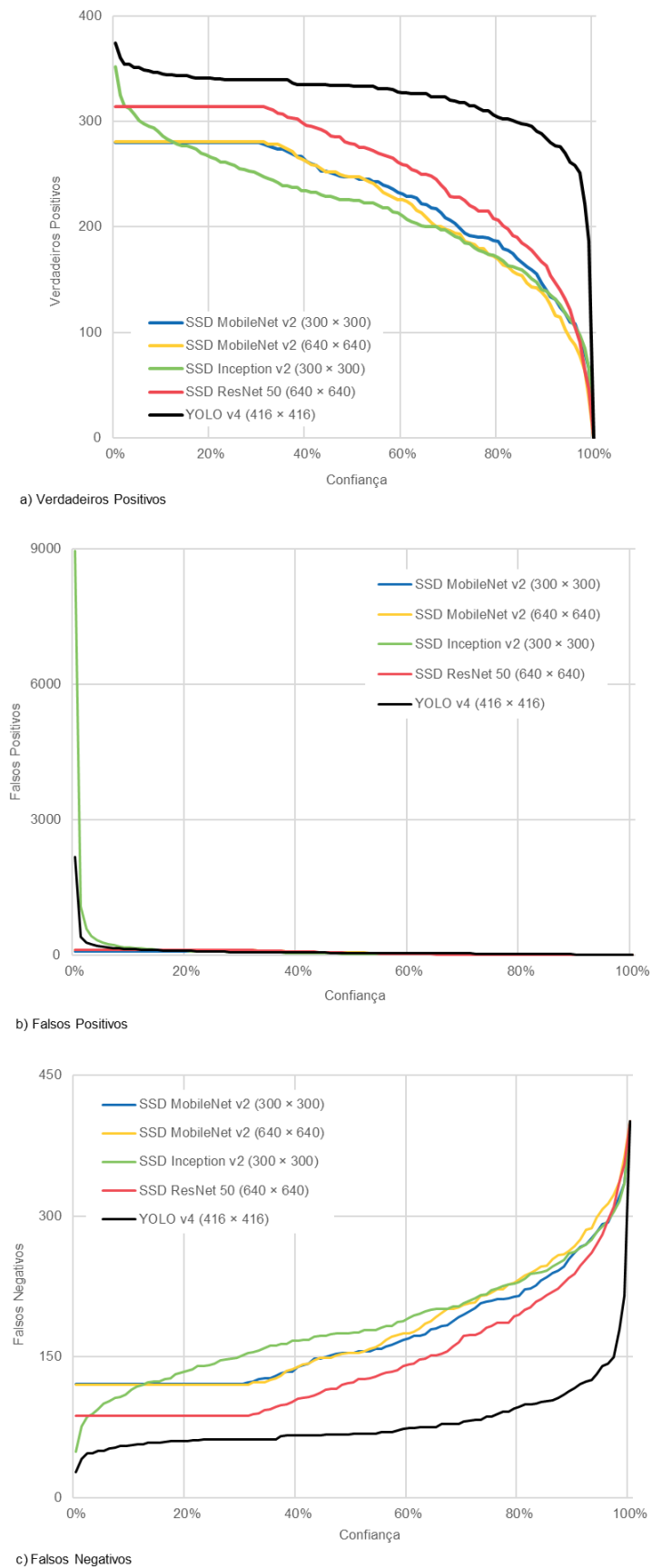
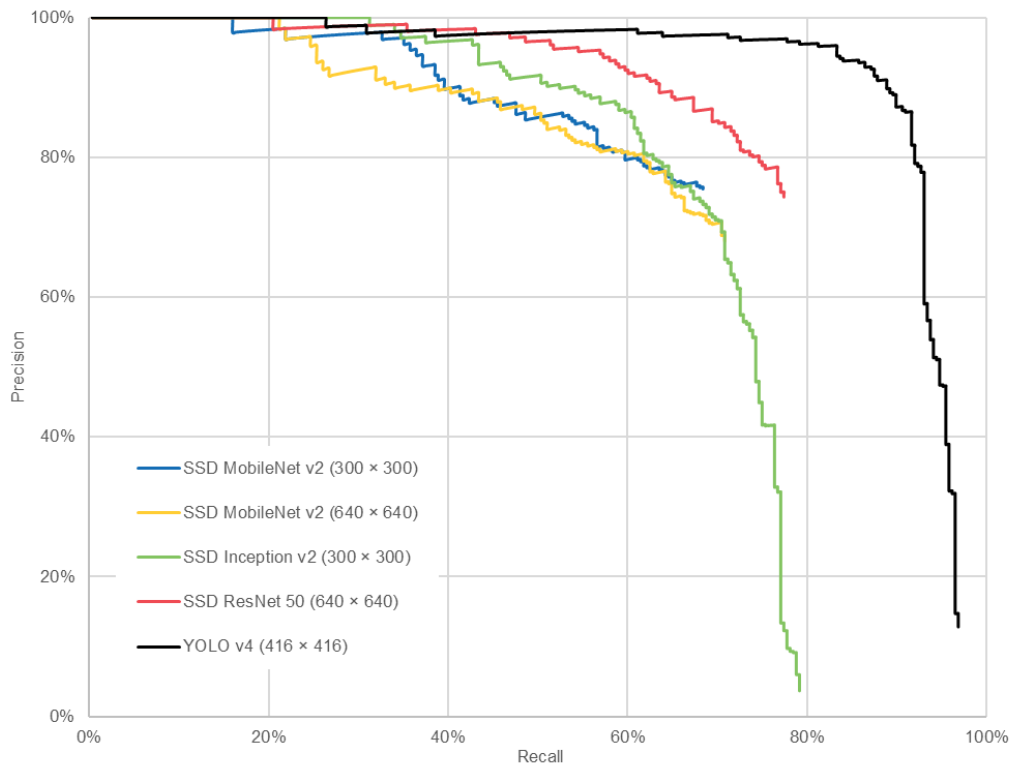


Figura II xiv – Evolução do número de: a) Verdadeiros Positivos , b) Falsos Positivos e c) Falsos Negativos com a evolução da confiança, na amostra de validação sem *augmentation*.

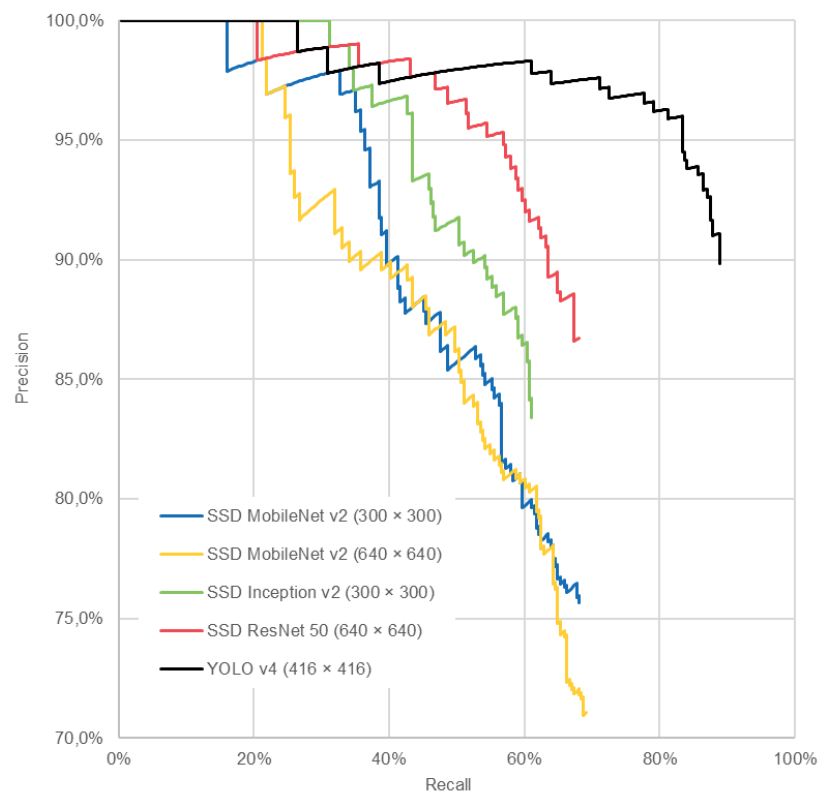
Tabela II xviii – Resultados das métricas consideradas para os modelos em estudo. Cálculo executado na amostra de teste sem *augmentation* para todas as detecções e na melhor confiança calculada.

Modelo	Confiança $\geq$	AP	Precision	Recall	F1
SSD MobileNet v2 (300 × 300)	0%	63%	75%	68%	72%
SSD MobileNet v2 (640 × 640)	0%	64%	68%	71%	69%
SSD Inception v2 (300 × 300)	0%	70%	4%	79%	7%
SSD ResNet 50 (640 × 640)	0%	74%	74%	77%	76%
YOLO v4 (416 × 416)	0%	93%	13%	97%	23%
SSD MobileNet v2 (300 × 300)	32%	63%	76%	68%	72%
SSD MobileNet v2 (640 × 640)	36%	63%	71%	69%	70%
SSD Inception v2 (300 × 300)	24%	59%	83%	61%	71%
SSD ResNet 50 (640 × 640)	57%	66%	87%	68%	76%
YOLO v4 (416 × 416)	54%	87%	90%	89%	89%

AP: Average Precision



a) Confiança = 0%



b) Confiança ótima

Figura II xv – Curvas *precision × recall* para a: a) confiança de 0% e b) confiança ótima. Avaliação executada na amostra de teste sem *augmentation*.

Tabela II xix – Matriz de confusão para os modelos SSD e YOLO em estudo na amostra de teste sem *augmentation*, para a confiança otimizada.

Conf. = 32%		SSD MobileNet v2 (300 × 300)				
Deteção		Referência			Não referenciada	
		<i>coty</i>	<i>smallleaves</i>	<i>turnip</i>	solo	infestante
	<i>coty</i>	76	4	0	1	1
	<i>smallleaves</i>	1	87	9	0	0
	<i>turnip</i>	0	9	67	0	0
Conf. = 36%		SSD MobileNet v2 (640 × 640)				
Deteção		Referência			Não referenciada	
		<i>coty</i>	<i>smallleaves</i>	<i>turnip</i>	solo	infestante
	<i>coty</i>	93	14	0	0	2
	<i>smallleaves</i>	6	69	2	0	1
	<i>turnip</i>	0	8	75	0	0
Conf. = 24%		SSD Inception v2 (300 × 300)				
Deteção		Referência			Não referenciada	
		<i>coty</i>	<i>smallleaves</i>	<i>turnip</i>	solo	infestante
	<i>coty</i>	38	1	0	0	0
	<i>smallleaves</i>	13	71	4	0	1
	<i>turnip</i>	0	10	71	0	0
Conf. = 57%		SSD ResNet 50 (640 × 640)				
Deteção		Referência			Não referenciada	
		<i>coty</i>	<i>smallleaves</i>	<i>turnip</i>	solo	infestante
	<i>coty</i>	48	4	0	0	1
	<i>smallleaves</i>	11	79	9	0	1
	<i>turnip</i>	0	9	62	0	0
Conf. = 54%		YOLO v4 (416 × 416)				
Deteção		Referência			Não referenciada	
		<i>coty</i>	<i>smallleaves</i>	<i>turnip</i>	solo	infestante
	<i>coty</i>	104	1	0	0	0
	<i>smallleaves</i>	5	84	2	0	0
	<i>turnip</i>	0	15	72	0	0

Conf.: confiança que otimiza o valor de F1 na amostra de validação.

Tabela II xx – Resultados das métricas exatidão, *precision* e *recall* para cada classe de alface para os modelos SSD e YOLO em estudo. Avaliação na amostra de teste sem *augmentation*.

<i>coty</i> n = 116	Exatidão	<i>Precision</i>	<i>Recall</i>	F1
MobileNet v2 (300 × 300)	84%	93%	66%	77%
MobileNet v2 (640 × 640)	86%	85%	80%	83%
Inception v2 (300 × 300)	71%	97%	33%	49%
ResNet 50 (640 × 640)	74%	91%	41%	57%
YOLO v4 (416 × 416)	96%	99%	90%	94%

<i>smallleaves</i> n = 92	Exatidão	<i>Precision</i>	<i>Recall</i>	F1
MobileNet v2 (300 × 300)	94%	90%	95%	92%
MobileNet v2 (640 × 640)	88%	88%	75%	81%
Inception v2 (300 × 300)	82%	80%	77%	78%
ResNet 50 (640 × 640)	85%	79%	86%	82%
YOLO v4 (416 × 416)	95%	92%	91%	92%

<i>turnip</i> n = 80	Exatidão	<i>Precision</i>	<i>Recall</i>	F1
MobileNet v2 (300 × 300)	92%	88%	84%	86%
MobileNet v2 (640 × 640)	95%	90%	94%	92%
Inception v2 (300 × 300)	91%	88%	89%	88%
ResNet 50 (640 × 640)	88%	87%	78%	82%
YOLO v4 (416 × 416)	92%	83%	90%	86%

n: número de referências na amostra de teste.

Rabanete (RAPSR)

Tabela II xxi – Confiância que otimiza o F1. Parâmetro calculado na amostra de validação sem *augmentation*.

Modelo	Confiância $\geq$	F1
SSD MobileNet v2 (300 × 300)	34%	86%
SSD MobileNet v2 (640 × 640)	37%	84%
SSD Inception v2 (300 × 300)	24%	88%
SSD ResNet 50 (640 × 640)	55%	93%
YOLO v4 (416 × 416)	66%	92%

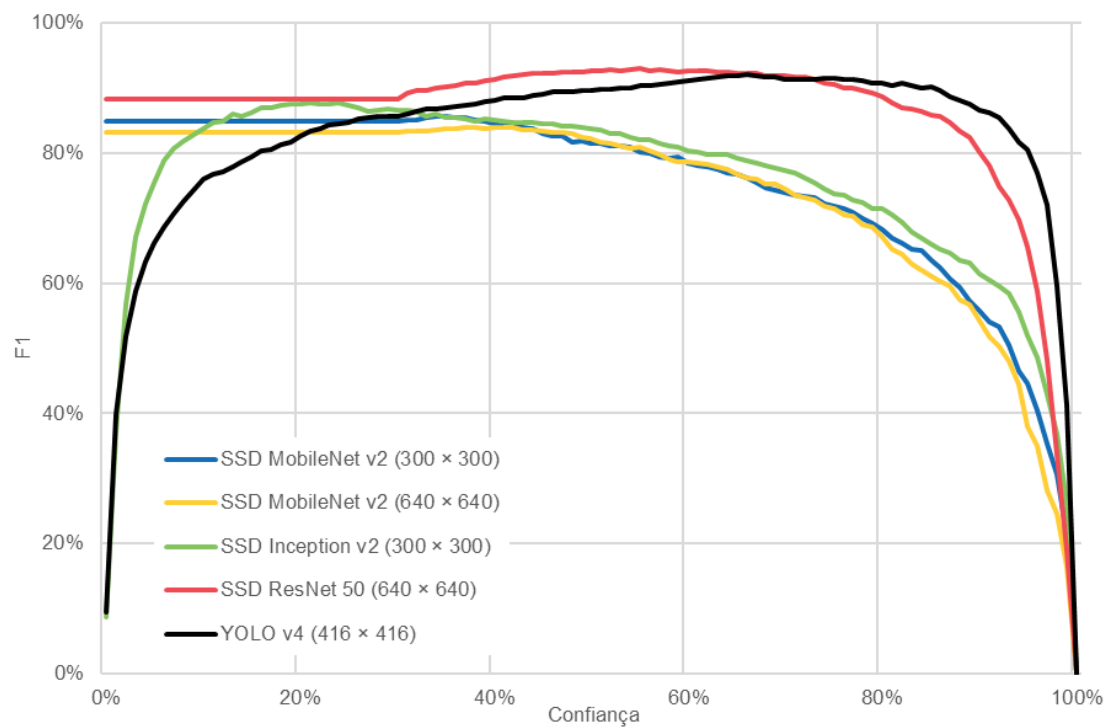


Figura II xvi – Evolução do F1 com o aumento da confiança, na amostra de validação sem *augmentation*.

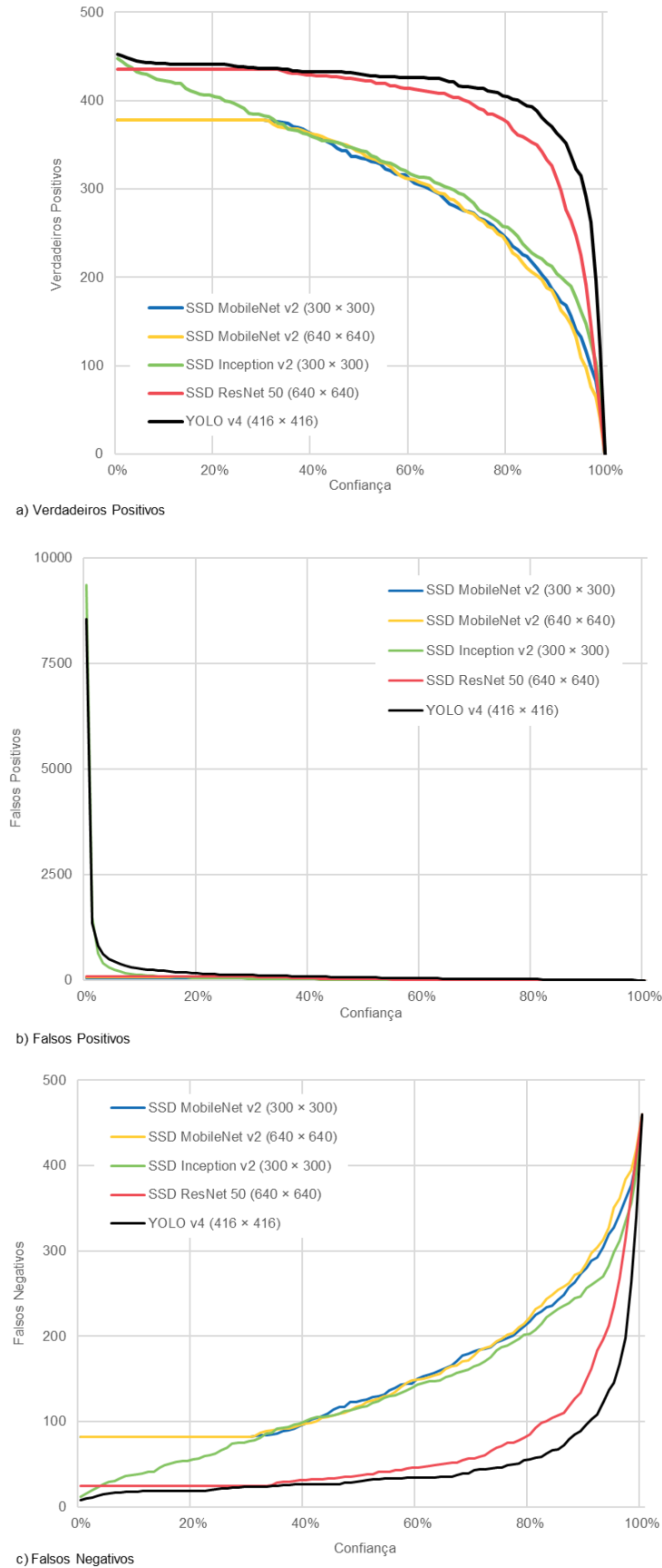
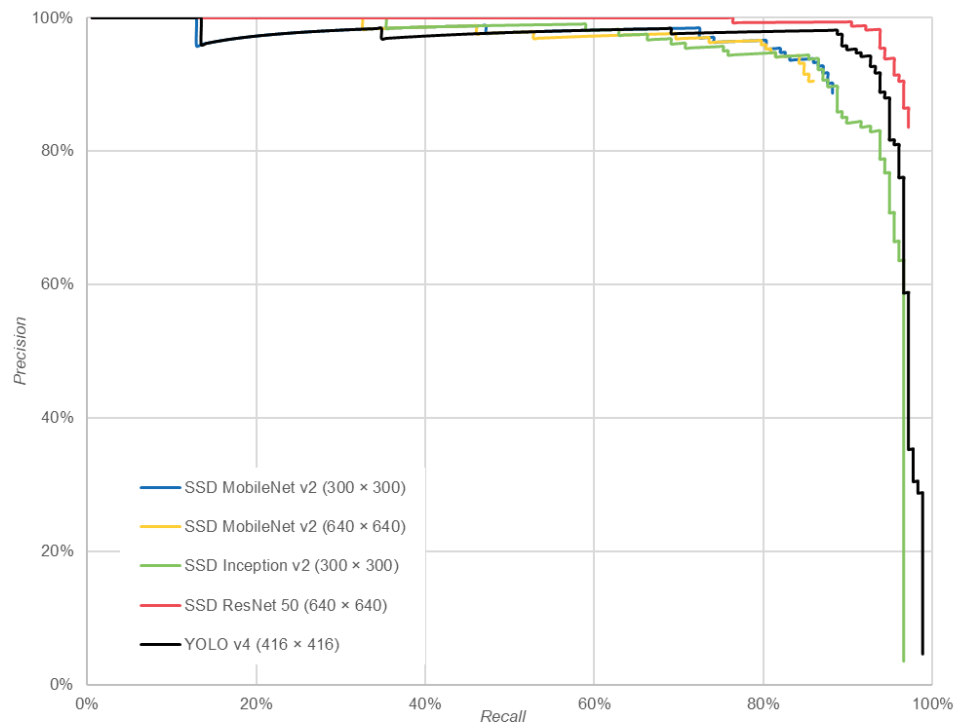


Figura II xvii – Evolução do número de: a) Verdadeiros Positivos , b) Falsos Positivos e c) Falsos Negativos com a evolução da confiança, na amostra de validação sem *augmentation*.

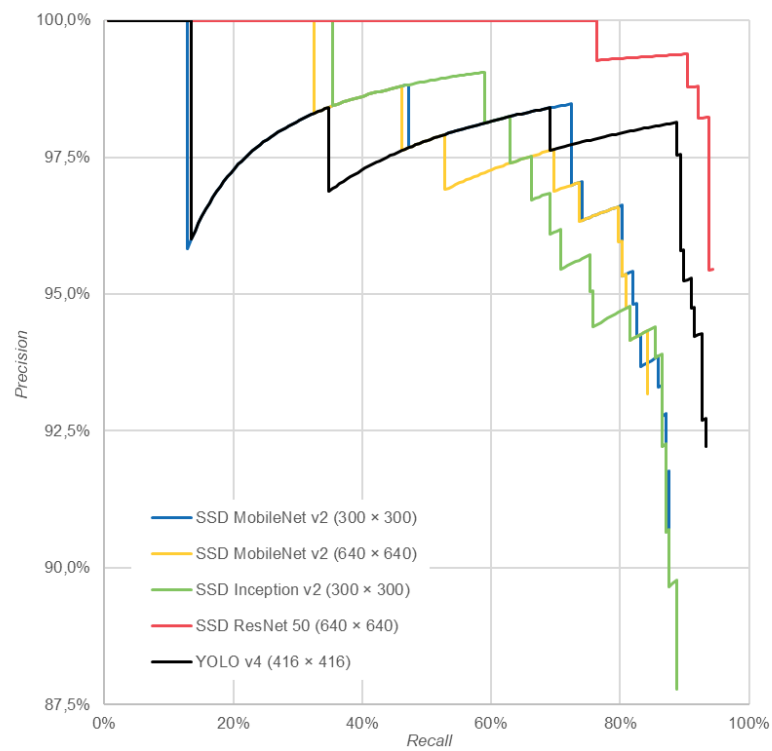
Tabela II xxii – Resultados das métricas consideradas para os modelos em estudo. Cálculo executado na amostra de teste sem *augmentation* para todas as detecções e na melhor confiança calculada.

Modelo	Confiança $\geq$	AP	<i>Precision</i>	<i>Recall</i>	F1
SSD MobileNet v2 (300 × 300)	0%	87%	89%	88%	88%
SSD MobileNet v2 (640 × 640)	0%	85%	91%	86%	88%
SSD Inception v2 (300 × 300)	0%	93%	4%	97%	7%
SSD ResNet 50 (640 × 640)	0%	97%	84%	97%	90%
YOLO v4 (416 × 416)	0%	95%	5%	99%	9%
SSD MobileNet v2 (300 × 300)	34%	86%	90%	88%	89%
SSD MobileNet v2 (640 × 640)	37%	83%	93%	84%	89%
SSD Inception v2 (300 × 300)	24%	87%	88%	89%	88%
SSD ResNet 50 (640 × 640)	55%	94%	95%	94%	95%
YOLO v4 (416 × 416)	66%	92%	92%	93%	93%

AP: *Average Precision*



a) Confiança = 0%



b) Confiança ótima

Figura II xviii – Curvas *precision × recall* para a: a) confiança de 0% e b) confiança ótima. Avaliação executada na amostra de teste sem *augmentation*.

Tabela II xxiii – Matriz de confusão para os modelos SSD e YOLO em estudo na amostra de teste sem *augmentation*, para a confiança otimizada.

Conf. = 34%		SSD MobileNet v2 (300 × 300)					
Deteção		Referência				Não referenciada	
		<i>coty</i>	<i>smallleaves</i>	<i>bigleaves</i>	<i>root</i>	solo	infestante
	<i>coty</i>	118	9	0	0	1	1
	<i>smallleaves</i>	1	529	35	30	1	14
	<i>bigleaves</i>	0	3	54	46	0	0
	<i>root</i>	0	0	0	24	0	0
Conf. = 37%		SSD MobileNet v2 (640 × 640)					
Deteção		Referência				Não referenciada	
		<i>coty</i>	<i>smallleaves</i>	<i>bigleaves</i>	<i>root</i>	solo	infestante
	<i>coty</i>	53	0	0	0	1	3
	<i>smallleaves</i>	17	519	7	2	3	27
	<i>bigleaves</i>	0	2	68	4	0	0
	<i>root</i>	0	0	15	116	0	0
Conf. = 24%		SSD Inception v2 (300 × 300)					
Deteção		Referência				Não referenciada	
		<i>coty</i>	<i>smallleaves</i>	<i>bigleaves</i>	<i>root</i>	solo	infestante
	<i>coty</i>	123	21	0	0	1	5
	<i>smallleaves</i>	0	476	14	2	0	11
	<i>bigleaves</i>	0	24	66	23	0	1
	<i>root</i>	0	1	28	70	0	0
Conf. = 55%		SSD ResNet 50 (640 × 640)					
Deteção		Referência				Não referenciada	
		<i>coty</i>	<i>smallleaves</i>	<i>bigleaves</i>	<i>root</i>	solo	infestante
	<i>coty</i>	61	13	0	0	0	0
	<i>smallleaves</i>	0	522	10	2	0	14
	<i>bigleaves</i>	0	9	57	26	0	0
	<i>root</i>	0	0	7	28	0	0
Conf. = 66%		YOLO v4 (416 × 416)					
Deteção		Referência				Não referenciada	
		<i>coty</i>	<i>smallleaves</i>	<i>bigleaves</i>	<i>root</i>	solo	infestante
	<i>coty</i>	138	44	0	0	4	9
	<i>smallleaves</i>	0	500	20	6	0	10
	<i>bigleaves</i>	0	6	11	2	0	0
	<i>root</i>	0	0	25	30	0	0

Conf.: confiança que otimiza o valor de F1 na amostra de validação.

Tabela II xxiv – Resultados das métricas exatidão, *precision* e *recall* para cada classe de alface para os modelos SSD e YOLO em estudo. Avaliação na amostra de teste sem *augmentation*.

<i>coty</i> n = 53	Exatidão	<i>Precision</i>	<i>Recall</i>	F1
MobileNet v2 (300 × 300)	96%	100%	87%	93%
MobileNet v2 (640 × 640)	96%	98%	89%	93%
Inception v2 (300 × 300)	94%	92%	89%	90%
ResNet 50 (640 × 640)	98%	100%	94%	97%
YOLO v4 (416 × 416)	97%	96%	94%	95%
<i>smallleaves</i> n = 52	Exatidão	<i>Precision</i>	<i>Recall</i>	F1
MobileNet v2 (300 × 300)	96%	92%	94%	93%
MobileNet v2 (640 × 640)	95%	92%	92%	92%
Inception v2 (300 × 300)	99%	98%	100%	99%
ResNet 50 (640 × 640)	99%	98%	100%	99%
YOLO v4 (416 × 416)	99%	98%	100%	99%
<i>bigleaves</i> n = 39	Exatidão	<i>Precision</i>	<i>Recall</i>	F1
MobileNet v2 (300 × 300)	98%	100%	92%	96%
MobileNet v2 (640 × 640)	98%	100%	90%	95%
Inception v2 (300 × 300)	97%	100%	87%	93%
ResNet 50 (640 × 640)	98%	100%	92%	96%
YOLO v4 (416 × 416)	99%	100%	95%	97%
<i>root</i> n = 34	Exatidão	<i>Precision</i>	<i>Recall</i>	F1
MobileNet v2 (300 × 300)	100%	100%	100%	100%
MobileNet v2 (640 × 640)	94%	100%	71%	83%
Inception v2 (300 × 300)	98%	94%	97%	96%
ResNet 50 (640 × 640)	99%	100%	97%	99%
YOLO v4 (416 × 416)	97%	100%	82%	90%

n: número de referências na amostra de teste.

Rúcula (ERUVE)

Tabela II xxv – Confiança que otimiza o F1. Parâmetro calculado na amostra de validação sem *augmentation*.

Modelo	Confiança $\geq$	F1
SSD MobileNet v2 (300 × 300)	31%	82%
SSD MobileNet v2 (640 × 640)	39%	80%
SSD Inception v2 (300 × 300)	22%	88%
SSD ResNet 50 (640 × 640)	60%	90%
YOLO v4 (416 × 416)	39%	91%

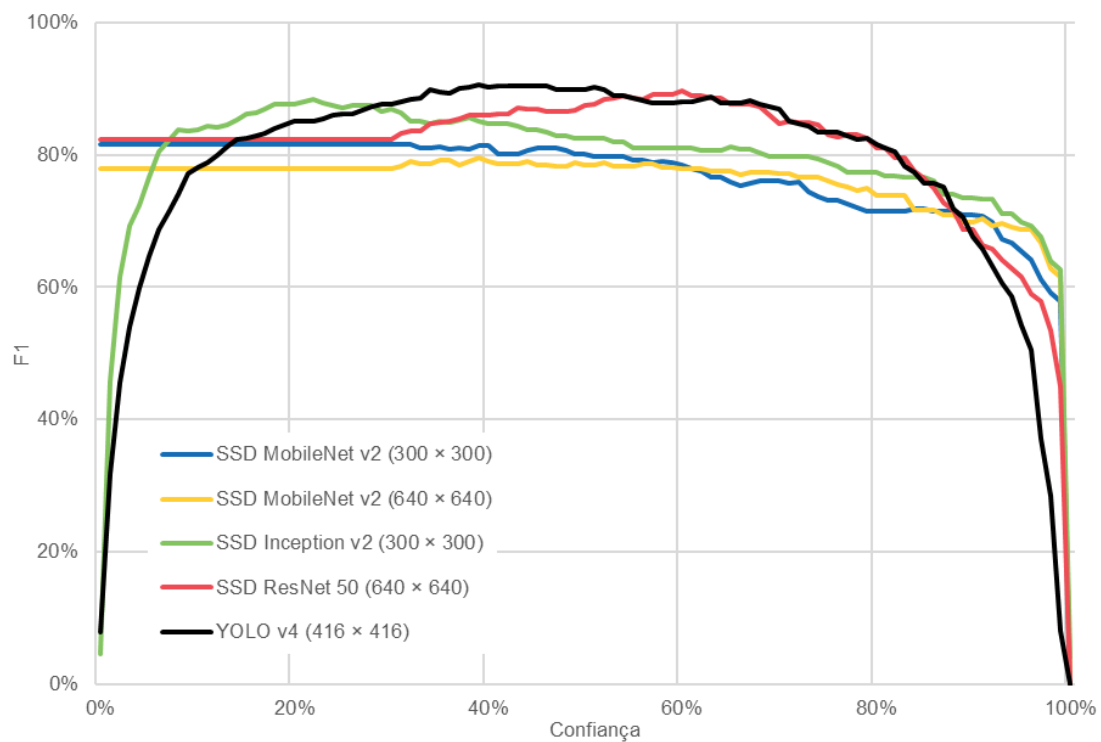
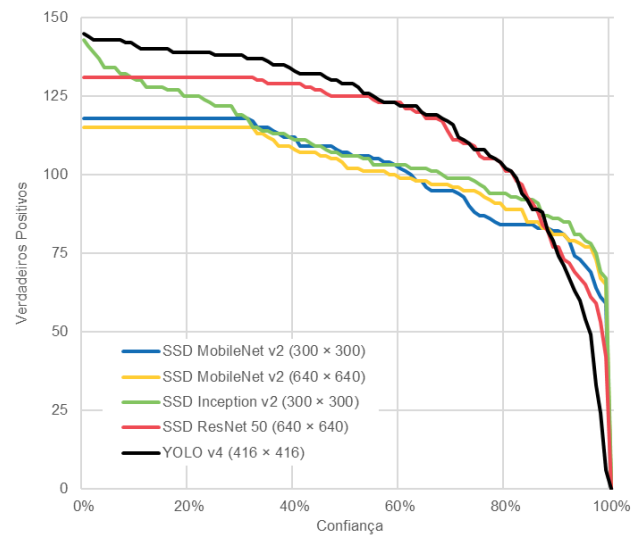
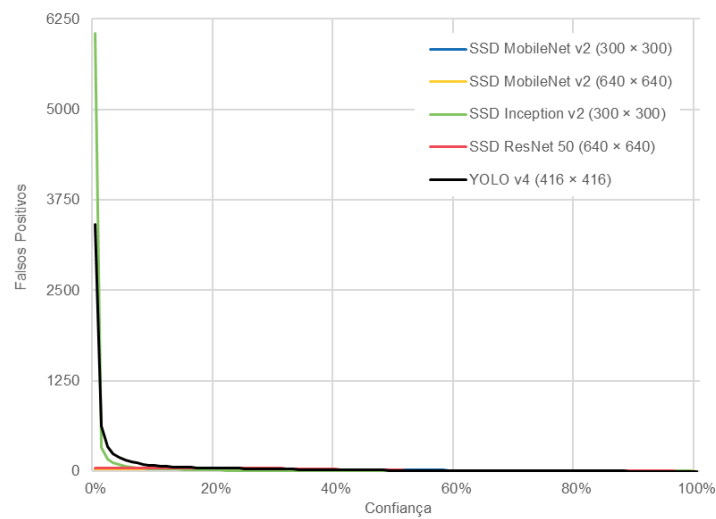


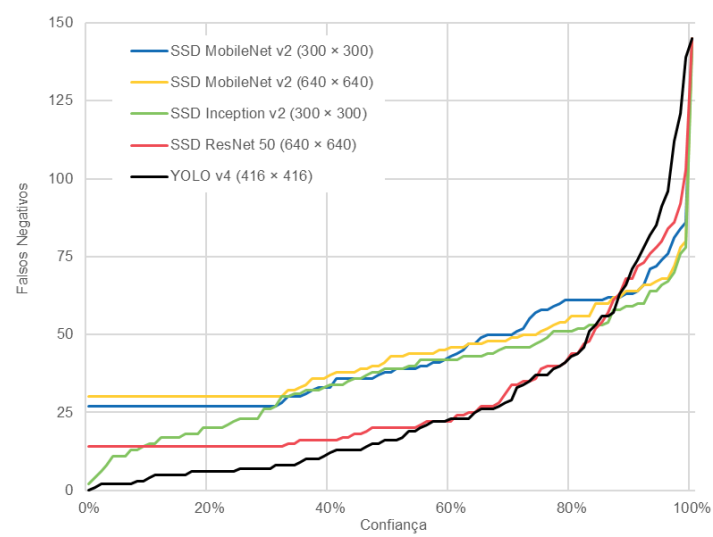
Figura II xix – Evolução do F1 com o aumento da confiança, na amostra de validação sem *augmentation*.



a) Verdadeiros Positivos



b) Falsos Positivos



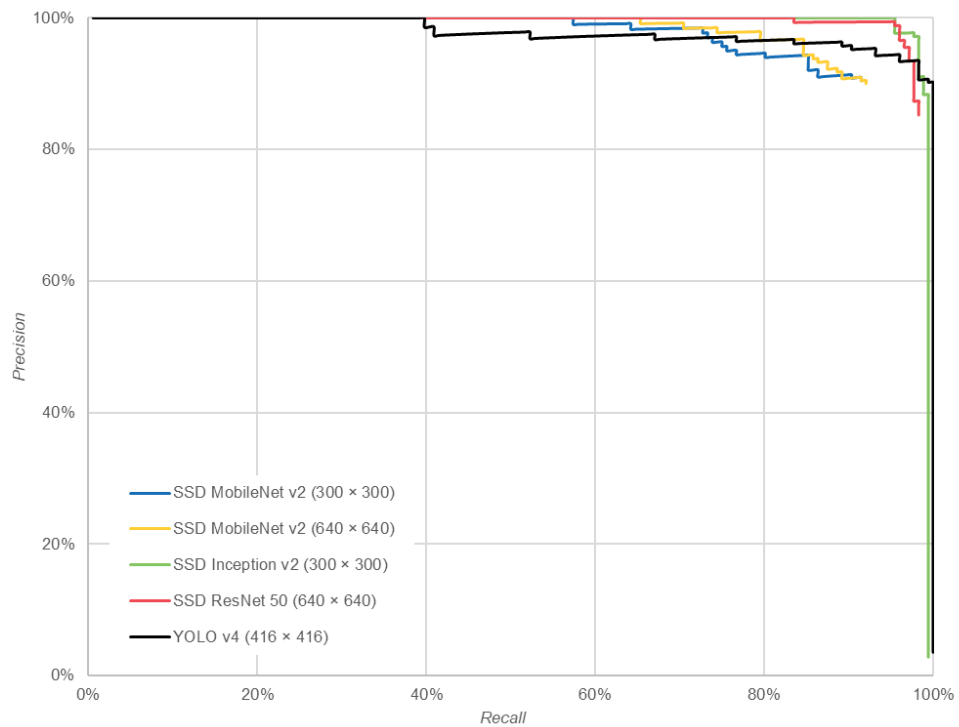
c) Falsos Negativos

Figura II xx – Evolução do número de: a) Verdadeiros Positivos , b) Falsos Positivos e c) Falsos Negativos com a evolução da confiança, na amostra de validação sem *augmentation*.

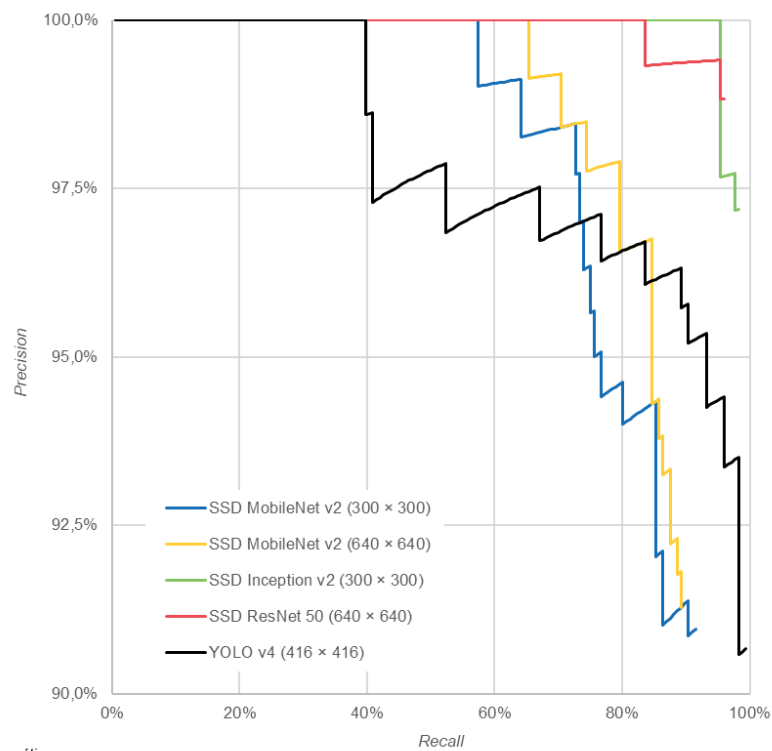
Tabela II xxvi – Resultados das métricas consideradas para os modelos em estudo. Cálculo executado na amostra de teste sem *augmentation* para todas as detecções e na melhor confiança calculada.

Modelo	Confiança $\geq$	AP	<i>Precision</i>	<i>Recall</i>	F1
SSD MobileNet v2 (300 × 300)	0%	90%	90%	91%	91%
SSD MobileNet v2 (640 × 640)	0%	91%	90%	92%	91%
SSD Inception v2 (300 × 300)	0%	99%	3%	99%	5%
SSD ResNet 50 (640 × 640)	0%	98%	85%	98%	91%
YOLO v4 (416 × 416)	0%	98%	4%	100%	7%
SSD MobileNet v2 (300 × 300)	31%	90%	91%	91%	91%
SSD MobileNet v2 (640 × 640)	39%	89%	91%	89%	90%
SSD Inception v2 (300 × 300)	22%	98%	97%	98%	98%
SSD ResNet 50 (640 × 640)	60%	96%	99%	96%	97%
YOLO v4 (416 × 416)	39%	98%	91%	99%	95%

AP: *Average Precision*



a) Confiança = 0%



b) Confiança ótima

Figura II xxi – Curvas *precision × recall* para a: a) confiança de 0% e b) confiança ótima. Avaliação executada na amostra de teste sem *augmentation*.

Tabela II xxvii – Matriz de confusão para os modelos SSD e YOLO em estudo na amostra de teste sem *augmentation*, para a confiança otimizada.

Conf. = 31%		SSD MobileNet v2 (300 × 300)				
Deteção		Referência			Não referenciada	
		<i>coty</i>	<i>minus9</i>	<i>plus9</i>	solo	infestante
	<i>coty</i>	65	0	0	0	0
	<i>minus9</i>	0	97	0	0	0
	<i>plus9</i>	0	0	9	0	0
Conf. = 39%		SSD MobileNet v2 (640 × 640)				
Deteção		Referência			Não referenciada	
		<i>coty</i>	<i>minus9</i>	<i>plus9</i>	solo	infestante
	<i>coty</i>	60	0	0	0	0
	<i>minus9</i>	0	92	0	0	1
	<i>plus9</i>	0	1	9	0	0
Conf. = 22%		SSD Inception v2 (300 × 300)				
Deteção		Referência			Não referenciada	
		<i>coty</i>	<i>minus9</i>	<i>plus9</i>	solo	infestante
	<i>coty</i>	68	0	0	0	0
	<i>minus9</i>	0	98	1	0	0
	<i>plus9</i>	0	0	9	0	0
Conf. = 60%		SSD ResNet 50 (640 × 640)				
Deteção		Referência			Não referenciada	
		<i>coty</i>	<i>minus9</i>	<i>plus9</i>	solo	infestante
	<i>coty</i>	65	0	0	0	0
	<i>minus9</i>	0	97	0	0	0
	<i>plus9</i>	0	0	9	0	0
Conf. = 39%		YOLO v4 (416 × 416)				
Deteção		Referência			Não referenciada	
		<i>coty</i>	<i>minus9</i>	<i>plus9</i>	solo	infestante
	<i>coty</i>	71	1	0	0	2
	<i>minus9</i>	0	97	0	0	2
	<i>plus9</i>	0	0	9	0	0

Conf.: confiança que otimiza o valor de F1 na amostra de validação.

Tabela II xxviii – Resultados das métricas exatidão, *precision* e *recall* para cada classe de alface para os modelos SSD e YOLO em estudo. Avaliação na amostra de teste sem *augmentation*.

<i>coty</i> n = 69	Exatidão	<i>Precision</i>	<i>Recall</i>	F1
MobileNet v2 (300 × 300)	98%	100%	94%	97%
MobileNet v2 (640 × 640)	95%	100%	87%	93%
Inception v2 (300 × 300)	99%	100%	99%	99%
ResNet 50 (640 × 640)	98%	100%	94%	97%
YOLO v4 (416 × 416)	98%	96%	100%	98%
<i>minus9</i> n = 98	Exatidão	<i>Precision</i>	<i>Recall</i>	F1
MobileNet v2 (300 × 300)	99%	100%	99%	99%
MobileNet v2 (640 × 640)	96%	99%	94%	96%
Inception v2 (300 × 300)	99%	99%	100%	99%
ResNet 50 (640 × 640)	99%	100%	99%	99%
YOLO v4 (416 × 416)	98%	98%	99%	98%
<i>plus9</i> n = 9	Exatidão	<i>Precision</i>	<i>Recall</i>	F1
MobileNet v2 (300 × 300)	100%	100%	100%	100%
MobileNet v2 (640 × 640)	99%	90%	100%	95%
Inception v2 (300 × 300)	100%	100%	100%	100%
ResNet 50 (640 × 640)	100%	100%	100%	100%
YOLO v4 (416 × 416)	100%	100%	100%	100%

n: número de referências na amostra de teste.