Vineyard yield estimation using image analysis – a review

Rosario Pascale

Dissertation to obtain a Master's Degree in

Viticulture and Oenology Engineering

PRESIDENT
PhD Joaquim Miguel Rangel da Cunha Costa, Assistant Professor at Instituto Superior de Agronomia, Universidade de Lisboa.

MENBERS
PhD Carlos Manuel Antunes Lopes, Associate Professor with Habilitation at Instituto Superior de Agronomia, Universidade de Lisboa;
PhD Ricardo Nuno da Fonseca Garcia Pereira Braga, Assistant Professor at Instituto Superior de Agronomia, Universidade de Lisboa.

2020

Abstract

Yield estimation is one of the main goals of the wine industry, this because with an accurate yield estimation it is possible to have a significant reduction in production costs and a better management of the wine industry. Traditional methods for yield estimation are laborious and time consuming, for these reasons in the last years we are witnessing to the development of new methodologies, most of which are based on image analysis. Thanks to the continuous updating and improvement of the computer vision techniques and of the robotic platforms, image analysis applied to the yield estimation is becoming more and more efficient. In fact the results shown by the different studies are very satisfying, at least as regards the estimation of what is possible to see, while are under development several procedures which have the objective to estimate what is not possible to see, due to bunch occlusion by leaves and by others clusters.

I this work the different methodologies and the different approaches used for yield estimation are described, including both traditional methods and new approaches based on image analysis, in order to present the advantages and disadvantages of each of them.

Key words: Yield estimation, Image Analysis, Bunch occlusion, Robot platforms, Precision Viticulture

Resumo

A estimativa de rendimento é um dos principais objetivos da indústria do vinho, pois com uma estimativa de rendimento precisa é possível ter uma redução significativa nos custos de produção e uma melhor gestão da indústria do vinho. Os métodos tradicionais de estimativa de rendimento são trabalhosos e demorados, pelo que nos últimos anos assistimos ao desenvolvimento de novas metodologias, muitas das quais baseadas na análise de imagens. Graças à contínua atualização e aprimoramento das técnicas de visão computacional e das plataformas robóticas, a análise de imagens aplicada à estimativa de rendimento está a tornar-se cada vez mais eficiente. De facto, os resultados apresentados pelos diferentes estudos são muito satisfatórios, pelo menos no que diz respeito à estimativa do que é possível ver, enquanto estão em desenvolvimento vários procedimentos que têm por objectivo estimar o que não é possível ver, devido à oclusão dos cachos pelas folhas e por outros cachos.

Neste trabalho são descritas as diferentes metodologias e as diferentes abordagens utilizadas para a estimativa de rendimentos, incluindo quer métodos tradicionais quer novas abordagens baseadas na análise de imagens, de forma a apresentar as vantagens e desvantagens de cada uma delas.

Palavras-chave: Análise de Imagem, Estimativa de Rendimento, Oclusão dos cachos, Plataformas de Robô, Viticultura de Precisão

Resumo alargado

Agricultura de precisão (AP) é um conceito que tem como objetivo principal a gestão da variabilidade espacial e temporal das parcelas de culturas. , de forma a obter um aumento do benefício económico e uma diminuição do impacto ambiental (Blackmore, 1999). A Viticultura de Precisão (PV) tem o mesmo objetivo da AP. Tradicionalmente, a gestão da vinha tem sido realizado como se fosse homogenea, apesar da variação espacial que pode ser observada em cada parcela. Na AP, cada parcela é gerida de forma diferente, seguindo as reais necessidades da mesma. Hoje em dia essa abordagem é fundamental, pois estamos presenciando, por um lado a crescente competição no mercado internacional, pelo que a redução dos custos e a preservação do ambiente tornam-se muito importantes.

Por estas razões, o desenvolvimento e a aplicação de técnicas inovadoras destinadas a monitorizar objetivamente a vinha é uma questão chave na investigação em viticultura, a fim de melhorar a sustentabilidade do cultivo da vinha, bem como a qualidade da uva e do vinho..

A estimativa de rendimento é um dos principais objetivos da indústria do vinho, pois com uma estimativa de rendimento precisa é possível ter uma redução significativa nos custos de produção e uma melhor gestão da indústria do vinho. Os métodos tradicionais de estimativa de rendimento são trabalhosos e demorados. Por exemplo, para aplicar o método tradicional baseado nas componentes do rendimento, é necessário ir ao campo todos os anos para contar o número real de videiras/ha, o número de cachos/videira e medir o peso do cacho.. Ao longo dos anos, alguns métodos alternativos também foram testados, como o método aeropolínicoe o método de medição da tensão dos arames, mas ambos os métodos apresentam várias desvantagens e, para além da serem muito trabalhosos, não são aplicáveis em todas as situações.

Por estas razões, nos últimos anos estamos a testemunhar o desenvolvimento de novas metodologias para a estimativa do rendimento, a maioria das quais baseadas na análise de imagens, de forma a termos novos métodos não invasivos, graças aos quais é possível evitar as desvantagens dos métodos tradicionais.

As previsões de produtividade podem ser tentadas a qualquer momento durante o ciclo biológico da videira. , Na literatura é possível encontrar diferentes tipos de abordagens, que vão desde os estados iniciais de crescimento até á colheita. Assim, existem alguns trabalhos em que se tentou estimar a produção a partir da contagem do nº de lançamentosdo número de flores por inflorescência e do número de bagospor cacho.

Graças à contínua atualização e melhoria das técnicas de visão computacional e das plataformas robóticas, a análise de imagens aplicada à estimativa de rendimento está se tornando cada vez mais eficiente. De facto, os resultados apresentados pelos diferentes estudos são muito satisfatórios, pelo menos no que diz respeito à estimativa do que é possível ver, enquanto estão em desenvolvimento vários procedimentos que têm por objectivo estimar o que não é possível ver, devido à oclusão do cacho.

A oclusão do cacho pode ser causada por folhas ou por outros cachos. As folhas, ao influenciarem a visualização dos cachos, afectam a estimativa de produção. Todavia, na maior parte das situações não é possível eliminar as folhas para ter uma visão livre dos cachos. A porosidade da sebe é uma característica importante , pois é um indicador da exposição dos cachos e da circulação do ar., Há alguns trabalhos em que os autores tentaram estimar a quantidade de uva encoberta pela folhagem e por outros cachos. Ambos os tipos de oclusão dependem de vários fatores, como variedade, sistema de condução, sistema de poda, etc. Portanto, a calibração desses dois tipos de oclusão é bastante difícil.

Neste trabalho são revistas as diferentes metodologias e as diferentes abordagens utilizadas para a estimativa do rendimento, incluindo quer métodos tradicionais quer novas abordagens baseadas na análise de imagens, de forma a apresentar as vantagens e desvantagens de cada uma delas.

Palavras-chave: Estimativa de rendimento. Análise de imagem, Oclusão dos cachos, Plataformas de Robô, Viticultura de Precisão

Table of contents

List of figures

Acronyms

a* greenness-redness coordinate

ANNs artificial neural networks

b* blueness-yellowness coordinate

BxB bunch by bunch

BR blue, red

CIE commission Internationale d'Eclairage

CNNs convolutional neural networks

F1 F1 score

FCN fully convolutional network

GDD growing degree days

GB green, blue

HBS hue, saturation and brightness

HVS human visual system

IOU Intersection Over Union

ISA Instituto superior de Agronomia

L* lightness coordinate

LCM linear canopy meters

NN neural network

PA precision agriculture

PV precision viticulture

PQA Point Quadrat Analysis

PR precision

R^2 determination coefficient

RC recall

RG red, green

RGB Red, Green, Blue

ROI region of interest

TPR true positive rate

TTMs Trellis tension monitors

UAV unmanned aerial vehicles

1. Introduction

Precision Agriculture (PA) is a management strategy that gathers, processes and analyzes temporal, spatial and individual data and combines it with other information to support management decisions according to estimated variability for improved resource use efficiency, productivity, quality, profitability and sustainability of agricultural production (International Society of Precision Agriculture; https://www.ispag.org/about/definition). Precision viticulture thus tries to find the widest range of available observations to describe the vineyard spatial variability and provide recommendations to improve management efficiency in terms of quality, production and sustainability (Matese & Di Gennaro, 2015)

Traditionally, vineyard management has been conducted as if it were homogeneous, even though the spatial variation that can be observed in each parcel. In fact vineyards are characterized by a high heterogeneity due to structural factors such as the pedo-morphological characteristics, and other dynamics such as cropping practices and seasonal weather (Bramley, 2003). This variability cause different response by vine, each of which have a direct consequence on the grape quality (Smart, 1985). Than using a PV approach, each parcel is management in different way, following the real needs of it. Nowadays this approach is fundamental, because we are witnessing on one hand to the growing competition on international market, so the reduction of the costs becomes mandatory, on the other hand to the climate change due to pollution, then chemicals should be used only when there is a real need.

For these reasons the development and application of innovative techniques aimed at objectively monitoring the vineyard is a key issue in viticulture research in order to improve grape-growing sustainability as well as grape and wine quality (Diago et al.,2014)

In this scenario, yield estimation has become one on the main goal of the wine industry. In fact with an accurate yield forecast it is possible to achieve a considerable cost saving and a higher gain realised if the fluctuation (in terms of kg of grape) is minimized (Dunn, 2010). In addition, it can be very useful for the management of both vineyard and the cellar (Dunn and Martin, 2003). For example with an accurate yield estimation it is possible to predict machinery and staff needed for harvest or the space needed in the cellar for the tanks positioning. Moreover, an early yield estimation could be important to planning canopy management (e.g. cluster thinning in order to avoid overcropping that leads to lower grapes quality), planning selling price of the grapes, developing marketing strategies and so on (Dunn and Martin, 2004). Furthermore fruit detection can also be a preliminary step for disease and nutrient deficiency monitoring (Barbedo, 2019). However, the spatial and seasonal variability due to biotic or abiotic factors, makes the early yield forecasting difficult to carry out with the appropriate accuracy.

According to the EU Strategic Research Agenda For Robotics in Europe 2014-2020, robotics technology will become dominant in the coming decade in every field (Lopes et al., 2017).   Among agricultural sectors Viticulture is one where aerial and ground robots can provide several services bringing numerous advantages. In fact, there are some cultural practices which have already been fully automated, like mechanical weeding (Vitirover, 2017), while others are under development, e.g. pruning (Visionrobotic, 2017).

In the last few years, autonomous robotic platforms got a lot of attention because in addition to the execution of cultural practices, they can be very useful for data collection, thus helping decision making. As said before, accurate yield estimation has become one of the main objectives of wine industry. Then, robotic platforms that can carry several types of sensors, have been used in numerous attempts to apply image analysis technologies for bunch and/or berries recognition in images and processing methods for grape yield estimation (Lopes et al., 2017).

Among these platforms there is Vinbot. Vinbot is an autonomous cloud-computing vineyard robot to optimize yield management and wine quality. It is an EU project (Powerful precision viticulture tool to break traditional yield estimation in vineyards) funded under the FP7 SME program. The Vinbot robot has the purpose of collecting real field data, as yield components, leaf area, and comparing them with the estimates provided by the autonomous VINBOT system.

VINBOT is an all-terrain autonomous mobile robot equipped with a set of sensors capable of capturing and analysing vineyard images and 3D data by means of cloud computing applications.

A new approach to yield estimation has become mandatory, because traditional methods are expensive and/or destructive Image analysis for yield estimation has several challenges to overcome, like the acquisition of images in the field (i.e. uncontrolled scenario) and bunch occlusion.

State-of-the-art computer vision system based on convolutional neural networks (CNNs) (which have replaced classic machine learning and pattern recognition) can deal with variations in pose, shape, illumination and large inter-class variability (He et al., 2016; Krizhevsky et al., 2012; Simonyan and Zisserman, 2014) which are essential features needed for an accurate recognition of objects in outdoor environments. The use of CNNs and their variants is become a standard for yield estimation, also enhanced by companion techniques like semantic segmentation, transfer learning and three-dimensional association to integrate and spatialize the detection results to overcome multiple counting and occlusions, and even to integrating with non-imaging approaches, for instance, using historical data (Coviello et al., 2020). Thanks to continuous updating and improvement of these technologies, image analysis is being increasingly used in viticulture.

Then, looking at state-of-the-art, it can be said that thank to deep learning machine algorithm it is possible to have an accurate estimation of what human eyes can see in the field; however, remains the problem to forecast what is not possible to see due to bunch occlusion (bunch-on-bunch and leaf induced occlusions).

1.1 Objectives

The aim of this work is to review the use of image analysis for vineyard yield estimation . During the last few years, different methods have been developed and tested, in order to overcome the various adversities that this technology presents, especially in uncontrolled environmental conditions, such as those of a vineyard, in which the variability of light and bunch occlusion make it difficult to identify the berries. In this thesis it will be review the major studies on the subject, illustrating the different approaches and the different technologies used in order to illustrate the advantages and disadvantages of these methods.

2. Literature Review

2.1. Traditional Methods

Traditionally, yield forecasting is based on the yield component sampling and counting, which are the number of vines/ha, number of nodes/vine, number of shoot/node, cluster/shoot, flowers/cluster, berries/cluster and berry weight (Martin et al., 2003). These components are crucial not only to define the final yield, but also to determine the quality of it (Diago et al., 2015). Traditional methods are easy enough to apply, however they require the collection of samples in the field. Therefore, are laborious, time consuming, expensive and destructive (Cunha, 2010; Dunn and Martin, 2004).

Methods based on yield components counts can be used at different stages all along the growing cycle. Among all the stages, the most used is at veraison, because the forecast at this phenological stage allows to obtain the yield forecast early enough to have adequate time for harvest and winery planning management, and because the forecasting at this stage can be quite accurate (Victorino et al., 2017)

However, yield forecast can be carry out also at early stages, which means any stage from budburst to flowering and, moreover, a very early forecast can be done estimating bud fruitfulness before budbreak (Clingeleffer et al., 2001).

One of the main problems of the traditional methods is the sampling, in fact in many cases it is not made accurately. To obtain an accurate sample, Dunn et al., (2010) proposed the following equation:

$$n = \frac{t^2 * CV^2}{PE^2} \qquad \text{Eq.1}$$

In which n is the number of samples required, t is value obtained with the Student t, CV is the coefficient of variation and PE is the percentage of error (in other words the doubt that we want tolerate).

Among the traditional methods, Sabbatini et al., (2012), describe three different methods: (1) one based on historical records of cluster weight at harvest; (2) one based on cluster weight during the “lag phase”; (3) one based on growing degree days (GDD) accumulation.

2.1.1 Harvest Cluster Weight Method

This method provide requires the use of the cluster weight and the number of clusters at the harvest time of the previous season:

$$PY = \frac{(\mathrm{ANV}*\mathrm{NC}*\mathrm{CW})}{(1000)} \qquad \text{Eq. 2}$$

Where:

PY = predicted yield (t/ha)

ANV = actual number of vines per ha

NC = number of clusters per vine

CW = cluster weight (in Kg) at harvest time

According with the formula, the grower needs to measure three parameter each year: (1) number of vine/ha; (2) number of cluster/vine; (3) cluster weight at harvest time.

The number of vines per hectare must be counted every year, because almost always there is a missing of vines due to several reasons such as diseases, replanting, etc.

Canopy management practices and intensity, mainly pruning and cluster thinning, affect the number of clusters, while the weight varies each year due to several factors, such as weather conditions, level of canopy management, fertilization, fungal disease, irrigation, insect feeding damage and bird depredation (Sabbatini et al., 2012).

2.1.2. Lag-phase method

2.1.3. Growing Degree Days (GDD) Method

GDD means the summation of the differences between the average daily temperature (T_i) and the zero vegetation temperature (T_b) of the considered species. The GDD equation is:

$$GDD = \sum_{i=1}^{n}(T_i - T_b) \qquad \text{Eq. 3}$$

This method is based on the correlation between GDD and the final weight of the cluster; this parameter changes for each variety, e.g. for an early varieties as Riesling and Pinot noir the GDD required to reach the 50% of the final weight are between 1000-1200, while for a late variety as Merlot it need about 1700 (Sabbatini et al., 2012).

2.2. Alternative yield estimation methods

2.2.1. Airborne pollen method

The use of aeropalynological data, in order go get an yield forecast, is based on the fact that the airborne pollen concentration allows one to simultaneously estimate several factors which influence the production process of the vine (pre-floral conditions, the effects of diseases on grape quantity and the strength of the plants vigor which are influenced by conditions in previous years) and the easy sampling of the yield components (monitoring the evolution of the crop are as in the production, avoiding counting in reference parcels) (Ribeiro et al., 2005).

With this method is possible to determine the potential production yield measuring the pollen concentration. Thanks to aeropalynological yield forecast model it is possible to explain from 97% to 99% of the annual variability (Cunha et al., 2003). The method is based on the trapping of pollen grains on gauze filters with an area of 400 cm^2 fixed on a wind-vane, in order that the filter is orientaded in according to the towards the wind direction. Following the Cour method (Cour and Villemur, 1985), the filters are changed two times per week and the pollen concentration is measured and expressed in number of pollen grain/m3. The strength point of the model is: the possibility of obtaining an accurate early forecast and the good results in crop prediction, but has the disadvantages which are represented by of the placement of the airborn pollen sampling device at regional level and by the complex laboratory process involved (Cunha et al., 2003).

Cristofolini and Gottardini, 2000 instead of using a Cour trap, have used a Hirst-type sampler, which allows to determine the daily airbone pollen concentration (pollen grain/m^3/day). To improve the model the authors have recorded the meteorological parameters, in particular rainfall and temperature during the main pollen season, which are the two most important parameters that influence this stage (Cristofolini and Gottardini, 2000).

2.2.2. Trellis Tension Monitoring Method

This method provides the automated measurement of the within−season changes in tension in the horizontal support wire of the trellis (Tarara,et al., 2004). The measurements are taken after the bloom, because before of that phenological stage, the tension variation is due only to the vegetative growth of the shoot, while after the bloom the variation is attributable to the fruit growth (Blom and Tarara, 2009). Blom and Tarara, (2009), in their work use the following equation to predict the yield:

$$Y_{t,c} = \left(\frac{Y_a}{T_{t,a}}\right) * T_{t,c} \qquad \text{Eq. 4}$$

Where $Y_{t,c}$ is the predicted yield at any time of the current year; Y_a is the yield from a previous year; $T_{t,a}$ is the trellis wire tension at time t from the antecedent year; $T_{t,c}$ is the wire tension at time t of the current year. Of course, given such equation, this method requires several historical data and with the collection of data in the different years the accuracy improves. The results obtained with this method show that it can replace the traditional yield estimation practices or to supplement longstanding practices with real-time information that can be applied to dynamic revision of static yield estimation. There is to say that beyond the historical data, other obstacles to the use of this method are necessity to have a trellis system in perfect condition and it requires many sensors (Fig.) across all rows.

Figure 1. Sensor for trellis tension measure (source: https://www.goodfruit.com/taking-the-guesswork-out-of-yield-estimating/)

2.3 Image Analysis

Image Analysis is the extraction of meaningful information from images, mainly from digital images (Solomon et al., 2010). Methods of image analysis vary from low-level processing of individual pixel or small neighbourhoods, to high-level processing made with the summary of the information contained in the entire image (Glasbey and Horgan, 2001).

Every image is subject to some degradation during the capture process. This degradation can be caused by noise, blurring or a distortion of image frame. To reduce these effects is used a process called “Image enhancement”, which compraises utilization of filters, unwarping and interpolation.

With filtering process it is possible to smooth or emphasize the edges, depending on objectives; filters are also commonly used to reduce the size of the image and emphasize the parts which are important to the final output, as it is in image segmentation (Glasbey and Horgan, 2001).

Image segmentation is the process that consist of the sub-division of an image into different region, which correspond to different objects or part of them. Every pixel is allocated in a region or category; a good segmentation is one in which pixels that have been placed in the same category, have analogous value and form a connected region and are dissimilar to neighbouring pixels which have been placed in a different category (Glasbey and Horgan, 2001).

2.3.1 Image related information

2.3.1.1. Colour space

A color space is a specific organization of colours. In combination with colour profiling supported by various physical devices, and supports reproducible representations of colour, whether such representation entails an analog or a digital representation (Poynton, 1995)

The choice of the colour space it's a fundamental decision that can dramatically influence the result of the image analysis process. According to Tkalčič and Tasič, (2003) colour is the way the human visual system (HVS) measures a part of the electromagnetic spectrum, between from 300 to 800 nm. It is possible to classify the colour in three categories (reference??):

1) HVS based colour spaces, which includes: RGB colour space, the opponent colours theory based colour spaces and the phenomenal colour spaces.
2) Application specific colour spaces include the colour spaces adopted from television (TV) systems (YUK; YIQ), photo systems (Kodak PhotoYCQ and printing systems (CMyK)
3) CIE colour spaces are spaces proposed by the Commission Internationale d'Eclairage (CIE) and have some properties of high importance like device-independency and perceptual linearity (CIE HZ, Lab and Luv colour spaces)

RGB colour space is any additive colour space based on the RGB colour model (Fig.2A), in which red, green, and blue light are added together in various ways to reproduce a broad array of colours (Hirsch, 2004). The main disadvantage of the RGB colour space is the high correlation (r value) between its components: about 0.78, 0.98 and 0.94 between BR, RG and GB, respectively (Tkalčič and Tasič, 2003).

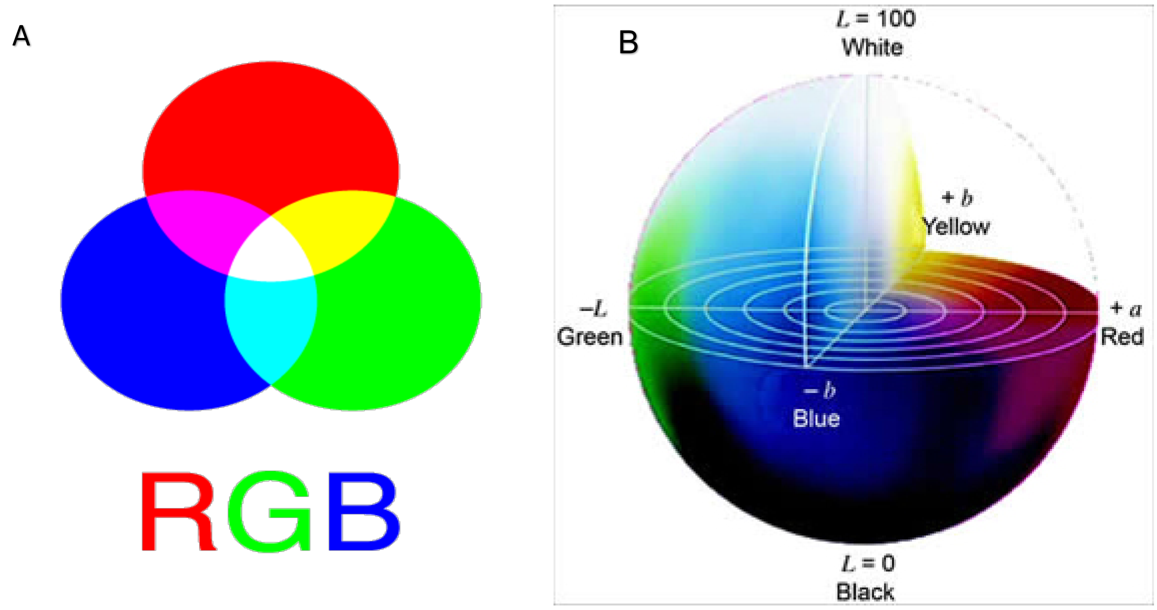


Figure 2: (A) RGB Colour space (source: https://www.customerlabs.co/blog/theory-of-colours/) ; (B) CIELab colour space (source: https://medium.com)

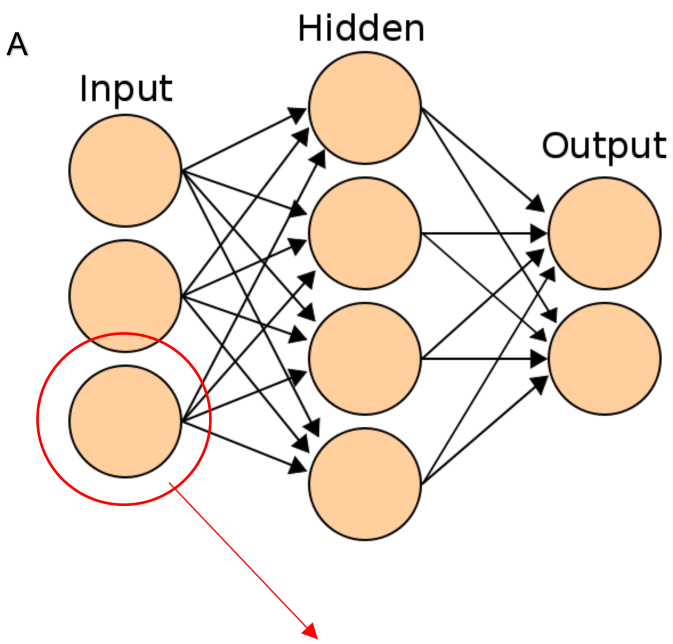
The phenomenal colour space uses three attributes to describe a colour: hue, saturation and brightness (HBS). Hue tells us which colour we are seeing (red, blu, yellow, etc.); the saturation is the level of non-whiteness. This means that a saturated colour is very pure, vivid. Conversely, the more unsaturated the colour, the more it tends towards white. Brightness is the measure of intensity of the light. Phenomenal colour spaces are linear transformation of RGB colour space, thus its deformation. For this reason their practical applications are limited (Tkalčič and Tasič, 2003).

CIELab colour space (Fig.2B) is a colour space defined by the International Commission on Illumination (CIE) in 1976. It expresses colour as three values: L* for the lightness from black (0) to white (100), a* from green (−) to red (+), and b* from blue (−) to yellow (+). CIELAB was designed so that the same amount of numerical change in these values corresponds to roughly the same amount of visually perceived change (CIE Colorimetry, 1986; Tkalčič and Tasič, 2003).

2.3.1.2 Machine learning

Artificial neural networks (ANNs), generally called neural networks (NNs) are simplistic mathematical models vaguely inspired by the biological neural networks that constitute animal brains (Fig. 3A) (Chen et al., 2019).

Neural networks components are: neurons, connection and weights, propagation function (Fig.3B)



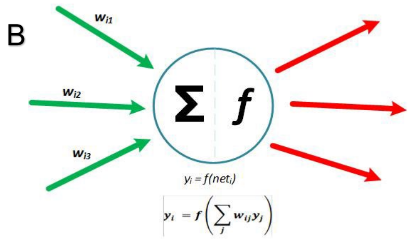


Figure 3. (A) Neural Networks structure (source: https://commons.wikimedia.org/wiki/File:Artificial_neural_network.svg) (B)Artificial neuron (source: Dike et al., 2018)

ANNs training can be divided in: supervised learning and unsupervised learning. Supervised learning is the machine learning task of learning a function that maps an input to an output based on example input-output pairs (Van Gerven and Bohte, 2017). It infers a function from labeled training data consisting of a set of training examples (Stuart and Norving, 2010). The main limitation of supervised learning consists in the fact that many of these algorithms works in a linear way. There are many conditions under which such an approximation is acceptable but not always.

Unsupervised learning is a type of machine learning that looks for previously undetected patterns in a data set with no pre-existing labels and with a minimum of human supervision. In contrast to supervised learning that usually makes use of human-labeled data, unsupervised learning, also known as self-organization allows for modeling of probability densities over inputs (Hinton et al., 1999).

There are several kinds of NN, like Fully Convolutional Network (FCN) and Convolutional Neural Networks (CNNs). These can be considered an evolution of NN, which allow learning descriptive criteria of the desired image regions just from the image data itself (Rudolph et al., 2018).

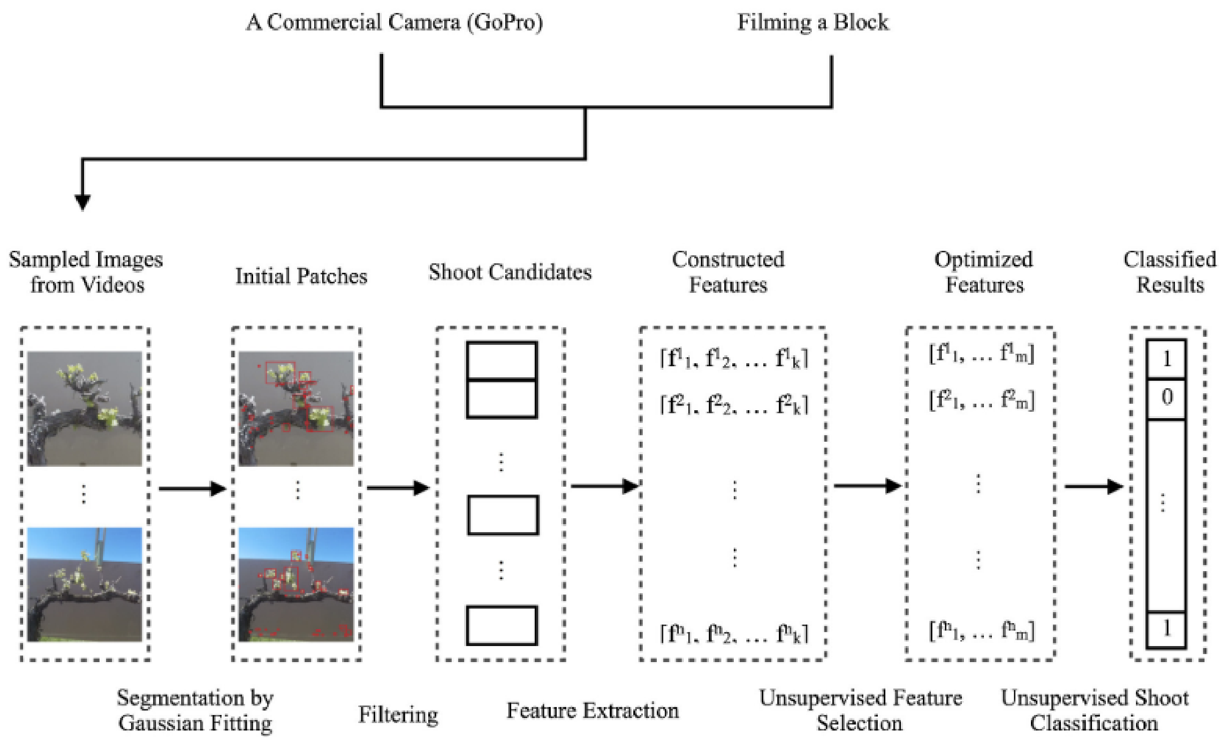
2.3.2 Application of image to vineyard yield forecasting

The computer vision techniques, based on the image analysis, have been used in several works to develop a non-invasive technology for yield estimation in order to avoid the disadvantages of the traditional conventional methods. One of the first attempts to apply these technologies to yield forecasting was done by Dunn and Martin, (2004); in this work, the authors, thanks to the use of image segmentation based on the red, green and blue technology (RGB) colour space, proved that fruits can be distinguished from other elements of the grapevine canopy. Since then, many works and strategies, many recently based on machine learning algorithms, have been developed. These algorithms are able to automatically extract information from collected images.

2.3.2.1 Shoot counting

An early shoots counting can be very usefull to help growers for the adjustment of the yield, in order to meet grape quality objectives. However, it is an expensive and time consuming measurement (Liu et al., 2017). An automated system might be able to reduce errors and costs. Liu et al., (2017) have investigated yield estimation at an early phenological stage prior to fruit set and demonstrated the ability to forecast yield using a low cost vision system along with a novel shoot detection framework (Fig.4) In their work there are several novelties (compared to previous works):

- the possibility to forecast yield and generate a map of potential yield variation at the early phenological stage of shoots;
- no manual labeling required to build a classifier;
- ability to deal with certain range of illumination changes and different scenarios with various noise based on unsupervised feature selection;
- the ability to use low-cost off-the-shelf image collection equipment;
- identification of the optimum phenological stage for imaging shoots.



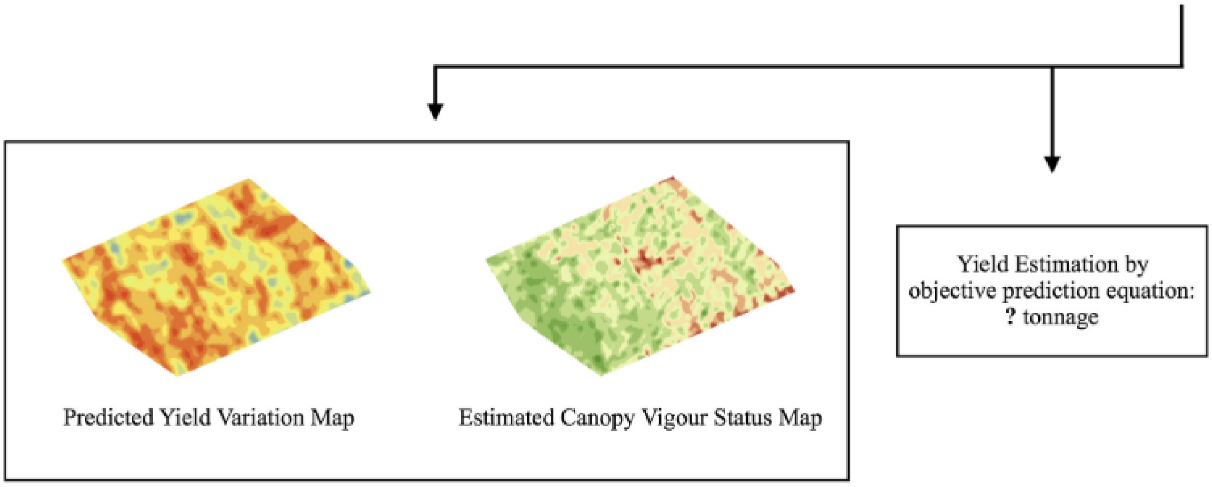


Figure 4. Flowchart of the system for yield estimation of Liu et al., 2017 (source: Liu et al., 2017)

They have developed a completely automatic system for yield forecasting based on shoots counting. Shoot detection has several adversities to overcome: varying lighting conditions, undesired objects in the field of view (e.g. posts, cordon, grass, animals, wire, reflections), change of shoot position in the field of view, shadows and barren cordon. To overcome these adversities, the authors proposed a method that segments the potential shoot patches by Gaussian fitting based on colour histograms for automatically locating the threshold value accurately. It then combines different scalar features into a feature vector to use as a descriptor of potential shoots. Since the data was collected during the day, illumination condition and scenario can vary significantly within a block, for that reason supervised learning approaches is not practicable because it would be very time consuming, thus very expensive (Liu et al., 2017). For that, reason authors have chosen an unsupervised selection based on three-correlation filter (Kendall et al., 1946; Higgins, 2003; He et al., 2005), which are compared in order to improve clustering performance and increase computational efficiency. To get a better yield estimation, two unsupervised clustering approaches were compared: K-means (MacQueen, 1967) and Agglomerative Hierarchical clustering (Hastie et al., 2005). The shoot detection framework achieved an accuracy of 86.83%, recall of 91.52%, precision of 89.18% and an F1-score of 0.90 on average.

Once the number of shoots can be estimated within a block, an early estimate of potential yield can be calculated. Liu et al. (2017) proposed the following equation (Eq. 5):

$$PY = NS * PRV * R_{BS} * BW * (1 - PR) * H_e * (1 - SP) \qquad \text{(Eq. 5)}$$

Where: PY is the total predicted yield (mass of fruit without rachis assuming machine harvested); NS is the number of shoots detected from videos; PRV is the proportion of recorded video of rows (VID) (should video of rows be missing or incomplete); RBS is the ratio of bunches to shoots from historical data; BW is the average bunch weight at harvest in previous seasons; PR is the proportion of rachis weight to bunch weight; He is the harvester efficiency factor; SP is the percentage of any destructively sampled fruit before harvest (Liu et al., 2017).

The authors report that the best results were achieved when videos where captured around E-L stage 9 (Dry and Coombe, 2009). At this stage, the error system range is between 1.18% to 36.02%, which is a very good result considering that these previsions were made five months before of the harvest. However there is a high variability between varieties, possibly making this method unreliable for all conditions.

2.3.2.2 Flowers and inflorescence counting

The number of flowers per inflorescence is a very relevant to estimate the fruit set rate, in order to get an early yield forecasting (Millan et al., 2017). However, is well-known that flower development, flowering and fruit set rate vary among cultivars, location and season (May, 2004; Dry et al., 2010; Galet, 1983; Carbonneau et al., 2007). Furthermore, fruit set also shows inter- and intra-vine variability (May, 2004). Thus, a precise count of the flower number per inflorescence is essential for an accurate yield estimation (Aquino et al., 2015 a). For these reasons, some methods have been presented, like those of May, (2000) and Keller et al. (2001), which proposed a method based on wrapping sample inflorescences with a fine mesh from the beginning of anthesis until fruit set completion. Then, the collected flower caps were manually counted to estimate the number of flowers per cluster. This method, despite being valid, is time consuming and labour demanding.

In the work of Poni et al. (2006) the authors have photographed each inflorescence using a digital camera applying a dark background. Then, they estimated initial flower number on tagged inflorescences using a linear regression between actual flower number and the flower number manually counted on photo prints established for twenty inflorescences taken from extra vines. Also in this case the labour demanding is very high.

In order to overcome these adversities, in recent years it has been developed some automatic methods based on image analysis. Among these there is the one of Diago et al., (2014), which have developed an image analysis methodology to be applied to RGB images taken under field conditions to estimate the number of flowers per inflorescence automatically. The method provides three stages: (1) image pre-processing; (2) flower counting; (3) image post-processing.

The first stage provides the conversion of the image from RGB to CIELab colour space and an initial segmentation to separate the flower from the background. The threshold used for the segmentation process was automatically set for every image based on the clusters present on the histogram of the b* coordinate; that because the pixel of the background had similar colour and thus similar values in b* coordinate, and the same occurred for pixels of the inflorescence; it means that the b* coordinate of each pixel in the background is different from the b* coordinate of the flowers, in this way it is possible to split the image in different region . The flower counting (second stage), is based on the fact that the flowers have a greater degree of light reflection than other areas fo the image, thus counting the number of the areas with the highest brightness usually correspond to number of flowers. In fact in according with Millan et al., 2017, this maximum is generated by light reflection on the surface of the 'quasi spherical' shape of the flower, which generates a maximum intensity where reflection takes place. To find and identify the brighter points of the image (L* coordinate), the authors used a computation of the extended-maxima transform, which was the regional maxima of the H-maxima transform (suppression of all local maxima lower than a threshold) (Soille, 1999). This process converts the initial image into a binary one. The threshold was manually selected based on a set of images. The third stage (image post-processing) is subdivided in three more steps (region size filtering; distance between brighter areas; shape of the brighter areas) and it was intended to eliminate material other than flowers from the image (Diago et al., 2014). In order to validate the method the actual flower number per inflorescence was determined manually after image acquisition by individually detaching the flowers from the rachis. Moreover, estimation of the flower number on imaged inflorescences was also done by manually counting the flowers on printed images. The R^2 of the relationships between the number of flowers estimated either manual or automatically with the developed method, to actual flower number per inflorescence, was always higher than 0.80 ($\rho \leq 0.001$) (Diago et al., 2014). Although the method of Diago et al., (2014) is valid, it must be said that since it requires an uniform background colour, it is difficult to apply in an entire vineyard.

The method developed by Diago et al., (2014) was used by Aquino et al., (2015 b) to develop vitisFlower®, which is an application for Android devices that allows to take a picture of a grapevine inflorescence for its analysis. The application was tested following a double approach. On the one hand, the application was tested by taking and analyzing 140 inflorescence images of 11 grapevine varieties using two different devices: a high-end and a mid-range device. In this way, not only its accuracy in detecting grapevine flowers was evaluated, but also its reliability to properly work on devices of different capabilities. On the second hand, the application's computational efficiency was also evaluated by performing a benchmarking study using four devices covering the whole market spectrum (Aquino et al., 2015 b). Results obtained indicate that more than 84% of the flowers in the images were identified with a less than 6% of detection error (Aquino et al. 2015 a). Also the method of Aquino et al., (2015 b) as the one of Diago et al., (2014) requires a distinguishable background which makes difficult and laborious the screening of entire vineyards (Rudolph et al., 2018). In order to facilitate and speed up the flowers estimation, Aquino et al. (2015 a) have proposed a new methodology based on an algorithm capable of working without the need of a dark background cardboard behind the inflorescence. The proposed methodology can be divided in two main phases:

ROI extraction and flower segmentation. In (Fig. 5) it is shown a flow-chart with the most relevant processes involved in the whole image analysis.

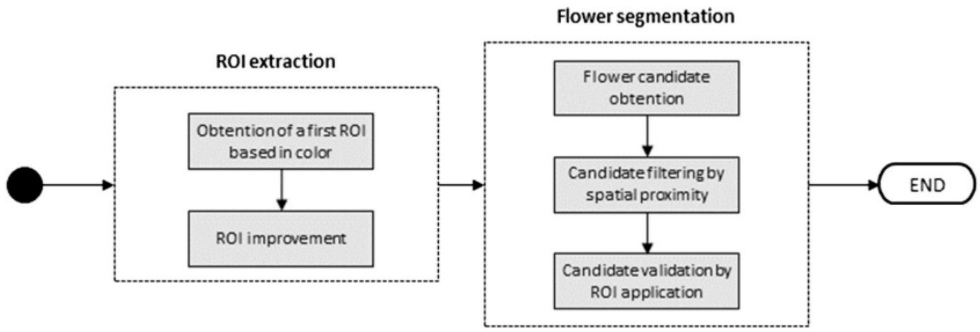


Figure 5. Algorithm Flow chart of Aquino et al., 2015

Once flowers have been counted on images, then they were counted manually both on the images and cutted from the inflorescence. Aquino et al., (2015 a) applied a linear regression with two different linear model: a variety-dependent and a variety-independent linear model to correlate the obtained information with the actual number of flowers. The R2 values obtained with both approaches were considerably high. Thus, the proposed method can facilitate the image acquisition and opens the door to its integration in vehicles and autonomous robotic platforms.

Millan et al., (2017) developed a nonlinear and linear variety-independent models for the estimation of the number of flowers per inflorescence, working with 11 variety (Viognier, Verdejo, Touriga Nacional, Tempranillo, Syrah, Riesling, Pinot Noir, Grenache, Cabernet Sauvignon, Albariño and Airen). They have taken 12 images of inflorescence for each variety, placing a black cardboard to get an homogeneous background. As pre-processing steps all the images were downscaled to a 9 Mpx resolution (they were taken to 24 Mpx) for computational workload optimization and noise reduction; moreover, images were converted to CIELab in order to decouple illumination and colour information. The first step consisted on the ROI (region of interest) extraction, which was done by binarization on the b* coordinate of the CIELab for separating the inflorescence from the background. The binarization threshold was set using Otsu's method. The second step involved the detection of flower candidates, and it was based on finding the local maximum of illumination in every flower. The third step consisted of false positive removal by discarding those flower candidates generated by reflection spots not caused by flower. To test the accuracy of this method, the authors allowed the complete development of the flowers and have collected them at the harvest time. The number of flowers and the one of the berries are related by the fruit set rate. This rate vary from one variety to another and it also depends on the weather conditions at berry set and the carbohydrate availability of each plant. It was determined as the total number of berries in the clusters of that variety (obtained by manually destemming the berries from the cluster) divided by the total number of flowers in the inflorescences (estimated with the image analysis algorithm and the nonlinear model) (Eq. 6):

$$Fruit\ set\ rate = \frac{\sum_{i=1}^{n} Bn_i^v}{\sum_{i=1}^{n} FI_i^v} \qquad \text{Eq.6}$$

Where Bn_i^v is the actual berry number value and FI_i^v is the predicted flower number value. Through this equation it is possible to obtain the estimated cluster berry number as follows (Eq. 7):

$$Estimated\ cluster\ berry\ number = FI_i^v\ * \text{Fruit set rate} \qquad \text{Eq.7}$$

Average berry weight was calculated as the weight of the berries (Kg) in the clusters divided by the number of berries in the same clusters, obtained by manual destemming:

$$Average\ berry\ weight = \frac{\sum_{i=1}^{n} Bw_i^v}{\sum_{i=1}^{n} Bn_i^v} \qquad \text{Eq. 8}$$

where Bw_i^v is the total weight of the berries (Kg), while Bn_i^v is the actual berry number value

From the previous equation it is possible to obtain the yield estimation:

$$Estimated\ yield = FI_i^v\ *\ Fruit\ set\ rate *\ Average\ berry\ weight \qquad \text{Eq. 9}$$

Flower detection results obtained were satisfactory, in fact, the overall recall (RC) value is 0.86 and the precision (PR) is 0.84. Also the results of the comparison beetwen the predicted flower number and the actual flower number are satisfactory, with an R^2 of 0.91, 0.94, 0.96 for linear, multivariable and non linear model, respectively (Fig.6) (Millan et al., 2017)

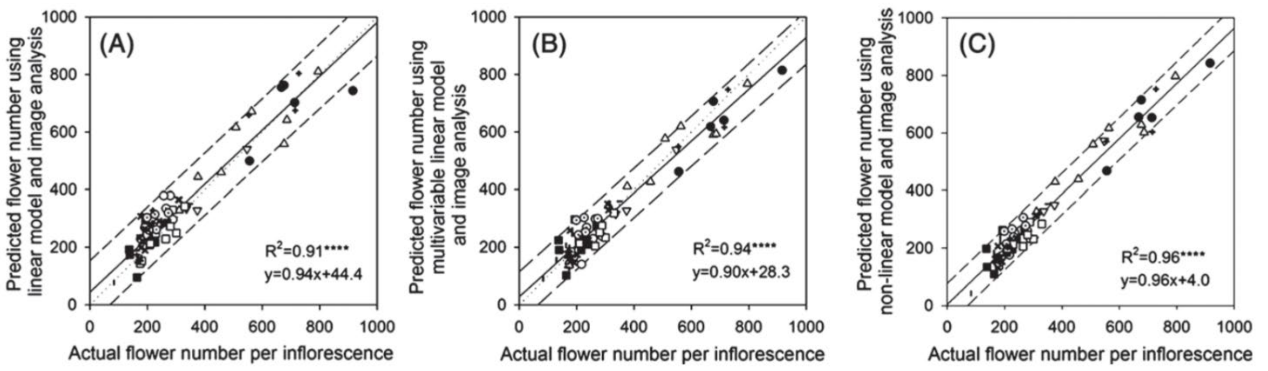


Figure 6. Graphical representation of the predicted flower number per inflorescence versus actual flower number per inflorescence using image analysis (source: Millan et al.,2017)

In Fig. 7 is possible to see three different graphics: in graphic (A) it is reported the relationship between berry weight and predicted flower number per inflorescence using image analysis and it can be see that this relationship is poor (R^2=0.49). This result was expected, due to the difference among varieties in the fruit set rate. In graphic (B) berry weight is related to predicted berry number per cluster obtained using equation (Eq. 8), while in graphic (C) total weight of berries is related to predicted berry weight obtained with estimation of berry number multiplied by average berry weight (Millan et al., 2017)

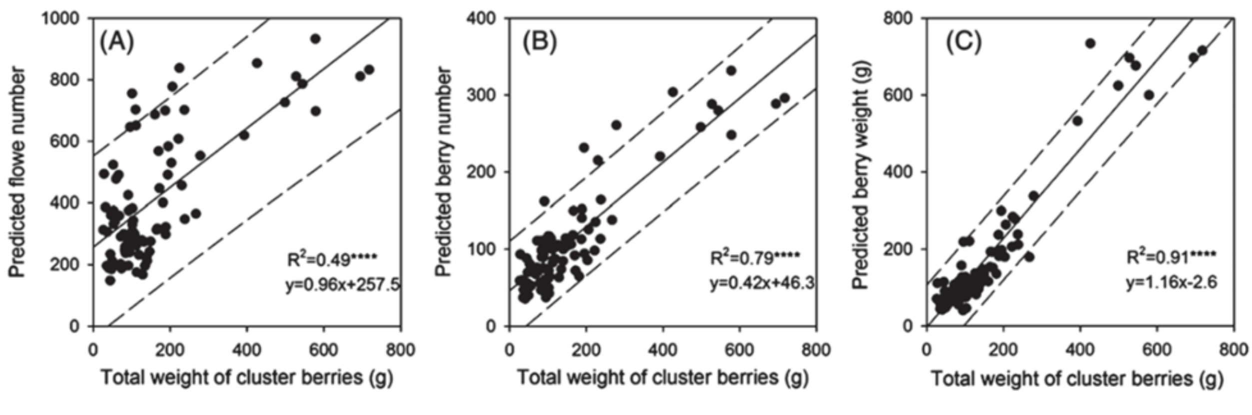


Figure 7. Graphical representation of the total weight of cluster berries versus (A) predicted flower number per inflorescence using image analysis (a fully automated algorithm processes the images to segment the flowers) and neural network model, (B) predicted berry number per cluster obtained using the estimation of flower number multiplied by berry set rate and (C) predicted berry weight per cluster obtained using the estimation of berry number multiplied by average berry weight. (source: Millan et al., 2017)

The results obtained show that to get a accurate yield estimation based on flowers counting, it is necessary to have a knowledge of fruit set rate. The number of flowers per inflorescence alone do not provide sufficient information for yield estimation, but procures promising results when combined with fruit set rate and especially when the average berry weight is known (Millan et al., 2017). Than there is the need of historical data; moreover, fruit set rate can be modified by weather conditions and viticultural practices, for these reasons there is a need to establish a protocol to obtain this factor, with precision, for the different varieties and sites.

Also the work of Rudolph et al., (2018), like the one of Aquino et al., (2015a), has been done in an unprepared vineyard (i.e. no artificial background or light was applied); the authors identified and localized inflorescence areas by partitioning the image into classes using a Fully Convolutional Network (FCN): the classes ”inflorescence” and “non-inflorescence” were segmented by assigning a class label to each individual pixel. The trained FCN used achieved a mean Intersection Over Union (IOU) of 87.6% on the test data set. Individual flowers were extracted from the areas representing the inflorescence using Circular Hough Transform. The flower extraction achieved a recall of 80.3% and a precision of 70.7% using the segmentation derived by the trained FCN model (Rudolph et al., 2018). The work presented by Rudolph et al., (2018) has proven an efficient screening of large sets of grapevines. This is important for studies regarding the development of early yield prediction models, and also for objective monitoring and evaluations of breeding material, genetic repositories or commercial vineyards.

In the work of Palacios et al., (2020), it has been done the first attempt to count flowers using images acquired on-the-go from a mobile platform. Images were automatically captured in RGB at night adopting artificial illumination. The developed algorithm have two general steps: inflorescences' segmentation, and individual flower detection. In both steps, the best results were obtained using the deep fully convolutional neural network SegNet architecture with a VGG19 network as the encoder. The F1 score values obtained was of 0.93 and 0.73, respectively in the inflorescences segmentation and the individual flower detection steps. The relationship between the detected number of flowers and the actual number of flowers per vine had an R^2 of 0.91. In this study it have been examined seven varieties (Cabernet Sauvignon, Malvasia, Muscat of Alexandria, Syrah, Tempranillo and Verdejo). In Fig. 8 it's possible to see the relationship between the estimated number of actual flowers and the final yield. It can be observed that the R^2 has a value above 0.70 in all the cases (Palacios et al., 2020)

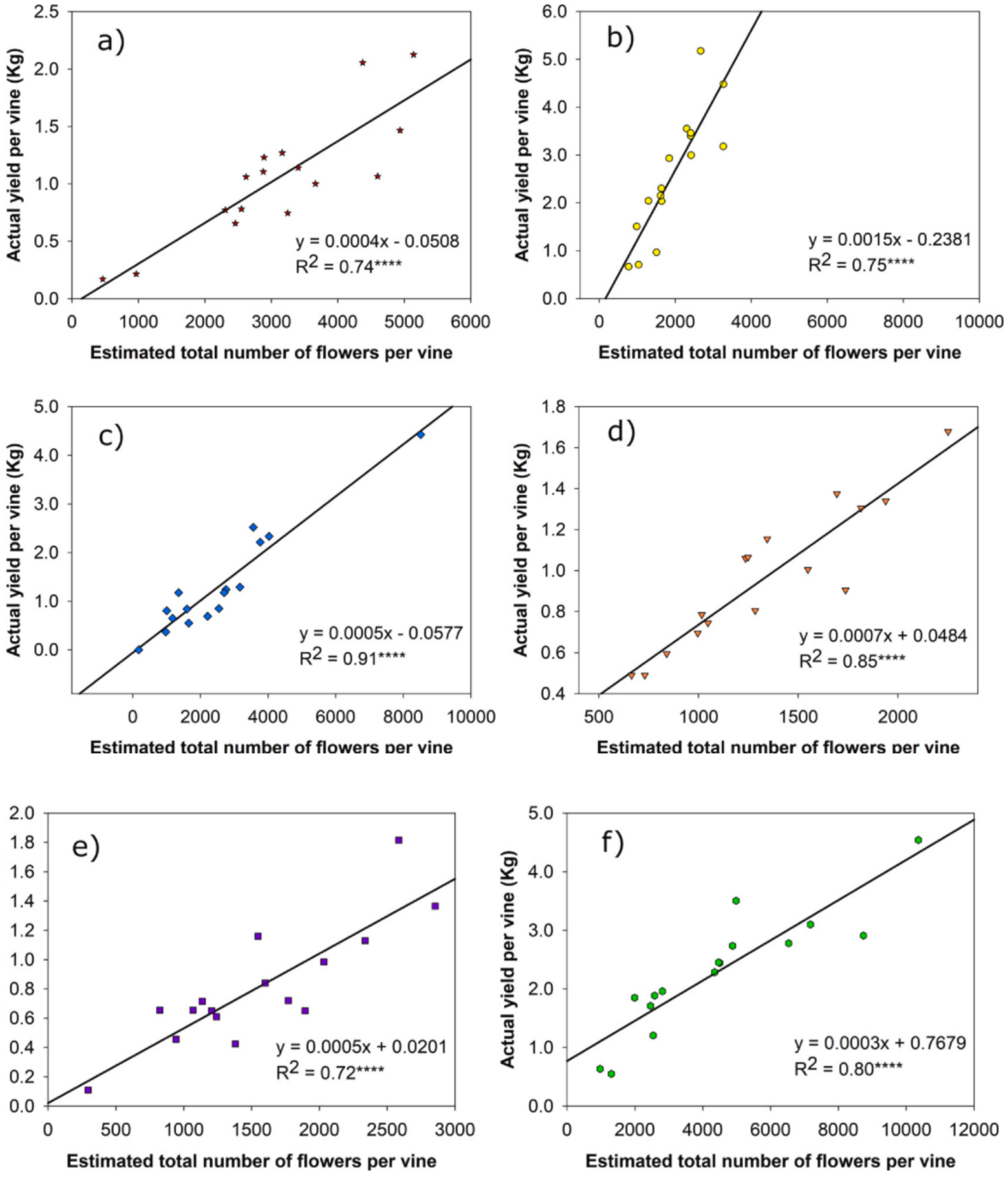


Figure 8. Correlation between the actual number of flowers (estimated by the linear model) and the yield weighted per vine, for each grape variety: a) Cabernet Sauvignon, b) Malvasia, c) Muscat of Alexandria, d) Syrah, e) Tempranillo and f) Verdejo

2.3.2.3 Yield estimation by berry detection with image analysis

Traditional methods for yield estimation try to sample the average number of grape clusters per vine, the average number of grape berries per cluster and average berry weight. Nuske et al., (2011) used a different approach combining clusters per-vine and berries per cluster in the one measurement (berries per vine); since the clusters per vine and berries per cluster account for 60% and 30% of variation in yield per vine respectively, an accurate berries count can reach the 90% of the variation in yield (the remaining 10% is due to berry weight variation). Moreover, the number of berries per vine is fixed from fruit set until harvest. The challenges of automatic berry detection are: their varying appearance under different lighting, the lack of colour contrast to the background, which is often similarly coloured to the grapes, and also occlusions causing not all grapes to be visible (Nuske et al., 2011). The lack of a colour contrast to the background makes berries detection very difficult for white varieties and for all the varieties before veraison; in order to overcome this problem, Nuske et al., (2011) used shape and texture cues for detection. While to overcome the occlusion they have calibrated berry count measurement to harvest yield from a set of vines, and apply this calibration to other vines not included in the calibration set, pointing to the fact that percentage of berries not detected is relatively constant from vine to vine.

The detection was done with a a sideways-facing camera and lighting on a small vineyard utility vehicle. The images captured the vines and were processed with an algorithm which works in three different stages: (1) detecting potential berry locations with a radial symmetry transform;(2) Identifying the potential locations that have similar appearance to grape berries; (3) group neighboring berries into clusters. Using the automatically berry counts with the actual harvest crop weights they obtained a linear relationship with an R2 value of 0.74. The average error obtained with their method was 9.8% of the actual harvest weight that already exceed what is possible obtain with the traditional practises used to estimate the yield (Nuske et al., 2011). The number of berries found by the algorithm (N_b^d) is the measurement that is passed to a yield forecasting function, f (.) which outputs an estimate (N_b) of the actual berry count (Eq. 10):

$$N_b = f(N_b^d) \qquad \text{Eq. 10}$$

In order to have an accurate estimation, the function f must model several biases that are inherent to the visual detection process; these biases were treated together as a single first order linear factor (Nuske at al., 2011). In a later work, Nuske et al. (2014) presented an extension of their studies and, in order to get a deeper understanding of the whole system, have considered the individual causes of bias. The authors have considered three different types of occluders that makes many berries invisible to the camera and bias the counts: self-occlusions; cluster-occlusions; vine-occlusions. Moreover, them they have introduced two bias due to visual detection process: detection (false positive and false negative in the detection of the berries); misregistration (errors, that can occur after several overlapping of the images, of double-counted or mistakenly not counted berries) (Nuske et al., 2014 a). Then the approach is consisted to combine all these bias terms as a linear factors. Once the grapes have been harvested and the berries counted (manually), the result has been correlated with that of three different image measurement: visible berry count, ellipsoidal model and convex hull model.

The visible berry count presented a high correlation coefficient with final yield and a low mean square error. The ellipsoidal model presents a high correlation score and the lowest mean squared error of 17%. The final image measurement model evaluate is the convex hull which had a R^2 = 0.92 which is the best of the three image measurements (Nuske et al.,2014 a)

In another work, Nuske et al. (2014 b) used a different approach: they worked during the night in order to get a better control on the imaging, since thanks to artificial illumination, imaging configuration is improved. Moreover, instead that take into consideration only one cues as has been done in other studies (Diago et al., 2012; Dunn and Martin, 2004; Rabatel and Guizard, 2007; Grossette et al., 2012) they focused on the three main visual cues that grape berries have in order to detect the berries in a variety of condition. Another innovation in their method is the introduction of a way to eliminate the double-counting due to overlapping imagery and also the challenge of geometrically referencing the measurements by estimating camera position along the row. The linear relationship between automatically generated berry counts and the actual harvest crop weight obtained a R^2 score ranged from 0.6 to 0.73 depending on dataset (Nuske et al., 2014 b). The main problem of this method is the requirement of highly specialized technologies, which limits its spread (Liu et al., 2017).

Aquino et al. (2018) have presented a solution based on image analysis for an early yield estimation at phenological stages previous to veraison, about 100 days before harvest. They have used an all-terrain vehicle modified with equipment to autonomously capture images, working during the night. In Fig. 9 it is possible to see how the algorithm works. This image analysis algorithm analyzes images based on mathematical morphology and pixel classification.

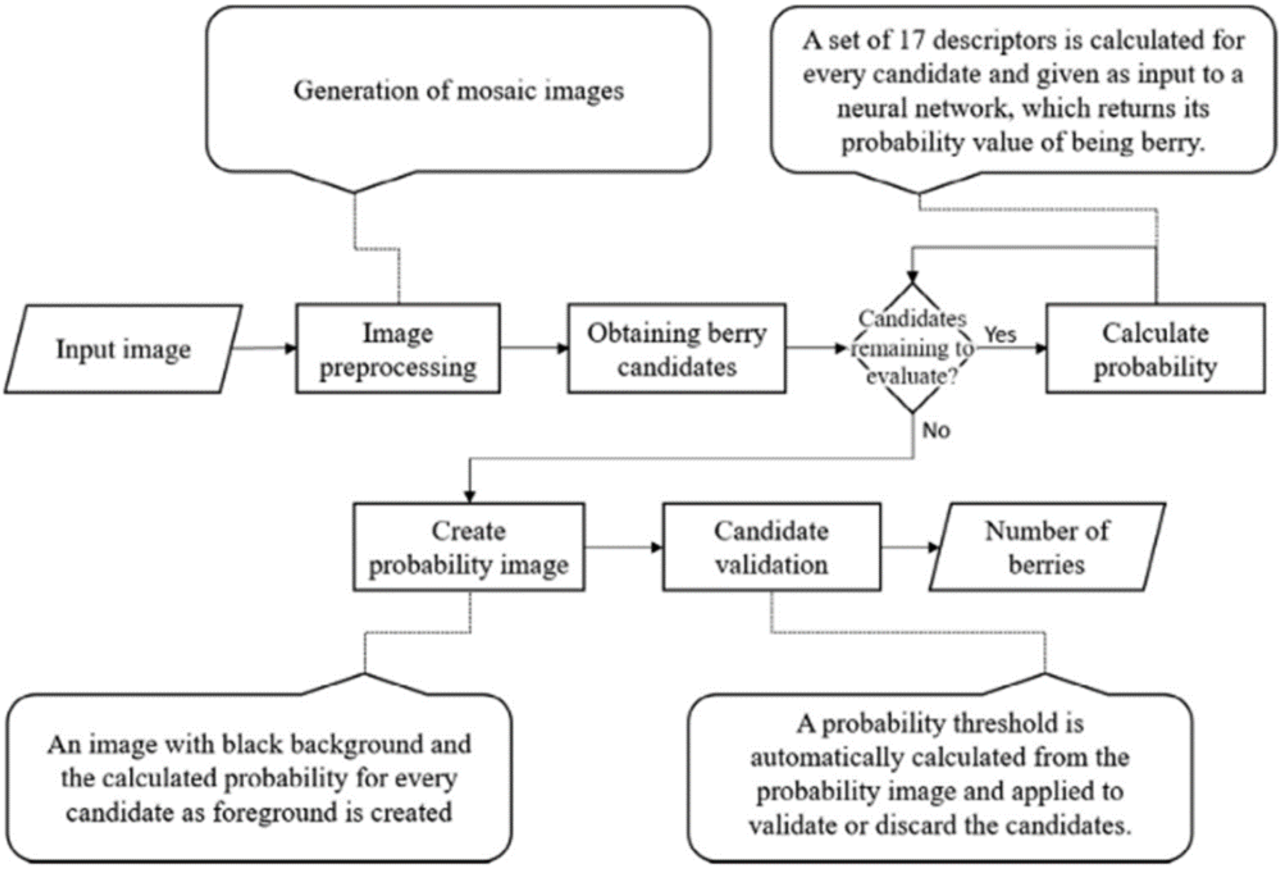


Figure 9. Operating method of the algorithm (Aquino et al., 2018)

Image analysis was carried out on five different varieties, obtained an overall average Recall and Precision values of 0.88 and 0.96, respectively. The number of detected berries was used for yield estimation. First of all an average berry weight per variety was calculated from samples taken the previous season. These average berry weights were then combined by multiplication with the number of visible berries detected per segment by the algorithm. Thus, the 'visible' yield per segment was predicted. These 'visible' weights per segment were correlated to the corresponding actual manually weighted yields to obtain a linear estimation model. Finally, this linear estimation model was applied to estimate the actual yield per segment and variety. Yield predictions was done using the Root-Mean-Square-Error. An average square error of 0.16 kg was obtained per vine (Aquino et al., 2018).

Milella et al.,(2019) have developed a method to get a canopy volume estimation and a clusters detection and counting. They used an agricultural vehicle equipped with RGB-D sensor, i.e. a depth sensing device that works in association with a RGB camera by augmenting the conventional image with distance information in a per - pixel basis. Grape clusters detection has been done using four convolutional neural network (CNN): namely the pre-trained AlexNet (Krizhevsky et al., 2012), VGG16 and VGG19 (Simonyan and Zisserman, 2014), and GoogLeNet (Szegedy et al., 2015); all these CNN allow for labelling input RGB images among 1000 different classes of objects. Field test results show that the proposed methods are able to correctly detect fruits, with all the used CNN; the maximum accuracy (91.52%) was obtained by the VGG19 deep neural network (Milella et al., 2019).

Even in the study of Santos et al., (2019) have been demonstrated that using a CNN the grape berries can be successfully detected. In a test set containing 408 grape clusters from images taken on a trellis-system based vineyard, they have reached an F1-score up to 0.9 for instance segmentation, a good separation of each cluster from other structures in the image, which allows a more accurate assessment of fruit size and shape. They also showed that 3-D models produced by structure-from-motion or SLAM can be employed to track fruits, avoiding double counts and increasing tolerance to errors in detection (Santos et al., 2019).

Zabawa et al., (2019), used a fully convolutional neural network on images acquired with Phenoliner (a field phenotyping platform). After images capture, every single berry was counted; the pixels were divided in three categories: “berry”, “edge” and “background”. The differentiation of the two categories “berry” and “edge” enables the discrimination between single berries and clusters, while remaining pixels are assigned to the class “background”. A connected component algorithm was applied to determine the number of berries in one image. Authors compared the automatically counted number of berries with the manually detected berries in two different training systems: vertical shoot positioned trellis (VSP) and semi minimal pruned hedges (SMPH). They obtained very good results. In fact, the accuracy of the detection for the VSP was 94.0%, while for the SMPH was 85.6%.

Di Gennaro et al., (2019) developed a fast and automated technology for automatic berry detection and counting using an UAV (Unmanned Aerial Vehicles). The UAV approach allows a fast monitoring of the vineyard taking into account a few images acquired in representative points of vineyard variability, which is fundamental to export this application to a large vineyard with moderate times for monitoring and image processing. The authors acquired two datasets: the first one in worst condition (W), with clusters covered by leaves, the second one in best condition (B), with leaves partially removed and cluster directly illuminated by the sun. An unsupervised recognition algorithm was applied to derive cluster number and size, which was used for estimating yield per vine. The classification performance of the unsupervised methodology was evaluated through the true positive rate (TPR) index (Eq. 11):

$$TPR = \frac{(\text{true cluster automatically classified /total clusters observed})}{100} \qquad \text{Eq. 11}$$

The UAV methodology applied in best condition (B) identified 100.0% of ripe grapes, while poorer performances were found with worst condition (W). The estimated yield was validated with the sampling obtained manually. The results were good, with a 12 % under-estimation in yield. (Di Gennaro et al., 2019). Despite these satisfactory results, this methodology have some drawback. Indeed, the authors have made a cost-benefit analysis and they say that in order to depreciate the fixed-cost investments the minimum vineyard size is 60 ha, then in country like the Italy (where the study was carried out) where the company that have more than 50 ha represent only the 0.2% (AGEA: Agenzia per le erogazioni in agricoltura, 2008), the diffusion of this methodology is very difficult. Moreover, results are good only when partial defoliation is practiced, then still remains the problem to estimate what is not possible to see.

Coviello et al., (2020), adapted Deep Learning algorithms, originally developed for crown counting, to develop a tool (Grape Berries Counting Net (GBCNet)) for smartphone cameras for vineyard yield estimation. GBCNet doesn't require any specific preparation for the images acquisition, enabling an easier and faster application. Thank to that, the authors have tested two different strategies: one based on the evaluation of the average number of berries per bunch and another taking a picture of the whole grape field, estimating the total number of berries and then simply multiply this for the average berry's weight. After the manual berries counting, it was evaluated the performance of GBCNet on seven different varieties. That has been done taking in consideration the Mean Absolute Error (MAE) and Mean Squared Error (MSE), which are two metrics that represent a measure of accuracy (MAE) and robustness (MSE) of the model. The authors obtained satisfactory results on all the seven varieties , although with a different accuracy level depending on the variety, in fact the MAE ranges from 0.85% for Pinot Gris to 11.73% for Marzemino. To estimate the final yield, the authors have proposed the following equation:

$$Y = N_v * N_b * N_a * P_a \qquad \text{Eq. 12}$$

Where: N_v is the number of vines per surface unit, N_b is the number of grape bunches per vine, N_a is the average number of berries per bunch and P_a is the average berry's weight (Coviello et al., 2020)

Liu et al., (2020) proposed an algorithm for 3D bunch reconstruction based on a single image for fast berry counting in vineyards. According to the flowchart in fig.10, the proposed approach starts with sub-bunch detection (a sub-bunch refers to the separate sections of the bunches which are visually disconnected in the image other than by rachis structure) then processes each sub-bunch before calculating the bunch sparsity factor (this step is mandatory, because the initial 3D model is built based on the assumption that the processed bunch is healthy and compact, this means that berries are solidly packed in the initial 3D bunch model and thus there is no space for any of the rachis structure, resulting in an overestimate of berry number) and reconstructing the 3D bunch model. In the proposed method, berries counting procedure is divided in two steps, which are parallel to each other: one provides for obtaining the initial 3D bunch model to get a berry number ; another provides fot calculating the sparsity factor based on the difference of two colour channels. Estimation of the bunch weights reached the 92% accuracy (Liu et al.,2020)

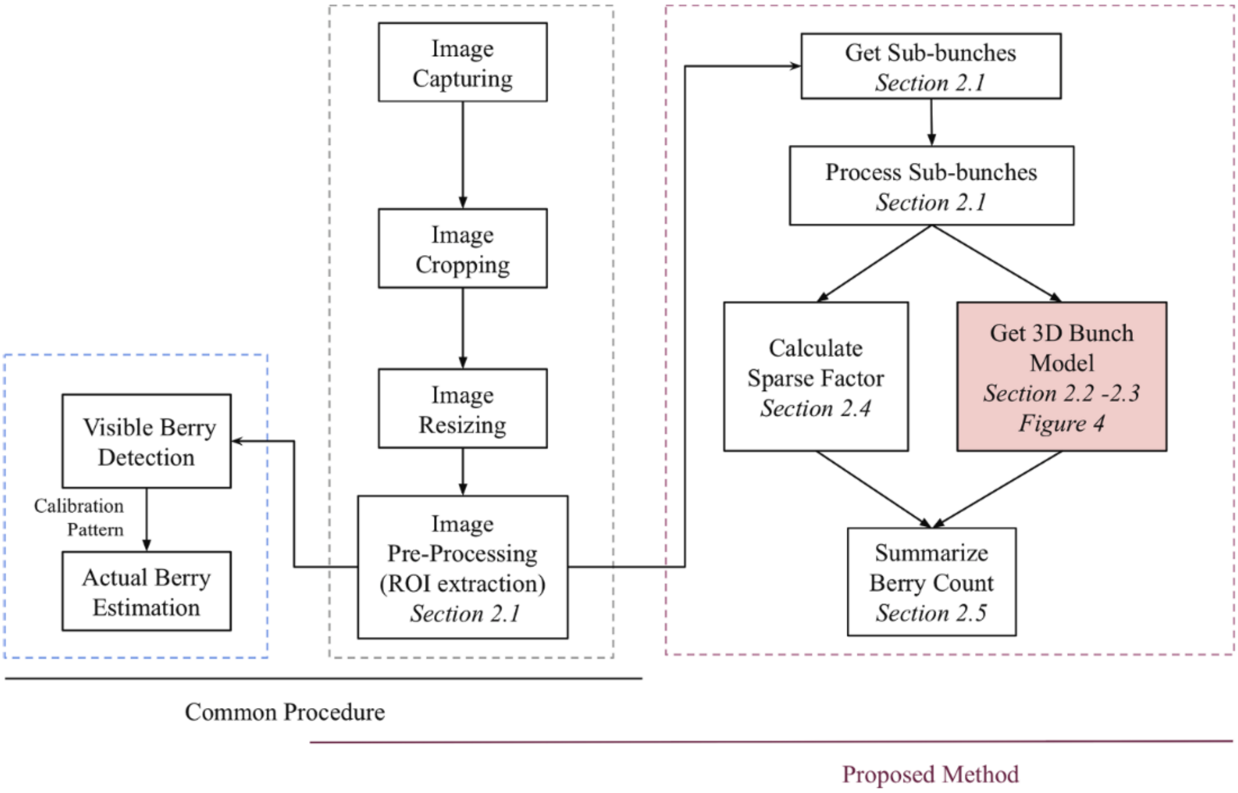


Figure 10. Berry counting approach by 3D bunch reconstruction (source: Liu et al.,2020)

The final yield has been determined, in three blocks, using two different methods: one using historical berry weights and another measuring the berry weights close to the harvest. The yield estimation made measuring the berries weight close to harvest was estimated within 6%, 16% and 3% of the actual yield respectively. While using the historical data, the error increase as there was a notable variation in average berry weight (Liu et al.,2020). The proposed methodology, needs the application of a dark background, the authors are trying to eliminate this necessity in order to make faster the whole process, moreover they are converting this algorithm into an app for mobile device (i.e. smartphone), in order to make it accessible to all farmers.

Victorino et al.,(2020) explored single bunch images taken in lab conditions to obtain the best visible bunch attributes to use as yield indicators. This has been done for three varieties (Encruzado, Arinto and Syrah), at different phenological stages. Authors observed that during the season there is an increase in bunch volume and weight for varieties Encruzado and Arinto, while for the Syrah there is a decrease, in these two parameters, from veraison to harvest, probably caused by water stress during berry ripening combined with berry dehydration close to harvest. Number of berries varies lightly during the season, while the length of the bunch remained stable for all the varieties. As for the bunch perimeter, it decreased from pea-sized berries to veraison for Arinto and Syrah varieties and then remained stable until harvest. Authors also have investigated the correlation between bunch attributes and bunch weight. All the bunch attributes considered, in each variety (area, number of berries, volume, max length, perimeter, weight) have a high coefficient of correlation with the bunch weight, in every stage considered. There are just some exceptions, for example the bunch max length for the Encruzado variety at the veraison stage, which presented a lower r-value.

In the same work, was evaluated the correlations between all the yield components present in this thesis (spurs, shoots, inflorescence, bunches) and final yield. Authors reported that the number of spurs presents a very low r with yield (only Arinto have a significant relationship). Regarding the number of shoots, only Syrah obtained an high and significantly positive r value with the final yield, while Arinto and Encruzado showed a lower and non-significant r. In all varieties, the number of inflorescences and bunches was significantly correlated with yield, with the highest r being observed on the number of bunches of the variety Syrah (Victorino et al.,2020). Results obtained indicate that an accurate detection of some of these yield components, can be useful for the estimation of the final yield (Victorino et al., 2020).

2.3.2.4 Methods to estimate bunch occlusion

Nuske et al., (2011 b) have proposed two methods for the calibration of the bunch occlusion ratio. The first is based on an destructive hand sample: vines have been imaged first, so on a small sample of vines the fruit was destructively removed and counted. Rather than counting individual berries, it has been measured the average berry weight and the total fruit weight of a vine. The process was repeated for all vines in the sample set. The hand fruit weight was projected to harvest using the ratio between current berry weight and expected berry weight at harvest. Taking the manual estimate against the image berry count for these specific vines, it was obtained an occlusion ratio that can be estimated well in advance of harvest, and applied to predict yield of the remaining vines that were not destructively sampled.

The second method is based on previous year harvest data as harvest data from vines trained and managed in a similar manner from year to year, presented a similar occlusion ratio. This method may not be applicable to all vineyards, especially in those where the occlusion may change during the season. However, for vineyards trained with VSP, that use leaf pulling for sun exposure, it was found a consistent occlusion ratio from year to year. The main advantage of calibrating from a prior harvest season is that hand samples are not necessary. The method simply takes a total measurement of the fruit harvested and compared against the berry count detected with images (Nuske et al., 2011 b).

As said before, the leaves can influence bunch detection, thus the accuracy of yield estimation. To determine the bunch occlusion by leaves, (Bonaria, 2019) have used two models: the first model, implemented the algorithm proposed by Lopes et al., (2017), which takes into account the percentage of porosity into the ROI, since higher percentages of porosity correspond to higher percentages of visible bunches. The linear regression between the fraction of visible bunches (dependent variable) and porosity (independent variable) of this model showed an R2 at pea-size, veraison and maturation respectively of 0.75, 0.88 and 0.88. These results show that the measurement of this parameter can be used to estimate the percentage of visible bunches area, thus to estimate the total area of bunches without leaf occlusion in all three phenological stages considered.

Canopy porosity is an important canopy feature as it is an indicator of bunch exposure and of the airflow circulation (Diago et al., 2019). Bunch exposure must be balanced, because if on one hand sun exposure improve the synthesis of aroma/flavour compounds (Reynolds and Wardle 1989, Diago et al., 2010) on the other hand, an excessive exposure can lead to berry sunburn and a loss of grape skin anthocyanins, hence berry colour (Kliewer 1977, Mori et al. 2007). According to Paliotti and Silvestroni (2004) the ideal canopy porosity values range from 10% to 20%, while for Coombe, (1987) the ideal range is from 20% to 40%. The standard method to evaluate grapevine canopy porosity is the Point Quadrat Analysis (PQA). The PQA consists in using a probe, which must be inserted through the canopy, and counting the number and parts of the vine with which the probe comes into contact: leaves, clusters, cane, gaps. The proportion of gaps can be quantified by dividing the number of gaps by the total number of insertions (Diago et al., 2016). However, PQA is laborious and time consuming, because, as recommended by Smart, (1987), the minimum number of passages through the canopy to accurately establish the gaps percentage is 50. Moreover, there is a possibility of damaging the fruit, thus limiting its utility to the wine industry. For these reasons Diago et al., (2016) have developed a new, objective, non-invasive method based on image analysis which has been compared with PQA. The authors have developed an adapted protocol for in-field image acquisition based on the method described by Tardaguila et al., (2010). Image analysis (as describe in Diago et al., 2012) was made with a clustering algorithm based on the Mahalanobis distance, in order to identify the pixels corresponding to the canopy gaps (porosity). The Mahalanobis distance measures the similarity between an unknown sample group and a known one; it takes into account the correlations of the data set with different variances for each direction, and it is scale-invariant. The modified algorithm used a known sample of values to classify an unknown batch of pixels into classes, based on a characteristic vector, in this case the RGB colour values of each pixel. The determination coefficient of the regressions between the percent of gaps, using both methods (image analysis and PQA), exceeded 0.90 ($p<0.05$) in each site, and R2 of the global regression was 0.93 ($p<0.05$) (Diago et al., 2016).

Diago et al. (2019) have used an on-the-go real time image acquisition and processing method to evaluate canopy porosity. They used the algorithm described in Diago et al. (2012), but in this case the hue saturation value (HVS) was included. Correlations between the image analysis and point quadrat analysis results for the proportion of gaps (R2 > 0.85; $p < 0.001$) and leaf to canopy area ratio ($R^2 > 0.57$; $p < 0.001$) were obtained for both sides of the canopy. For the bunch to canopy area ratio the best relationship was found on the western side of the canopy ($R^2= 0.79$; $p < 0.001$) (Diago et al., 2019).

In Victorino et al.,(2020) was evaluated the visibility of the yield components throughout the vine growing cycle. The graph shows that spurs and early-stage shoots are the most visible yield component during the growning season, this was predictable given the lack of leaves early in the season. Regarding inflorescences, can be observe that them were more visible in the Syrah (occlusion = 59 %) than in the Arinto (68 %) and Encruzado (64 %). Another fact that can be observed regarding inflorescences occlusion, it is that at the stage of flowers separated, for Arinto the occlusion decreases by about 20%, while remain stable for the other two varieties; this can be attribute to inflorescence size of Arinto, in fact at this stage, they are wide open with several blank spaces within, thus being easier to detect than other varieties. About the bunch occlusion, results shown that it was very variable for all the varieties. At the stage of pea-sized, the occlusion is similar to the previous stage for the varieties Encruzado and Syrah, but this is not true for Arinto, probably due to this variety's previous occlusion drop. From the pea-size stage to the harvest, the visibility remained stable, with a slightly decrease close to the harvest. However, for the Syrah this decrease is more evident. That is due to the high water stress to which this plot was subjected, that accelerated the leaves senescence.

As said before, bunch occlusion can influence the accuracy of the yield estimation, thus is important to estimate this parameter. Bonaria, (2019) used another model that had the objective to estimate the weight of the bunches, it was built using the total bunches area per linear canopy meter, which was obtained by adding to the total projected bunches area without leaf occlusion, i.e. the calculated percentage of bunch on bunch occlusion per linear canopy meter, and the total bunches weight per linear canopy meter. The R^2 value of the linear regression between the projected visible bunches area (independent variable) and the percentage of visible bunches (dependent variable) at pea-size, veraison and maturation was respectively of 0.85, 0.57 and 0.69.

In the same work of Bonaria, (2019) it was estimated the bunch by bunch occlusion in order to obtain an estimate of the whole area of the bunches occluded by the other bunches, in this way it is possible to have a better estimation of the total area of the bunches per linear canopy meter (LCM) and then a better estimation of the yield. On the images collected to evaluate the bunch by bunch occlusion (Fig.9), the fruiting areas (all visible bunches in the picture) composed by all layers together and the areas of the bunches layers in first plane were projected. From the image with all layers (Fig. 11A) we can get the union area ($A \cup B$), that represents the union between the bunches area of the first layer (A) (Fig. 11B) and the area of the bunches in background (B) (Fig. 11C); (A) was get from the same image from where was get ($A \cup B$), then from fig. 11A. Instead (B) was get from the image took after the removal on the first layer, so (B) represents also ($A \cup B$) but of the second set of photographies. Then to calculate the percentage of bunch occlusion, it was necessary the calculation of the intersection between the two sets ($A \cap B$) that can be calculated with equation:

$$A \cap \mathrm{B} = \mathrm{A} + \mathrm{B} - \mathrm{A} \cup \mathrm{B} \qquad \text{Eq. 12}$$

The intersection ($A \cap B$) represents the area of (B) covered by (A), to obtain the percentage of bunch by bunch occlusion, we have just to use the equation 13:

$$\% \mathrm{B} \times \mathrm{B} = \frac{\mathrm{A} \cap \mathrm{B}}{A \cup B} * 100 \qquad \text{Eq. 13}$$

To obtain the total percentage of bunch by bunch occlusion it must sum the percentages of occlusion caused by each layer.

Figure 11. (A) Selected area of bunches representing the $A \cup B$; (B) area of the layer that was took out from the plant representing the set A; (C) area of bunches without the occlusion caused by the layer took out representing B. (Source: Bonaria, 2019)

The values of the average percentage of bunch-by-bunch occlusion were 8%, 6% and 12% respectively at pea-size, veraison and maturation. Then to estimate the yield at harvest from the estimated weight for the different phenological stages it were used the growth factors.

3. Conclusion

This review presents the state-of-the-art of the use of image analysis for vineyard yield estimation. Methodologies can be divided according to when they are applied during the growing season, then we have some methods applied before flowering, others between flowering and veraison and others from veraison to harvest. Early forecastes can be very helpful because vineyard managers have more time to adjust the final yield, for example to obtain the desired quality. However, a yield estimation made closer to the harvest (for example at veraison) is more likely to be more accurate. Another difference between the methods lies in the type of algorithm used, which can be supervised or unsupervised learning. Methodologies containing a supervised learning approach, albeit with excellent results, require a lot of work in the laboratory to train the algorithms; on the other hand, with unsupervised learning the results obtained are in any case excellent, but, in addition, there is no need (or at least very little) for laboratory work. Among the different methods it is possible to see that fruit detection has high percentages of accuracy. The challenges to be overcome are still different: accurate calibration of bunch occlusion; calibration of various methods, suitable for every situation; above all the estimation of the bunches or berries that cannot be seen by the cameras. For this reason, several studies are focusing their attention on this aspect.

The needed to calibrate different methods for any situation is due to the fact that different training systems lead to different difficulties in the yield estimation. For example, in the work of Zabawa et al., (2020) were investigated two training systems; vertical shoot positioning training system, presented compact bunches and few overlapping leaves; this because there is only one main branch and many leaves are cut down (at least in those areas where the heat is not excessive), moreover, most grapes grown on the bottom part of the canopy. While, the semi minimal pruned hedge training systems had many branches and a thick layer of canopy, thus berries can be in all places of the canopy.

Thanks to the continuous improvement in the field of machine learning and the development of increasingly efficient robotic monitoring technologies and platforms, these technologies will be increasingly used in all fields and viticulture cannot be an exception, as thanks to their use it is it is possible to reduce production costs and more easily reach own quality objectives. All this is mandatory in a market that sees the continuous increase of international competition, even from those countries where, until not many years ago, it was thought that the cultivation of vines was not possible. The main problem with these techniques lies in the fact that many of them are not economically sustainable by most wineries. In addition to the expensive and highly accurate analytics instruments used in the lab, sensors on portable devices (among which the most attractive are the smartphone) are constantly being developed in precision viticulture in order to support quality control, to reduce costs and obtain results that are comparable to the ones obtained in labs with traditional technologies. Another aspect to take into consideration is the request from the winegrowers to complement the yield estimation with the monitoring of grape's quality, so there is the need to combine these two tools.

4. References

AGEA: https://www.agea.gov.it/portal/page/portal/AGEAPageGroup/HomeAGEA/home (Accessed 10 November 2020)

Aquino, A., Millan, B., Diago, M. P., & Tardaguila, J. (2018). Automated early yield prediction in vineyards from on-the-go image acquisition. *Computers and electronics in agriculture*, *144*, 26-36. doi: https://doi.org/10.1016/j.compag.2017.11.026

Aquino, A., Millan, B., Gaston, D., Diago, M. P., & Tardaguila, J. (2015). VitisFlower®: Development and testing of a novel android-smartphone application for assessing the number of grapevine flowers per inflorescence using artificial vision techniques. *Sensors*, *15*(9), 21204-21218. doi: https://doi.org/10.3390/s150921204

Aquino, A., Millan, B., Gutiérrez, S., & Tardáguila, J. (2015). Grapevine flower estimation by applying artificial vision techniques on images with uncontrolled scene and multi-model analysis. *Computers and Electronics in Agriculture*, *119*, 92-104. doi: https://doi.org/10.1016/j.compag.2015.10.009

Barbedo, J. G. A. (2019). Plant disease identification from individual lesions and spots using deep learning. *Biosystems Engineering*, *180*, 96-107. doi: https://doi.org/10.1016/j.biosystemseng.2019.02.002

BLACKMORE, B. (1999). Developing the principles of precision farming [on line]. *Available in http://www. cpf. kvl. dk/Papers [27 September, 2001]*.

Bonaria R. (2019). Grapevine yield estimation using image analysis for the variety Arinto. Dissertation to obtain a Master Degree in Viticulture and Enology Engineering. Instituto Superior de Agronomia, Lisboa, 33-34.

Bramley R. Smarter thinking on soil survey. Australian and New Zealand Wine Industry Journal. 2003;18(3):88–94. Available from http://iwrdb. org/cgi-bin/koha/opac-detail.pl?biblionumber=30800.

Bramley, R. G. V. (2005). Understanding variability in winegrape production systems 2. Within vineyard variation in quality over several vintages. *Australian Journal of Grape and Wine Research*, *11*(1), 33-42. https://doi.org/10.1111/j.1755-0238.2005.tb00277.x

Carbonneau, A., Deloire, A., & Jaillard, B. (2007). *La vigne: physiologie, terroir, culture*. Dunod.

Clingeleffer, P. R., Martin, S. R., Dunn, G. M., & Krstic, M. P. (2001). Crop development, crop estimation and crop control to secure quality and production of major wine grape varieties: a national approach: final report to Grape and Wine Research & Development Corporation/principal investigator, Peter Clingeleffer;[prepared and edited by Steve Martin and Gregory Dunn]. https://publications.csiro.au/publications/publication/PIprocite:e3b20ec6-9806-471c-a8f1-939515a1f626

Commission Internationale d'Eclairage. (2007). CIE Colorimetetry. Part 4: 1976 L* a* b* colour space. Vienna, Austria: CIE.

Cour, P., & Villemur, P. (1986). Fluctuations des émissions polliniques atmosphériques et prévisions des récoltes de fruits. http://pascal-francis.inist.fr/vibad/index.php?action=getRecordDetail&idt=8367926

Coviello, L., Cristoforetti, M., Jurman, G., & Furlanello, C. (2020). GBCNet: In-Field Grape Berries Counting for Yield Estimation by Dilated CNNs. *Applied Sciences*, *10*(14), 4870. https://doi.org/10.3390/app10144870

Cunha, M., Abreu, I., Pinto, P., & de Castro, R. (2003). Airborne pollen samples for early-season estimates of wine production in a Mediterranean climate area of northern Portugal. *American Journal of Enology and Viticulture*, *54*(3), 189-194. https://www.ajevonline.org/content/54/3/189

Di Gennaro, S. F., Toscano, P., Cinat, P., Berton, A., & Matese, A. (2019). A low-cost and unsupervised image recognition methodology for yield estimation in a vineyard. Frontiers in plant science, 10, 559. https://doi.org/10.3389/fpls.2019.00559

Diago, M. P., Aquino, A., Millan, B., Palacios, F., & Tardáguila, J. (2019). On-the-go assessment of vineyard canopy porosity, bunch and leaf exposure by image analysis. *Australian journal of grape and wine research*, *25*(3), 363-374. https://doi.org/10.1111/ajgw.12404

Diago, M. P., Krasnow, M., Bubola, M., Millan, B., & Tardaguila, J. (2016). Assessment of vineyard canopy porosity using machine vision. *American Journal of Enology and Viticulture*, *67*(2), 229-238. DOI: 10.5344/ajev.2015.15037

Diago, M. P., Sanz-Garcia, A., Millan, B., Blasco, J., & Tardaguila, J. (2014). Assessment of flower number per inflorescence in grapevine by image analysis under field conditions. *Journal of the Science of Food and Agriculture*, *94*(10), 1981-1987. https://doi.org/10.1002/jsfa.6512

Diago, M. P., Tardaguila, J., Aleixos, N., Millan, B., Prats-Montalban, J. M., Cubero, S., & Blasco, J. (2015). Assessment of cluster yield components by image analysis. *Journal of the Science of Food and Agriculture*, *95*(6), 1274-1282. https://doi.org/10.1002/jsfa.6819

Diago, M. P., Vilanova, M., & Tardaguila, J. (2010). Effects of timing of manual and mechanical early defoliation on the aroma of Vitis vinifera L. Tempranillo wine. *American Journal of Enology and Viticulture*, *61*(3), 382-391. https://www.ajevonline.org/content/61/3/382

Dike, H. U., Zhou, Y., Deveerasetty, K. K., & Wu, Q. (2018, October). Unsupervised learning based on artificial neural network: A review. In 2018 IEEE International Conference on Cyborg and Bionic Systems (CBS) (pp. 322-327). IEEE. 10.1109/CBS.2018.8612259

Dry, P. R., Longbottom, M. L., McLoughlin, S., Johnson, T. E., & Collins, C. (2010). Classification of reproductive performance of ten winegrape varieties. *Australian Journal of Grape and Wine Research*, *16*, 47-55. https://doi.org/10.1111/j.1755-0238.2009.00085.x

Dunn, G. M., & Martin, S. R. (2004). Yield prediction from digital image analysis: A technique with potential for vineyard assessments prior to harvest. *Australian Journal of Grape and Wine Research*, *10*(3), 196-198. https://doi.org/10.1111/j.1755-0238.2004.tb00022.x

Galet, P. (1983). Precis de Viticulture. 4ème Ed. Imprimerie Charles Dehan. Montpellier. *Francia*.

Grossetete, M., Berthoumieu, Y., Da Costa, J. P., Germain, C., Lavialle, O., & Grenier, G. (2012, July). Early estimation of vineyard yield: site specific counting of berries by using a smartphone. In *International Conference of Agricultural Engineering—CIGR-AgEng*.

Guzman, R., Navarro, R., Beneto, M., & Carbonell, D. (2016). Robotnik—Professional service robotics applications with ROS. In *Robot Operating System (ROS)* (pp. 253-288). Springer, Cham.

He, X., Cai, D., & Niyogi, P. (2006). Laplacian score for feature selection. In *Advances in neural information processing systems* (pp. 507-514).

Higgins, J. J. (2003). Introduction to Modern Nonparametric Statistics . Cengage Learning.

Hinton, Geoffrey; Sejnowski, Terrence (1999). Unsupervised Learning: Foundations of Neural Computation. MIT Press. ISBN 978-0262581684.

International Society of Precision Agriculture: https://www.ispag.org/about/definition (Accessed 03/03/2021)

Kendall, M. G. (1946). The advanced theory of statistics. *The advanced theory of statistics.*, (2nd Ed). Charles Griffin & Company Limited, London

Kliewer, W. M. (1977). Effect of high temperatures during the bloom-set period on fruit-set, ovule fertility, and berry growth of several grape cultivars. *American Journal of Enology and Viticulture*, *28*(4), 215-222. https://www.ajevonline.org/content/28/4/215

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105).

Liu, S., Cossell, S., Tang, J., Dunn, G., & Whitty, M. (2017). A computer vision system for early stage grape yield estimation based on shoot detection. *Computers and Electronics in Agriculture*, *137*, 88-101. https://doi.org/10.1016/j.compag.2017.03.013

Liu, S., Zeng, X., & Whitty, M. (2020). A vision-based robust grape berry counting algorithm for fast calibration-free bunch weight estimation in the field. *Computers and Electronics in Agriculture*, *173*, 105360. https://doi.org/10.1016/j.compag.2020.105360

Lopes, C. M., Torres, A., Guzman, R., Graça, J., Reyes, M., Vitorino, G., ... & Barriguinha, A. (2017). Using an unmanned ground vehicle to scout vineyards for non-intrusive estimation of canopy features and grape yield. In *GiESCO International Meeting, 20th, Sustainable viticulture and wine making in climate chenge scenarios, 5-10 November 2017*. GiESCO.

Martin, S., Dunstone, R., & Dunn, G. (2003). How to forecast wine grape deliveries using grape forecaster excel workbook version 7. *GWRDC, Adelaide, Australia*, 100.

Matese, A., & Di Gennaro, S. F. (2015). Technology in precision viticulture: A state of the art review. *Int. J. Wine Res*, *7*, 69-81. https://doi.org/10.2147/IJWR.S69405

May, P. (2004). Flowering and Fruitset in Grapevines Lythrum Press.

Milella, A., Marani, R., Petitti, A., & Reina, G. (2019). In-field high throughput grapevine phenotyping with a consumer-grade depth camera. *Computers and electronics in agriculture*, *156*, 293-306. https://doi.org/10.1016/j.compag.2018.11.026

Millan, B., Aquino, A., Diago, M. P., & Tardaguila, J. (2017). Image analysis-based modelling for flower number estimation in grapevine. *Journal of the Science of Food and Agriculture*, *97*(3), 784-792. https://doi.org/10.1002/jsfa.7797

Mori, K., Goto-Yamamoto, N., Kitayama, M., & Hashizume, K. (2007). Effect of high temperature on anthocyanin composition and transcription of flavonoid hydroxylase genes in 'Pinot noir'grapes (Vitis vinifera). *The Journal of Horticultural Science and Biotechnology*, *82*(2), 199-206. https://doi.org/10.1080/14620316.2007.11512220

Nuske, S., Achar, S., Bates, T., Narasimhan, S., & Singh, S. (2011 a). Yield estimation in vineyards by visual grape detection. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 2352-2358). IEEE. 10.1109/IROS.2011.6095069

Nuske, S., Achar, S., Gupta, K., Narasimhan, S., & Singh, S. (2011 b). Visual yield estimation in vineyards: experiments with different varietals and calibration procedures. *Robotics Institute, Carnegie Mellon University Technical Report: CMURI-TR-11-39, Dec*.

Nuske, S., Gupta, K., Narasimhan, S., & Singh, S. (2014 a). Modeling and calibrating visual yield estimates in vineyards. In *Field and Service Robotics* (pp. 343-356). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-40686-7_23

Nuske, S., Wilshusen, K., Achar, S., Yoder, L., Narasimhan, S., & Singh, S. (2014 b). Automated visual yield estimation in vineyards. *Journal of Field Robotics*, *31*(5), 837-860. https://doi.org/10.1002/rob.21541

Palliotti, A. S. O., & Silvestroni, O. (2004). Ecofisiologia applicata alla vite. *Viticoltura ed Enologia Biologica. Ed agricole*.

Poni S, Casalini L, Bernizzoni F, Civardi S and Intrieri C, Effects of early defoliation on shoot photosynthesis, yield components, and grape composition. Am J Enol Vitic 57:397–407 (2006). https://www.ajevonline.org/content/57/4/397

Poynton, C. A. (1995, February). A Guided Tour of Colour Space. In *New Foundation for Video Technology: The SMPTE Advanced Television and Electronic Imaging Conference* (pp. 167-180). SMPTE. 10.5594/M00840

Rabatel, G., & Guizard, C. Grape berry calibration by computer vision using elliptical model fitting. 6*th European Conference on Precision Agriculture*, Jun 2007, Skiathos, Greece. p. 581 - p. 587.

Reynolds, A. G., & Wardle, D. A. (1989). Influence of fruit microclimate on monoterpene levels of Gewürztraminer. *American Journal of Enology and Viticulture*, *40*(3), 149-154. https://www.ajevonline.org/content/40/3/149

Rudolph, R., Herzog, K., Töpfer, R., & Steinhage, V. (2018). Efficient identification, localization and quantification of grapevine inflorescences in unprepared field images using fully convolutional networks. https://arxiv.org/abs/1807.03770

Sabbatini P., Dami I. and Howell G.S., 2012. Predicting Harvest Yield in Juice and Wine Grape Vineyards. Extension Bulletin 3186, Michigan State University

Santos, T. T., de Souza, L. L., dos Santos, A. A., & Avila, S. (2020). Grape detection, segmentation, and tracking using deep neural networks and three-dimensional association. *Computers and Electronics in Agriculture*, *170*, 105247. https://doi.org/10.1016/j.compag.2020.105247

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. https://arxiv.org/abs/1409.1556

Smart, R. E. (1985). Principles of grapevine canopy microclimate manipulation with implications for yield and quality. A review. *American Journal of Enology and Viticulture*, *36*(3), 230-239. https://www.ajevonline.org/content/36/3/230

Smart, R. E. (1986, August). Influence of light on composition and quality of grapes. In *Symposium on Grapevine Canopy and Vigor Management, XXII IHC 206* (pp. 37-48). 10.17660/ActaHortic.1987.206.1

Soille, P. (2013). *Morphological image analysis: principles and applications*. Springer Science & Business Media. Berlin, Heidelberg

Solomon, C., & Breckon, T. (2011). *Fundamentals of Digital Image Processing: A practical approach with examples in Matlab*. John Wiley & Sons, Inc., Hoboken, NJ, USA

Stuart J. Russell, Peter Norvig (2010) Artificial Intelligence: A Modern Approach, Third Edition, Prentice Hall ISBN 9780136042594.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. Boston, Massachusetts, 2015, pp. 1-9

Tarara, J. M., Ferguson, J. C., Blom, P. E., Pitts, M. J., & Pierce, F. J. (2004). Estimation of grapevine crop mass and yield via automated measurements of trellis tension. *Transactions of the ASAE*, *47*(2), 647. doi: 10.13031/2013.16028

Tardaguila, J., Blanco, J. A., Poni, S., & Diago, M. P. (2012). Mechanical yield regulation in winegrapes: Comparison of early defoliation and crop thinning. *Australian Journal of Grape and Wine Research*, *18*(3), 344-352. doi: https://doi.org/10.1111/j.1755-0238.2012.00197.x

Tardaguila, J., de Toda, F. M., Poni, S., & Diago, M. P. (2010). Impact of early leaf removal on yield and fruit and wine composition of Vitis vinifera L. Graciano and Carignan. *American journal of enology and viticulture*, *61*(3), 372-381.

Van Gerven, M., & Bohte, S. (2017). Artificial neural networks as models of neural information processing. Frontiers in Computational Neuroscience, 11, 114. https://doi.org/10.3389/fncom.2017.00114

Victorino, G., Braga, R., & Lopes, C. M. (2017). The effect of topography on the spatial variability of grapevine vegetative and reproductive components. *Actas Portuguesas de Horticultura, nº 29, p. 510-516*. http://hdl.handle.net/10400.5/16292

Victorino, G., Braga, R., Santos-Victor, J., & Lopes, C. M. (2020). Yield components detection and image-based indicators for non-invasive grapevine yield prediction at different phenological phases. *OENO One*. http://hdl.handle.net/10400.5/20753

Zabawa, L., Kicherer, A., Klingbeil, L., Milioto, A., Topfer, R., Kuhlmann, H., & Roscher, R. (2019). Detection of single grapevine berries in images using fully convolutional neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (pp. 0-0).

VISIONROBOTICS, 2017. Intelligent Autonomous Grapevine Pruner https://www.visionrobotics.com/vr-grapevine-pruner. Accessed 15 November 2020

VITIROVER, 2017. Découvrez le VitiRover https://www.vitirover.fr/. Accessed 15 November 2020