Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/issn/15375110

Research Paper

A robust automated flower estimation system for grape vines

Scarlett Liu [a,b,*], Xuesong Li [a], Hongkun Wu [a], Bolai Xin [a], Julie Tang [a], Paul R. Petrie [c], Mark Whitty [a]

[a] School of Mechanical and Manufacturing Engineering, UNSW, Sydney, NSW 2052, Australia
[b] School of Traffic and Transportation Engineering, Central South University, 210075 Changsha, Hunan, China
[c] The Australian Wine Research Institute and the South Australian Research & Development Institute, Australia

ARTICLE INFO

Article history:
Received 13 May 2017
Received in revised form
25 February 2018
Accepted 25 May 2018
Published online 26 June 2018

Keywords:
Flower counting
Image processing
Grape vine
Computer vision
Grape yield estimation

Automated flower counting systems have recently been developed to process images of grapevine inflorescences, which assist in the critical tasks of determining potential yields early in the season and measurement of fruit-set ratios without arduous manual counting. In this paper, we introduce a robust flower estimation system comprised of an improved flower candidate detection algorithm, flower classification and finally flower estimation using calibration models. These elements of the system have been tested in eight aspects across 533 images with associated manual counts to determine the overall accuracy and how it is affected by experimental conditions.

The proposed algorithm for flower candidate detection and classification is superior to all existing methods in terms of accuracy and robustness when compared with images where visible flowers are manually identified. For flower estimation, an accuracy of 84.3% against actual manual counts was achieved both *in-vivo* and *ex-vivo* and found to be robust across the 12 datasets used for validation. A single-variable linear model trained on 13 images outperformed other estimation models and had a suitable balance between accuracy and manual counting effort. Although accurate flower counting is dependent on the stage of inflorescence development, we found that once they reach approximately EL16 this dependency decreases and the same estimation model can be used within a range of about two EL stages. A global model can be developed across multiple cultivars if they have inflorescences with a similar size and structure.

1. Introduction

Flower number per inflorescence is one of the main determinants of grapevine yield (May, 2000). The number of flowers varies between cultivars, locations and seasons, therefore accurately assessing flower number is a key opportunity to determine the potential yield early in the season (Dry, Longbottom, McLoughlin, Johnson, & Collins, 2010). The manual counting of flowers has been undertaken in both a research (Dry et al., 2010; Dunn & Martin, 2007; Petrie & Clingeleffer, 2005) and industrial context (Dunn & Martin,

* *Corresponding author*. School of Mechanical and Manufacturing Engineering, UNSW, Sydney, NSW 2052, Australia.
E-mail addresses: scarlett.liu@unsw.edu.au (S. Liu), xuesong.li@unsw.edu.au (X. Li), hongkun.wu@unsw.edu.au (H. Wu), bolai.xin@unsw.edu.au (B. Xin), julie.tang@unsw.edu.au (J. Tang), Paul.Petrie@awri.com.au (P.R. Petrie), m.whitty@unsw.edu.au (M. Whitty).
URL: http://www.robotics.unsw.edu.au/srv/team.html
https://doi.org/10.1016/j.biosystemseng.2018.05.009
1537-5110/

Nomenclature

AC	estimation accuracy
$\mathbf{C}$	the classification result obtained by minimising inner distance of classes
e	size of element used to conduct morphological processing
$F1$	F1 score
I	image in RGB colour space
I_b	binary mask
I_{mask}	stem mask
k	K-means clustering algorithm
N_c	true positive value
N_{fn}	false negative value
N_{fp}	false positive value
$\mathbf{P}$	a list containing FREAK feature vectors for all flower candidates
$\mathbf{p}_i$	a single element of $\mathbf{P}$, representing FREAK feature vectors for one flower candidate
PE	percentage error
R^2	coefficient of determination
$\widehat{R}^2$	adjusted R-square value
u	coordinate of pixel in image with respect to horizontal axis
U	width of image in pixels
v	coordinate of pixel in image with respect to vertical axis
V	height of image in pixels
χ	precision
ζ	recall
$\mu_{i,d}$	mean value of a feature in class c_i

2007). In all situations the counting of flowers is an onerous proposition, and many destructive and non-destructive methods have been developed to assist this process. These include placing gauze bags over the developing inflorescences and then manually sorting and counting the flower caps (May, 2000), or photographing the inflorescence and then manually counting the flowers in the image (Poni, Casalini, Bernizzoni, Civardi, & Intrieri, 2006). While these methods allow the counting process to be completed in the laboratory, as opposed to the vineyard, they are still time consuming. Another option has been to count the number of flowers on the first branch of the inflorescence (Bennett, Jarvis, Creasy, & Trought, 2005; May, 1987) or to count the number of branches on the inflorescence (Dunn & Martin, 2007; May, 2000; Shavrukov, Dry, & Thomas, 2004) and relate this to the number of flowers on the inflorescence. These methods require calibration that is likely to vary between varieties, seasons and possibly sampling dates within a season, once again limiting their utility. An easier and more efficient and accurate method to count grape flowers would facilitate research on the impact of management practices on grapevine flowering and the use of flower counts for commercial yield estimation. This has provided the impetus for approaches based on image analysis.

Among the state-of-the-art approaches, the method proposed by (Diago, Sanz-Garcia, Millan, Blasco, & Tardaguila, 2014) based on the extended H-maxima transform has been widely applied, particularly since the method has been implemented in a free smartphone app (Aquino, Millan, Gaston, Diago, & Tardaguila, 2015). In their work, images are first segmented in the LAB colour space and then morphological operations are applied to segmented binary images and the flower number is generated after several filtering operations. At that stage, a strong relationship between detected visible flowers (flowers which are not occluded by stems or other flowers in the image) and actual flowers (total flowers in current bunch) was proven, but the actual ability of the algorithm to estimate the number of flowers per inflorescence was not characterised.

To achieve the goal of estimating the number of flowers per inflorescence by a single image, Aquino, Millan, Gutiérrez, and Tardáguila (2015) further developed visible flower detection and estimation models upon the image processing techniques. Their procedure is sensitive to colour and image size because the required threshold value and the size of structuring element varies significantly between images taken in the field. Despite this both Diago et al. (2014) and Aquino, Millan, Gutiérrez, et al. (2015) set values manually in their algorithms, and when this approach is taken for images such as that in Fig. 1 – with vigorous canopy and green grass in the background the ROI extraction fails. As for the flower segmentation, the main idea displayed by Aquino, Millan, Gutiérrez, et al. (2015) is detecting the peak reflections of flowers by Gaussian pyramidal decomposition. Following visible flower detection, three estimation models and five extracted features were investigated in regards to variety independence. Final R^2 values for the flower estimation ranged from 0.85 to 0.99.

A similar approach was presented in the work Millan, Aquino, Diago, and Tardaguila (2017), who adapted the standard procedure from Diago et al. (2014) but instead of focussing on extracting Regions Of Interest (ROIs), investigated the effects of different features on estimation models and extended the test data to 11 cultivars. The results from these improved models showed R^2 values from 0.19 to 0.99. However, the aforementioned methods are not robust or replicable because some key threshold values in their algorithms are manually set, despite the attainment of high R^2 values. Furthermore, multiple peaks existing on one flower candidate will lead to inaccuracy in flower counting. Therefore, some post-processing is required, including the elimination of false-positive detections and adding false-negative flowers.

Fig. 1 – An example of canopy in a vineyard in Australia, showing the challenging conditions faced by flower detection methods. The flagging tape visible was used for marking bunches and cordon in a separate experiment.

The assessment criteria for estimation results in both papers does not sufficiently characterise the performance of the algorithms; in particular, the extremely high R^2 values generated from estimation models are misleading and the Root Mean Square Error (RMSE) is meaningless without the presence of the mean value of the dataset. Providing another statistically correct benchmark is necessary to assess the performance of the flower estimation.

In addition to the large variation between the morphology of grapevine inflorescences between different varieties (as shown in Fig. 2) (Shavrukov et al., 2004), they also expand and grow very rapidly in the lead up to flowering (May, 2000). Figure 3 demonstrates how the form of the inflorescence changes over time. This change in morphology is likely to perturb the relationship between visible and actual flowers even if the same inflorescence is repeatedly photographed over a number of weeks. So the first step to count flowers by image processing is to identify: is there a relationship between visible flowers (only flowers shown on the photo) and actual flowers (the total number of flowers per inflorescence)? Then, when is the optimal time to take a photo and estimate the flower number from a photo?

Hence, a robust flower estimation system is proposed in this paper in order to achieve better flower counting performance and to analyse the impact of variety and inflorescence development. The contributions made in this paper are:

1. A robust algorithm for visible flower detection that utilises texture information from images, with F1 score 0.93 generated from four cultivars (204 inflorescences).
2. A completed flower estimation system, which is tested on 353 images collected *in-vivo* or *ex-vivo* and achieved 84% accuracy in average. Here, we use *in-vivo* to refer to live inflorescences attached to vines in the field and *ex-vivo* to indicate inflorescences which have been detached and imaged under stable lighting conditions and camera setup, such as in an office or laboratory.
3. A detailed analysis of the dependence of flower estimation system on cultivar, stage of inflorescence development and estimation model.
4. The highlighting of the correct metric for assessing the performance of flower estimation by introducing *adjust* – R^2 and the accuracy instead of purely R^2.
5. Confirmation that a global estimation model can be built among cultivars sharing similar inflorescence structures around EL-stage 16 and the best EL stage (Coombe, 1995) for imaging inflorescence is around EL stage 16 regarding flower estimation.

2. Materials and methods

2.1. *Image acquisition and experimental details*

In total, 533 images were captured under various illumination conditions at multiple times at different locations by different people with a black backing board. The development stage of sampled inflorescences ranged between EL stage 12–21 (Coombe, 1995). The collected data covers four cultivars and

(a) Chardonnay (b) Shiraz (c) Cabernet Sauvignon (d) Merlot

Fig. 2 – Morphologies of inflorescences from different cultivars. The difference in flower compactness, the shape of the flowers, and the number of branches. (a) CHA 7, (b) SHI 5, (c) CAS 5, (d) MER, referring to Table 1.

(a) Chardonnay, date 1 (b) Chardonnay, date 2 (c) Chardonnay, date 3

(d) Shiraz, date 1 (e) Shiraz, date 2 (f) Shiraz, date 3

Fig. 3 – Morphology changes of inflorescences over time. The difference in flower compactness, the shape of the flowers and colour. (a) CHA 1, (b) CHA 2, (c) CHA 3, (d) SHI 1, (e) SHI 2, (f), SHI 3, referring to Table 1.

were taken *in-vivo* or *ex-vivo* in Australian vineyards. The camera is either a normal commercial RGB camera or mobile phone. So there is no uniform image size or resolution in the datasets. Figures 2 and 3 show example images to illustrate the problems that the proposed approach is able to solve.

2.2. Algorithms description

The proposed system for automated flower counting is demonstrated in Fig. 4. This system is comprised of three major components:

- Flower candidate detection
- Visible flower classification
- Flower number estimation

Figure 2c is taken as an example to demonstrate each step described in this system. It contains elements which commonly cause failure in existing methods: part of the inflorescence is exposed under the sun; parts of it are blurred due to camera movement or poor focus; and it was sampled *in-vivo*.

2.2.1. Initial segmentation of flower cluster

A bright spot on each flower is the key attribute to detect visible flowers in most previous work (Diago et al., 2014) (the method is meant to be used at the phenological stage EL-16) (Aquino, Millan, Gutiérrez, et al., 2015; Millan et al., 2017), but this spot is not always present or distinct when the flower's surface is matt or not spherical. For instance, the florettes of Chardonnay at an early growing stage (Fig. 5a) have no light reflection in their centres; the florettes of Shiraz have double reflected spots at a more mature stage (Fig. 5b). To provide a more robust solution for visible flower detection, the SURF (Speeded-Up Robust Features (Bay, Ess, Tuytelaars, & Van Gool, 2008)) detector is adopted for flower candidate detection in this paper.

Given a colour image containing a single inflorescence in the red–green–blue (RGB) space,

$$I(u,v) = \{R(u,v), G(u,v), B(u,v)\}, \quad u \in (1,U), v \in (1,V), \tag{1}$$

The threshold value is automatically allocated by Otsu's method (Otsu, 1975) based on the intensity distribution of all pixels in channel $A(u,v)$ of the *LAB* colour space. This step excludes the flowers incorrectly detected in the background and calculates the area of the inflorescence in pixels.

2.2.2. Stem detection

Algorithm 1 shows how the $L(u,v)$ and $A(u,v)$ channels from the LAB colour space are used to estimate the stem area, which is used to mask misidentified flowers.

For the purpose of emphasising boundary features, and therefore avoiding the disturbance from the flower textures, the gradient of $L(u,v)$ is extracted as the initial candidate of stem area, as shown in Fig. 6a. The gradient is then binarised for further morphological investigation by using Otsu's method (Otsu, 1975), as shown in Fig. 6b. By synthesising the binarisation result from the channel $A(u,v)$ and the gradient, the stem area can be separated from the flower area without dramatic morphological degradation.

The connected components in the binary mask I_b, are then labelled and the morphological features of the stem and flower are obtained. Three features: equivalent diameter, eccentricity and area were selected as the morphological criteria for identification of flowers and the stem as they are distinct for these components. An element size e is calculated and

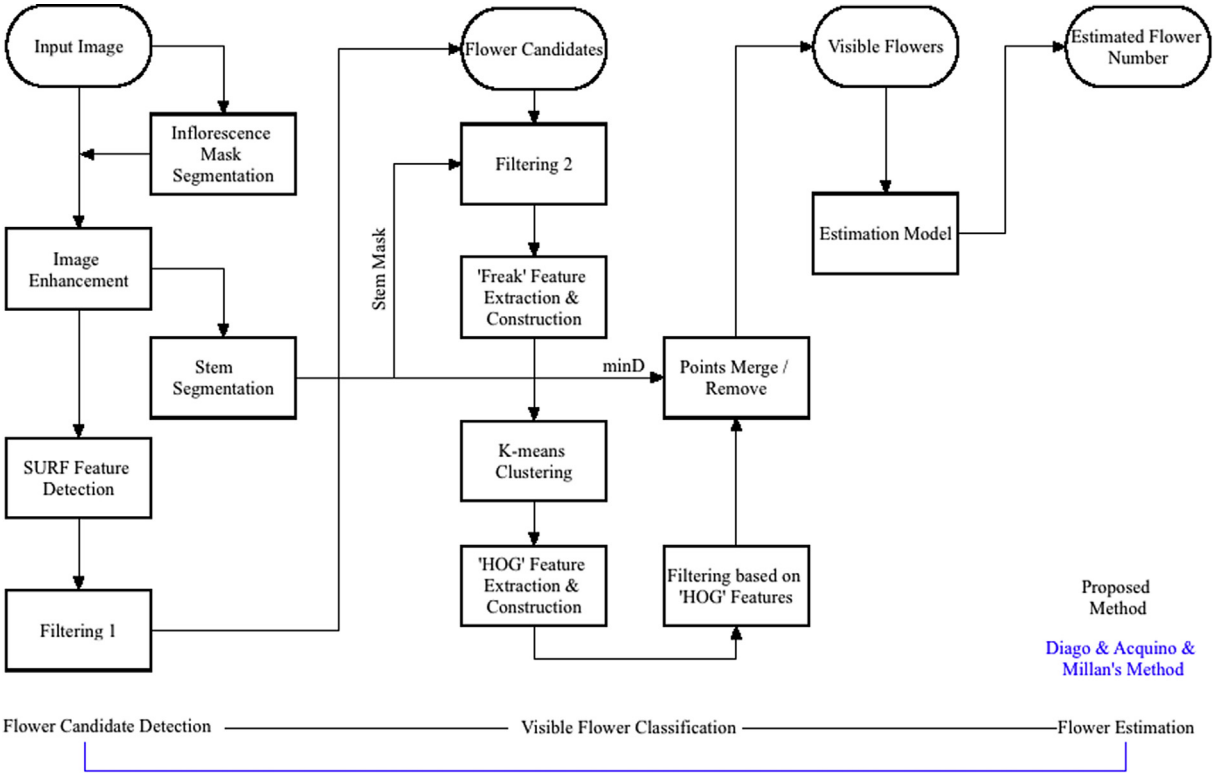
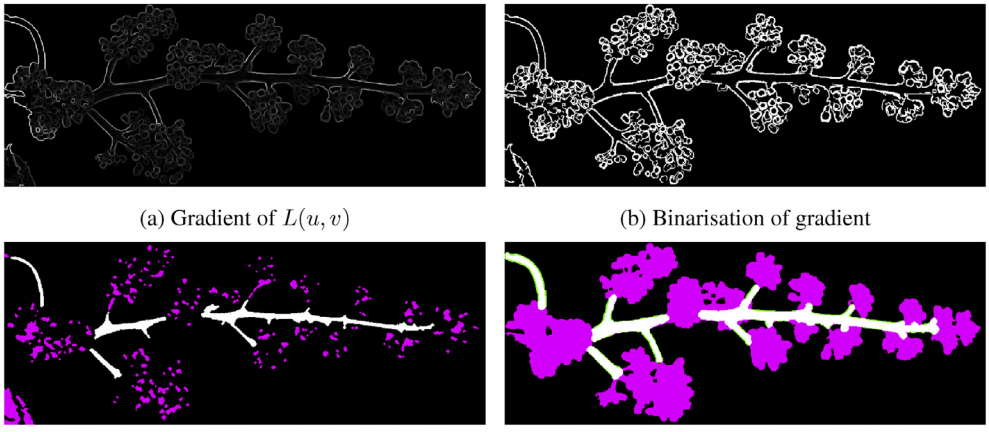


Fig. 4 – The proposed system of automated flower counting.

(a) Chardonnay (b) Shiraz

Fig. 5 – Visible flowers have varying compactness and appearance which stymies existing methods that rely on a peak reflectance on each flower. (a) EL stage 15, matt surface (b) EL Stage 17, multiple peaks in one flower.



(a) Gradient of $L(u,v)$ (b) Binarisation of gradient

(c) Eroded stem (white) and flower (purple) (d) Conditional growing result

Fig. 6 – Stem area detection.

applied to further separate the flower area and stem area. The median diameter of a flower is then estimated and followed by an opening operation to eliminate the noisy pixels. By selecting the regions with the largest aspect ratio and equivalent diameter, the stem area can be extracted, shown as white pixels in Fig. 6c. Finally, the integral form of the stem and flower are recovered via a conditional growing process, as shown in Fig. 6d. The white area marked in Fig. 6d, which is saved as I_{mask} will subsequently be used to remove misidentified flower blobs.

Algorithm 1. Stem detection.

Input: Raw image: $L(u,v)$ and $A(u,v)$
Output: Stem mask: I_{mask}

$I_{mask}(u,v) = 0, u \in (1,U), v \in (1,V)$
$I_{b_1} \leftarrow Otsu's\ Thresholding[A(u,v)]$
$I_{b_2} \leftarrow Otsu's\ Thresholding[gradient[L(u,v)]]$
$I_b \leftarrow I_{b1}\ \&\ \sim I_{b2}$
$R[EquiD, Eccen, Area] \leftarrow$ *Component Labeling* $[I_b]$
$e \leftarrow mean[\ I_R.EquiD], I_R \in I_R(I_R.Eccen > 4^th\ quartile)$
$I_{mask} \leftarrow Erode[I_b, e]$
$I_{mask} \leftarrow Open[I_{mask}, e]$
$R[EquiD, Eccen, Area] \leftarrow$ *Component Labeling*$[I_{mask}]$
$ID_1 \leftarrow R.eccentricity < 0.7$
$ID_2 \leftarrow quartile[R.area, 0.6] < R.area < quartile[R.area, 0.9]$
$ID = ID_1 \& ID_2$
$m_d = mean[R(ID).equivDiameter] + e \times 2$
$I_{mask} = Open[I_{mask}, e(m_d)]$
$I_{mask} \leftarrow components\ with\ target\ morphological\ features$

2.2.3. Flower candidate detection

The SURF detector concentrates on block-like structure in a grey scale image. It detects blobs by convolving the grey image with the determinant of the Hessian (DoH) matrix, as defined in Equation (2):

$$\mathscr{H}(\mathbf{x},\sigma)=\begin{bmatrix} L_{uu}(\mathbf{x},\sigma) & L_{uv}(\mathbf{x},\sigma) \\ L_{uv}(\mathbf{x},\sigma) & L_{vv}(\mathbf{x},\sigma)\end{bmatrix} \tag{2}$$

where

$$\begin{aligned} L_{uu}(\mathbf{x},\sigma) &= I_i(\mathbf{x})\times\frac{\partial^2}{\partial u^2}g(\sigma) \\ L_{uv}(\mathbf{x},\sigma) &= I_i(\mathbf{x})\times\frac{\partial^2}{\partial uv}g(\sigma)\end{aligned} \tag{3}$$

$L_{uu}(\mathbf{x},\sigma)$ in Equation (3) is the convolution of the image with the second derivative of the Gaussian. $I_i(\mathbf{x})$ is an integrated image where each point $\mathbf{x}=(u,v)$ is calculated by Equation (4):

$$I_i(\mathbf{x})=\sum_{i=0}^{i<=u}\sum_{j=0}^{j<=v} I_i(u,v) \tag{4}$$

The flower candidates initially detected by SURF are shown in Fig. 7a. The flower detection result through this stage of image processing can be very imprecise, due to uncontrolled illumination. To increase the count precision, the false-positive results, representing the flowers detected that do not belong to a real flower, are deleted using several filters, before sending the remaining candidates for feature extraction.

- Filter 1, filtering all dark blobs over bright background In some cases, blobs in the dark area between branches/flower are detected as flowers. To remove these false positive flowers, the sign of Laplacian $\nabla^2 L$ is calculated as in Equation (5). All points with negative value are filtered from the flower candidates.

$$\begin{aligned}\nabla^2 L &= tr(\mathscr{H})=L_{xx}(\mathbf{x},\sigma)+L_{yy}(\mathbf{x},\sigma) \\ &= \begin{cases} + & \text{bright blob over dark background} \\ - & \text{dark blob over bright background}\end{cases}\end{aligned} \tag{5}$$

- Filter 2, removing points on the stem by applying the stem mask I_{mask} segmented in Section 2.2.2.
- Filter 3, deleting outliers. For all candidates, the SURF algorithm calculates the scale S of each potential flower. The outliers are calculated using Equation (6); points with scales beyond *MinValue* and *MaxValue* are deleted.

$$\begin{cases} Q1 = \text{the 25th percentile of } S \\ Q3 = \text{the 75th percentile of } S \\ Spread = 1.5*(Q3-Q1) \\ MinValue = Q1 - Spread \\ MaxValue = Q3 + Spread \end{cases} \tag{6}$$

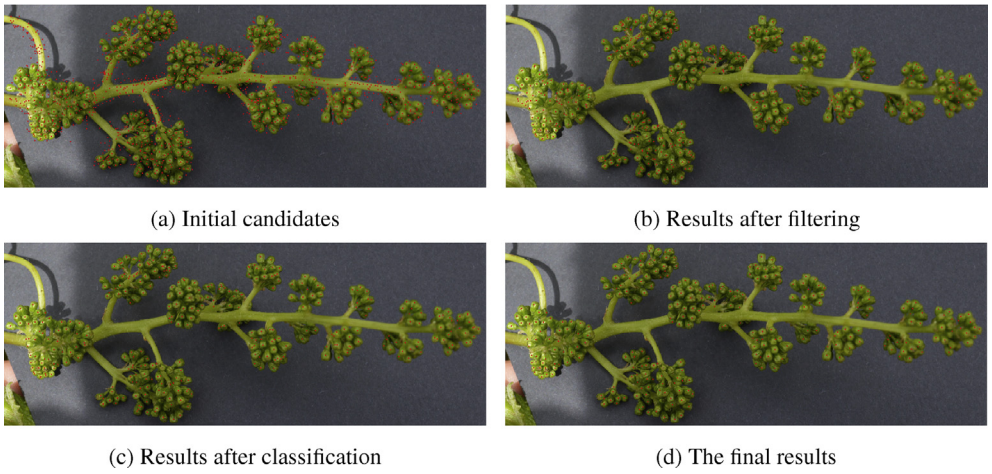
2.2.4. Visible flower classification

Based on the data obtained in flower candidate detection, there are still redundant points remaining on each flower. To distinguish the points with weak performance, the FREAK descriptor (Alahi et al., 2012) is calculated for each candidate. In this case, a feature vector with 64 elements is constructed for every point.

Let $\mathbf{P}=[\mathbf{p}_1,\mathbf{p}_2,\ldots,\mathbf{p}_N]$ denote N points (N flower candidates), $\mathbf{p}_i=F(f_1,f_2,\ldots,f_{64})$ refers to the constructed feature vector with 64 elements for each point. The proposed system adopts an unsupervised learning method, the K-means (MacQueen et al., 1967) clustering algorithm *k*. *N* points are classified into two classes $\mathbf{C}=\{c_1,c_2\}$ by minimising an objective function which we chose to be the Within-Cluster Sum of error Squares (WCSS) as seen in Equation (7).

$$\begin{cases} \mathbf{Y}=k(\mathbf{P},2,d(\mathbf{p},\mu),J), \quad \mathbf{Y}\in\mathbf{C} \\ d(\mathbf{p},\mu)=\sum_{d=1}^{64}\left(\mathbf{p}_d-\mu_{i,d}\right)^2 \\ J=\text{argmin}_{\mathbf{C}}\sum_{i=1}^{2}\sum_{\mathbf{p}\in c_i}\|\mathbf{p}-\mu_i\|^2 \end{cases} \tag{7}$$

where μ_i is the mean of c_i. $\|\mathbf{p}-\mu_i\|^2$ is a chosen distance measured between a data point $\mathbf{p}$ to its cluster centre μ_i. $\mathbf{Y}=[y_1,y_2,\ldots,y_N]$ are the classified results. The smaller group is defined as non-flower candidates after K-means classification.



(a) Initial candidates (b) Results after filtering

(c) Results after classification (d) The final results

Fig. 7 – The demonstration of the proposed visible flower detection. (a) Initial candidates detected by SURF, (b) flower candidates after removing outliers and noise on the stem, (c) further removing noise by classification based on FREAK (Alahi, Ortiz, & Vandergheynst, 2012) and HOG (Dalal & Triggs, 2005) features, (d) the final results by merging candidates with one flower based on the average radius generated from Section 2.2.2.

The next step is extracting HOG (Dalal & Triggs, 2005) features for each candidate. This is for generating the direction of gradient changes around each candidates as real flowers are not homogeneous in direction. A window with four cells is constructed around each candidate. The cell size is determined by the *minD* calculated in Section 2.2.2. In each cell, nine directions are generated with their magnitude from HOG features. The dominant direction for each cell is selected as the one with the greatest magnitude. In short, each candidate has four directions regarding intensity gradient changes in each window. If four directions of one candidate are homogeneous with a confidence of 90%, then this candidate is disqualified. The result after this step is demonstrated in Fig. 7c.

The main reasons that traditional flower estimation methods fail include the ineffective and non-adaptive feature extraction, as well as repetitive counting of an individual flower candidate, with redundant points remaining after the classification. In order to enhance the accuracy of visible flower counting, the next step is to merge all points within a candidate if they are within a spacing of *minD*. Given that *minD* is an average number there are still some candidates remaining in flowers with a larger diameter. The last step is checking the existence of a bridge between two points. If the there is no bridge between two points, then those two are merged into one point. The final results of visible flower detection is illustrated in Fig. 7d.

2.2.5. *Flower number estimation*

Once we have an accurate visible number, a single-variable linear model is adopted for the final flower estimation. The single linear model is utilised in the proposed system because of its satisfactory performance. It presents a more stable and simple solution for estimating the number for one inflorescence based on a single image. The visible flower number is the only input for this model. The coefficients of the linear model, slope and intersect vary between cultivars, and a detailed comparison of different estimation models are presented in Section 3.3.

2.3. *Main estimation factors*

Eight aspects (A1 – A8) of the novel flower estimation system presented in Section 2.2 were analysed and compared with prior work. The aspects were:

- A1, the accuracy of flower detection
- A2, the selection of features for estimation model building
- A3, the selection of estimation model
- A4, the training size to build an estimation model
- A5, the impact of inflorescence development stage.
- A6, the impact of cultivar
- A7, the accuracy of flower estimation with self-calibration
- A8, the accuracy of flower estimation using global models comparing both *in-vivo* and *ex-vivo* images

Hence we designed several experiments to evaluate the performance of the proposed approach. Table 1 illustrates the details of collected datasets and how were they tested for validation.

2.4. *Assessment criteria*

Before starting the experimentation, the criterion to assess the quality of estimation derived by the models needs to be defined, and the coefficient of determination (R^2) and the Root Mean Square Error (*RMSE*) are normally used. However, both values are unable to explicitly represent the accuracy of estimation. A larger data range can lead to bigger R^2, even though the estimation model is the same. To avoid generating a

Table 1 – A reference table of datasets tested in this paper. CHA is short for Chardonnay; SHI is short for Shiraz; CAS is short for Cabernet Sauvignon; MER is short for Merlot; EL is short for EL stage, the phenological stage according to the modified Einhorn-Lorenz system (Coombe, 1995); Ex is short for *ex-vivo*; In is short for *in-vivo*.

Dataset	Images	Date	EL	Location	Ex/In	Purpose
CHA 1	22	14/10/2015	Stage 14	Barossa	Ex	A1, A5 – A7
CHA 2	26	20/10/2015	Stage 16–17			A1, A5 – A8
CHA 3	29	23/10/2015	Stage 19–20			A1 – A7
CHA 4	30	23/10/2015	Stage 17–18			A7
CHA 5	20	12/01/2015	Stage 16	Coonawarra		
CHA 6	30	23/10/2015	Stage 17–18	Barossa	In	A6 – A8
CHA 7	24	15/10/2013	Stage 15	Barossa		
SHI 1	27	20/10/2015	Stage 15	Barossa	Ex	A1, A5 – A7
SHI 2	30	23/10/2015	Stage 17			A1, A5 – A8
SHI 3	25	28/10/2015	Stage 20			A1 – A7
SHI 4	20	12/01/2015	Stage 16	Coonawarra		A7
SHI 5	28	22/10/2013	Stage 16–17	Barossa	In	A6 – A8
CAS 1	25	02/11/2015	Stage 17–18	Barossa	Ex	A1, A7
CAS 2	20	10/01/2015	Stage 16–17	Coonawarra		A6, A7
CAS 3	20	12/01/2015	Stage 17			A7
CAS 4	85	28/10/2014	Stage 15–16	Padthaway		
CAS 5	31	28/10/2013	Stage 16–17	Barossa	In	A6 – A8
MER	20	12/01/2015	Stage 16–17	Coonawarra	Ex	A1, A6 – A8

misleadingly high R^2, and having better assess the estimation models, we adopt the adjusted R-square ($\widehat{R}^2$) (Miles, 2014) to demonstrate how good the fit is between detected flowers and gold standard. The gold standard was comprised of images with each visible flower manually labelled. Percentage error (*PE*) was adopted to assess the performance of the flower estimation system, which is more direct and meaningful from the perspective of implementation in viticulture. The definition of $\widehat{R}^2$ is given by

$$\begin{cases} R^2 = 1 - \dfrac{\sum_i (y_i - \widehat{y}_i)}{\sum_i (y_i - \overline{y}_i)} \\ \widehat{R}^2 = 1 - \dfrac{(1-R^2)(n-1)}{n-k-1} \end{cases} \tag{8}$$

where R^2 is the R-Square value, n is the number of observations in the datasets and k is the number of independent predictors.

The definition of percentage error is given by

$$PE = \sum_{i=1}^{n} \frac{(|y_i - \widehat{y}_i|)}{y_i} \times \frac{1}{n} \times 100 \tag{9}$$

where y_i is the ith actual flower number counted manually (ground truth) and $\widehat{y}_i$ is the ith estimated value. *PE* provides a better estimate of the accuracy (*AC*, $AC = 100 - PE$.) of the estimation model than simply using $\widehat{R}^2$ values.

The most common assessment criteria regarding evaluating the performance of detection is *recall* and *precision* calculated based on the confusion matrix (Fawcett, 2006). In this paper we adopt these criteria as well as their combination, the F1 score, to evaluate the accuracy of the proposed visible flower detection, which is the result after the visible flower classification step.

- *Recall*. The recall metric ζ is evaluating the percentage of actual flowers detected,

$$\zeta = \frac{N_c}{N_c + N_{fn}} \tag{10}$$

where N_c is the true positive value, standing for the number of flowers detected and corresponding to an actual flower in the gold standard set, N_{fn} is the false negative, representing an actual flower, labelled in the gold standard set but not found by the proposed visible flower detection algorithm. Additionally, the denominator of Equation (10) is equal to gold standard value *N*.

- *Precision*. Measures the percentage of flowers correctly targeted, and is derived from

$$\chi = \frac{N_c}{N_c + N_{fp}} \tag{11}$$

where N_{fp} is the false positive, measuring the number of flowers detected that does not correspond to an actual flower in the gold standard set.

- *F1 Score*. Reflects the accuracy of classification by combining *recall* and *precision*.

$$F1 = \frac{2 * \zeta * \chi}{\zeta + \chi} \tag{12}$$

3. Results and discussion

In this section, the experimental results are presented with respect to the eight aspect (A1 – A8) aforementioned in Section 2.3. Each of the images took less than one second to process using Matlab (R2016b, Mathworks, Natick, MA, USA) on a standard desktop machine without any particular optimisation technique.

3.1. A1, the accuracy of visible flower detection

To assess the performance of the proposed visible flower detection, flowers from eight datasets (CHA 1-3, SHI 1-3, CAS 1 and MER, in total 204 photos) were manually selected as points in images on the screen for providing the gold standard. A visual demonstration to show the performance of the proposed visible flower detection is illustrated in the in Fig. 8c and d. The method described by Diago et al. (2014) is effective; however the precision of visible flower number detected is not high, as their results were obtained from very small datasets (15 images for three cultivars). We replicated the method, and tested it on the six datasets (CHA 1-3 and SHI 1-3) used to develop the proposed system. A key parameter in that approach is the threshold of H-maximum *H*, which is not specified in that paper since it was defined manually according to their datasets. A set of values of $H = [30, 40, 50]$ were tested in our experiments based on our datasets and results show that the method presented by Diago et al. (2014) always achieves better results when $H = 40$. Its performance is demonstrated in the second row of Fig. 8. The related metrics generated by our proposed method and the method presented by Diago et al. (2014) are listed in Table 2. The results shown in Fig. 8 and Table 2 clearly indicate the detection results by the proposed method are closer to the gold standard based on $\widehat{R}^2$ and errors and in some cases much better than Diago's results.

By comparing the detection results with the gold standard, the *Recall* and *Precision* metrics were generated and the experimental results are shown in Table 3 below. They demonstrate the proposed visible flower detection has a substantially better performance than related previous work. Given the proposed method has been tested on four cultivars at different development stages under various illumination conditions, the robustness of the proposed visible flower detection provides a solid foundation for flower number estimation.

3.2. A2, comparison of features for model building

Features determine much of the success of a machine learning application, because a model is only as good as its features. The method presented by Millan et al. (2017) employs

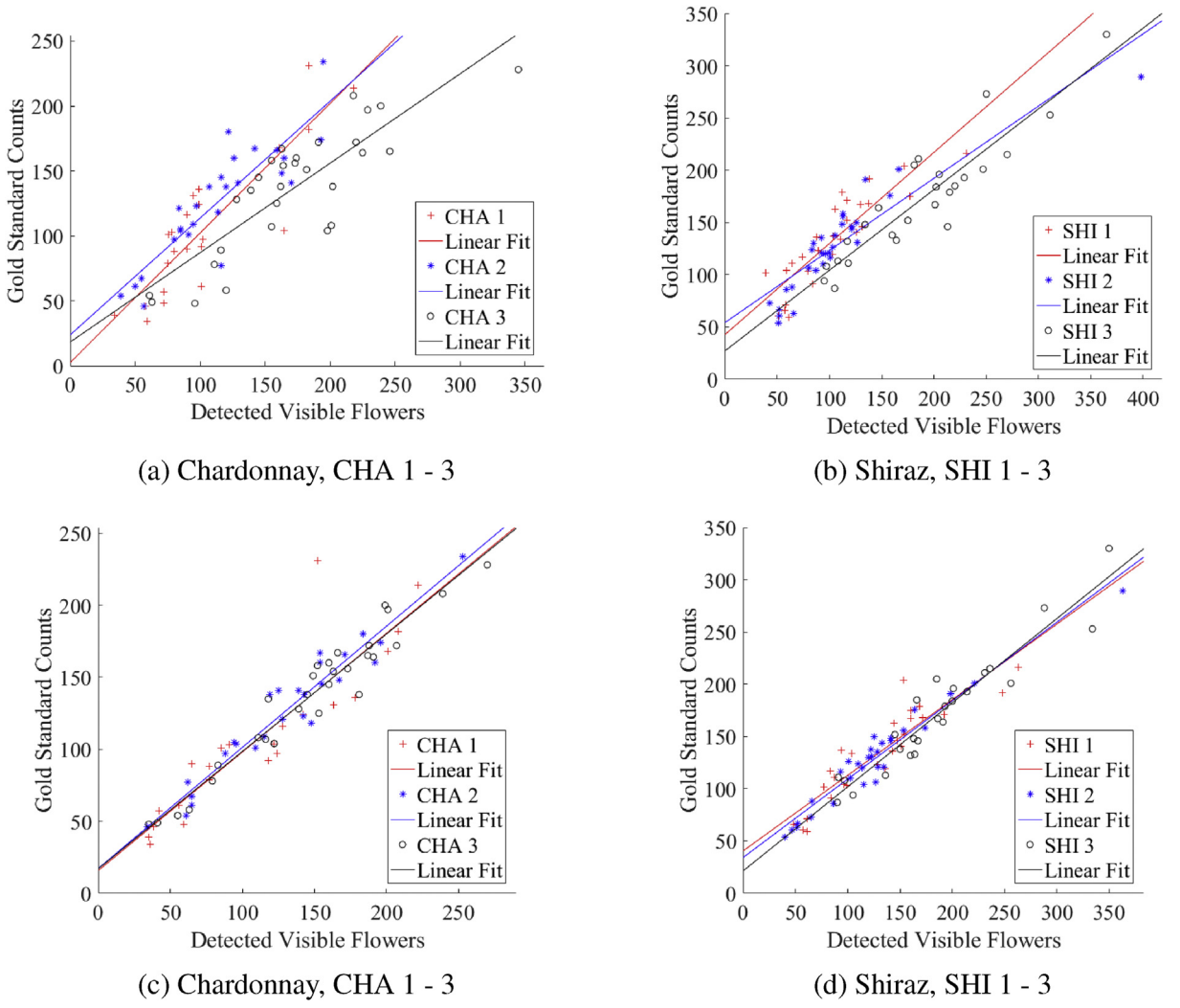


Fig. 8 – Presentation of the performance of the proposed visible flower detection on two cultivars (Chardonnay and Shiraz) in three different development stages. Gold standard versus visible flower count detected by the proposed method. (c) and (d) are tested by the proposed visible flower detection, (a) and (b) are tested by the method proposed by Diago et al. (2014).

Table 2 – A comparison of the performance of visible flower detection in terms of $\widehat{R}^2$, Average Error (*AE*) and Percentage Error (*PE*).

Metrics	Methods	CHA 1	CHA 2	CHA 3	SHI 1	SHI 2	SHI 3
$\widehat{R}^2$	Proposed	0.79	0.93	0.94	0.82	0.94	0.90
	Diago et al. (2014)	0.75	0.75	0.70	0.76	0.81	0.83
AE(%)	Proposed	4.34	1.34	5.53	−5.54	−5.47	11.35
	Diago et al. (2014)	5.47	−8.86	30.89	−22.32	−18.71	8.98
PE(%)	Proposed	24.62	10.12	10.29	13.06	11.95	15.41
	Diago et al. (2014)	23.45	18.04	31.19	23.17	21.54	15.24

five features as model inputs: the number of visible flowers, ROI area, flower radius, flower density and flower area, four of which are directly related to the number of pixels. As the number of pixels has an obvious relationship with image resolution and distance between flower and camera, rather than the relationship between visible flowers and actual flowers per inflorescence, we propose a set of new features: the number of visible flowers, average contrast, smoothness, uniformity and entropy. The last four features originally defined by Gonzalez and Wintz (1977), are less dependent on flower size in the image and aim to more accurately detect the texture of the flowers to improve robustness. Five features were chosen in order to directly compare with Millan's approach. Similarly, a neural network with five input neurons, a hidden layer with 12 hidden neurons and a single output – the total number of flower estimated per inflorescence – was

Table 3 – A comparison of visible flower detection between the proposed method and the results presented by Diago et al. (2014) and Aquino, Millan, Gutiérrez, et al. (2015) in terms of *Recall*, *Precision* and F1 score.

Methods	Cultivar	Images	Recall ζ	Precision χ	F1
Proposed method	CHA 1	22	0.9612	0.9512	0.9555
	CHA 2	26	0.9560	0.9352	0.9446
	CHA 3	29	0.9611	0.9241	0.9417
	SHI 1	27	0.9367	0.9228	0.9284
	SHI 2	30	0.9262	0.8904	0.9069
	SHI 3	25	0.9638	0.8788	0.9188
	CAS 1	25	0.9173	0.8996	0.9069
	MER	20	0.9819	0.8916	0.9344
Diago et al. (2014)	Graciano	5	0.7230	0.9240	0.8112
	Carignan	5	0.7120	0.9570	0.8165
	Tempranillo	5	0.7780	0.9150	0.8410
Aquino, Millan, Gutiérrez, et al. (2015)	Arien	10	0.8793	0.8320	0.8550
	Albarino	10	0.8016	0.8320	0.8165
	Tempranillo	10	0.7516	0.8733	0.7923
	Verdejo	10	0.8817	0.8974	0.8895

constructed and trained. The results are shown in Fig. 9a and b, where the accuracy is compared with the number of training samples.

From the results in the Fig. 9a and b, we can know that for the neural network model, the proposed combination of features can achieve better quality of estimation of flower number than that of Millan for almost all training set sizes. The improvement is due to the features being independent of the number of pixels per inflorescence, which varies with camera resolution and distance to inflorescence.

3.3. A3, comparison of estimation models

Besides examining pixel-independent features for a multi-variable neural network model, we investigated a series of other models for the purpose of estimating the actual number of flowers from the number of visible flowers. Previously, two main estimation models have been used: single-variable model (Diago et al., 2014; Grossetete et al., 2012; Poni et al., 2006), taking only the number of visible flowers as input; and the multi-variable model introduced by Millan et al. (2017) and described in Section 3.2.

Poni et al. (2006) and Diago et al. (2014) suggested that the relationship between visible flower number and actual flower counts could be approximated by a single-variable linear model. Grossetete et al. (2012) built a single-variable polynomial model to calculate the real berry number from the visible number, although this was for berry counts after berry set.

In this paper, we compare four categories of model: single-variable linear model, single-variable nonlinear model, multi-variable linear model and multi-variable nonlinear model.

As expected, the single-variable models used the number of visible flowers as the input, whereas the multi-variable models used the five image features proposed in Section 3.2. The single-variable nonlinear model was built on a second

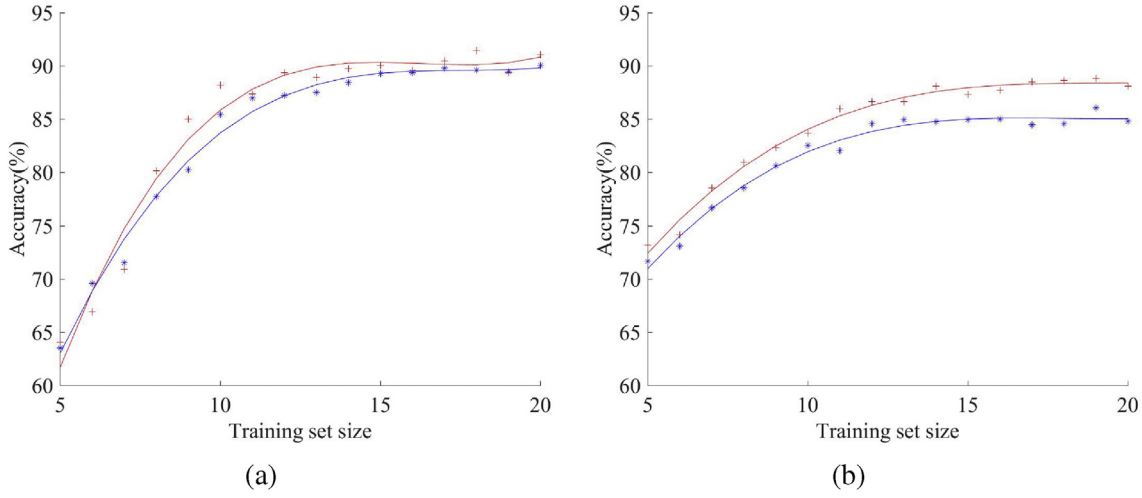


Fig. 9 – A comparison of accuracy between the proposed features (shown with red cross and line) and features proposed in paper Millan et al. (2017) (shown with blue star and line) when they are used to train the multi-variable neural network model for actual flower number estimation. (a) Experiment on Chardonnay dataset CHA 3, (b) Experiment on Shiraz dataset SHI 3. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

degree polynomial. Two multi-variable non-linear models were used, a neural network and Gaussian Process Regression (GPR).

From Fig. 10, it can be seen that single-variable linear model is the most robust estimation model for one Shiraz and one Chardonnay dataset and this was representative of all the datasets tested. For the Chardonnay data, the single-variable linear model is the most accurate model for whole range of tested training sizes, and when the training size is larger than 13, the neural network and single-variable nonlinear models converge to the same accuracy. For the Shiraz data, the single-variable linear model is also the most accurate model for the whole training size. Therefore, single-variable linear models outperform the other tested estimation models when estimating actual flower numbers from relatively small amounts of training data. This confirmed the results from Poni et al. (2006) and Diago et al. (2014) albeit with a larger dataset. In the following experiments, the single-variable linear model was selected to predict flower number.

3.4. A4, selection of training size

To find the optimal training number, the accuracy as a function of training set size was plotted for two representative datasets and it is shown in Fig. 10. We can see that accuracies of the best models like single-variable linear, single-variable nonlinear and the neural network model almost stop increasing when the training set size is over 13. Given the fact that building estimation models needs people to manually count the number of flowers in a destructive manner, we recommend a training set size of 13 inflorescences as a balance between effort and accuracy.

3.5. A5, impact of inflorescence development stage

To verify the impact of inflorescence development stage on estimation models, samples were collected from two cultivars at three different stages of development, ranging from EL14 to EL20. Flowers of CHA 1 and CHA 2 were at a similar development stage as that of SHI 1 and SHI 2 respectively as shown in Table 1. In this experiment, each model was built with 13 samples randomly chosen among one cultivar picked on one date then the model was tested on other datasets of the same cultivar from the same block but at different stages of development.

For Chardonnay (see Table 4), the first dataset was of generally inferior quality as shown by the lower accuracy (75%) when tested on the remainder of the whole dataset. This was due to the flowers being relatively small and in compact groups against the stem of the inflorescence. It also resulted in poor accuracy (45–68%) when tested against datasets containing more developed inflorescences – which were sampled at least 6 days later. However, for the second pair of datasets, taken 3 days apart (light blue shaded cells), and estimated at being EL16-17 and EL19-20, the models could almost be used interchangeably, with a minimum accuracy of 83%.

For Shiraz (See Table 5), with the first two datasets being collected 3 days apart and the second two 5 days apart, with EL stages 15, 17 and 20 respectively, an accuracy of 82% was achieved between adjacent pairs of datasets. When comparing datasets 1 and 3 directly, this dropped to 73 or 78% (light red shaded cells), suggesting that it would be less appropriate to use a model trained on one dataset at EL15 and another at EL20. Given how few days elapsed between these development stages, we conclude that overall flower counting by image processing is dependent on the stage of inflorescence development. However, once they reach EL16 (coinciding with BBCH 55 by Millan et al. (2017)), the dependency reduces to a tolerable level and the same model can be used within a range of about 2 EL stages.

3.6. A6, cultivar dependency

For this experiment, the estimation model was built on one cultivar, then tested on another cultivar. From Tables 6 and 7, the accuracy of each estimation model was reduced due to differences in the cultivars, as discussed below. In particular, from Table 6, training on CHA 1 gave poor results overall, most

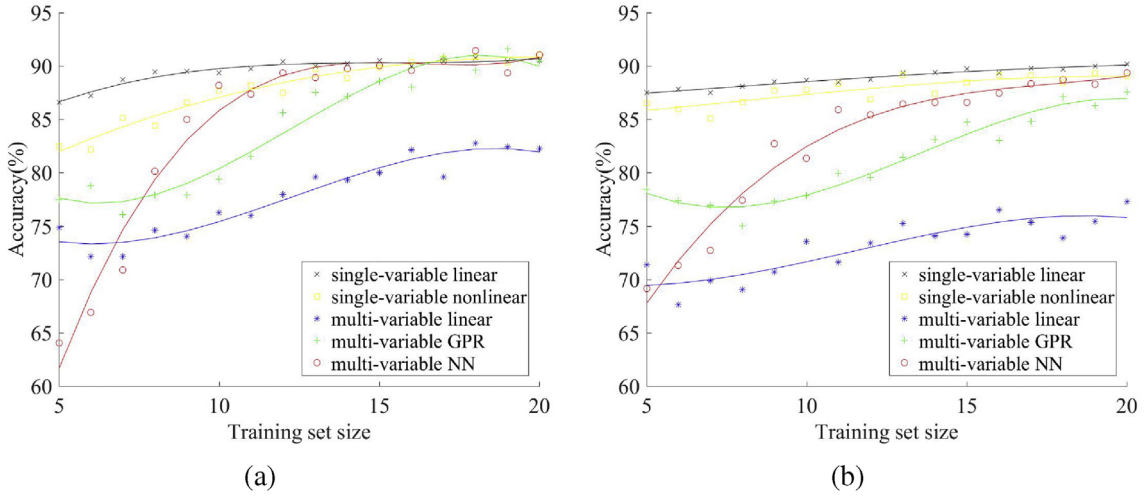


Fig. 10 – Comparison of five types of estimation models and their associated accuracies for different training set sizes as described in Sections 3.3 and 3.4. (a) Experiment with Chardonnay dataset CHA 3, (b) Experiment with Shiraz dataset SHI 3.

Table 4 – Experiments on flowers of different growth stages on Chardonnay. AC is estimation accuracy. Light blue shaded cells show the datasets based on more developed inflorescences have interchangeable models. For EL stages refer to Table 1.

Testing Data	Training Data		
$AC(\%)$	CHA 1	CHA 2	CHA 3
CHA 1	75.39	68.26	57.45
CHA 2	57.31	84.78	83.53
CHA 3	45.96	83.51	90.88

Table 5 – Experiments on flowers of different growth stages on Shiraz, AC is estimation accuracy. Light red shaded cells highlight the reduction in accuracy for datasets further apart in terms of development stage. For EL stages refer to Table 1.

Testing Data	Training Data		
$AC(\%)$	SHI 1	SHI 2	SHI 3
SHI 1	82.84	83.55	78.16
SHI 2	82.06	88.72	86.71
SHI 3	73.31	82.20	87.92

likely due to the very early stage of development of this dataset, which was shown to give poor results in Section 3.5.

Testing models calibrated from Shiraz or Chardonnay on Merlot inflorescences gave poor results overall, due to the noticeable difference in structure of the Merlot inflorescence, with a lower flower density and more spread than the other cultivars. The models built upon more advanced Chardonnay (CHA 3) and Shiraz (SHI 3) performed better with the Merlot despite being at a later EL stage (EL20 versus EL17), suggesting that the estimation model might be better allowing for hidden flowers as the flowers grow in size and obscure each other to some degree.

Table 6 – Experiment with the model built on SHI 1 – SHI 3 and tested on CHA 1 – CHA 3, CAS 2 and MER. AC is estimation accuracy, $AC = 100 - PE$. Highlighted cells are datasets where the inflorescences are at similar stages of development. For EL stages refer to Table 1.

Testing Data	Training Data		
$AC(\%)$	CHA 1	CHA 2	CHA 3
SHI 1	74.41	80.60	72.40
SHI 2	67.69	88.93	82.43
SHI 3	61.83	83.53	87.43
CAS 2	67.81	87.07	84.34
MER	54.45	71.47	81.32

Table 7 – Experiment with the model built on SHI 1-3 and tested on CHA 1-3, CAS 2 and Mer. AC is estimation accuracy, $AC = 100 - PE$. Highlighted cells are datasets where the inflorescences are at similar stages of development. For EL stages refer to Table 1.

Testing Data	Training Data		
$AC(\%)$	SHI 1	SHI 2	SHI 3
CHA 1	78.88	75.16	66.50
CHA 2	72.32	83.60	85.24
CHA 3	64.02	77.60	88.49
CAS 2	80.68	87.15	87.62
MER	64.34	72.30	82.26

Overall, the results suggests that a single estimation model for all varieties is not achievable unless they are of similar inflorescence structure which correlates with the findings by Millan et al. (2017). However in some cases, if we arbitrarily take 80% as a threshold for success, several of the trained models on one cultivar were successful when tested on another cultivar, especially when inflorescences at a similar stage of development were compared – as shown by the highlighted cells in Tables 6 and 7.

3.7. A7, the accuracy of flower estimation

The proposed flower estimation system has been developed based on the datasets of CHA 1-3 and SHI 1-3. To evaluate its performance, the remaining datasets (four cultivars, in total 353 images) were tested for validation. We followed the scheme demonstrated in Fig. 4.

From the analysis of the dependency on development stage (Section 3.5) and cultivar dependency (Section 3.6) we found that the estimation model is more stable when inflorescences reach around EL stage 16. According to Table 8, the estimation model built in the same cultivar and a similar stage of development has high accuracy as expected.

To prove this conclusion, we randomly trained 13 images (Section 3.4) to get an estimation model for each dataset by the single linear model analysed in Section 3.2 and Section 3.3. The experiment was repeated 70 times for each dataset and the average accuracy calculated and presented in Table 8. This experiment achieved an accuracy of 84.30% across 12 datasets (four cultivars with mixed development stages) showing that image based flower counting with self-calibration using 13 images is practical and achievable.

3.8. A8, testing on global models

However, the need to build a new estimation model for every block or cultivar is not desirable. The final experiment completed in this paper investigated the possibility of using a global estimation model. As per Section 3.6, the estimation model was generated from one dataset and this model applied to the other datasets where both datasets were around EL17 and the results are shown in Table 9.

Table 8 – Experimental results of the proposed flower estimation system for individual datasets using a self-calibrated model, as would be used in practice. AC is estimation accuracy calculated based on the metric PE mentioned in Equation (9). AC = 100 – PE.

Datasets	AC (%)	Datasets	AC	Datasets	AC (%)	Datasets	AC (%)
CHA 4	87.87	CHA 7	85.13	CAS 1	85.29	CAS 4	82.73
CHA 5	85.72	SHI 4	84.38	CAS 2	87.05	CAS 5	86.51
CHA 6	80.65	SHI 5	84.12	CAS 3	85.58	MER	82.54

Table 9 – Experimental results of the global estimation model for some *ex-vivo* datasets around EL stage 17. AC is estimation accuracy calculated based on the metric PE mentioned in Equation (9). AC = 1 – PE, the dataset at the top of each column was used for building the model and it was tested on the dataset at the left of the row.

Testing Data	Training Data			
AC(%)	**CHA 2**	**SHI 2**	**CAS 2**	**MER**
CHA 2	84.75	82.96	83.6	82.12
SHI 2	88.64	88.72	91.02	84.23
CAS 2	87.57	87.16	87.05	84.23
MER	70.67	74.66	77.84	82.54

Table 10 – Experimental results of the global estimation model for all *in-vivo* datasets. AC is estimation accuracy calculated based on the metric PE mentioned in Equation (9). AC = 1 – PE, the dataset at the top of each column was used for building the model and it was tested on the dataset at the left of the row.

Testing Data	Training Data			
AC(%)	**CHA 6**	**CHA 7**	**SHI 5**	**CAS 5**
CHA 6	80.65	49.53	56.69	55.71
CHA 7	72.6	85.13	84.47	47.18
SHI 5	79.77	71.46	84.12	68.7
CAS 5	79.19	62.31	72.84	86.51

They indicate that for cultivars of similar development stage and structure, good performance can be obtained if we set a tolerance of 80% for accuracy. All of these datasets contained images taken in office or lab conditions.

We also tested the global estimation models on data collected in the field. The clear message reflected by Table 10 is that for field application, the estimation performance by self-calibrated estimation model is far superior than the model calibrated on other cultivars. The average of self-calibration from *ex-vivo* images (Table 9) was 85.77%, not very different from the average of 84.10% from self-calibration models using *in-vivo* images (Table 10). These high numbers demonstrate that the flower counting system presented in this paper is robust to lighting conditions, cultivar and development stage when self-calibration is used. Equivalently, a global model is not achievable for *in-vivo* images, as shown by the poor off-diagonal results in Table 10.

4. Conclusion

This paper proposes a robust grape vine flower estimation system in order to improve flower counting performance from single images of inflorescences. The system is comprised of an improved flower candidate detection algorithm which includes stem or rachis segmentation; an unsupervised learning method for classifying detected flowers; and a simple linear model for converting classified visible flower counts into estimated counts.

The system was found to be more robust to varying image quality than existing work and gave improved performance in terms of detection and classification of visible flowers at different development stages.

A newly proposed combination of features has been shown to outperform existing flower estimation models for training sizes up to 20 images, with accuracy seen to converge to between 85 and 90%. A single-variable linear model was shown to outperform other estimation models for smaller numbers of training data and 13 training images is recommended as an appropriate balance between effort and accuracy. Following this guideline, the system achieved accuracies of 84.3% under cross-validation. This was achieved across twelve datasets that were separate from the six datasets used to develop the method. This accuracy held for both *ex-vivo* and *in-vivo* datasets, showing the applicability of the proposed flower estimation system for in-field measurement of flower counts.

We conclude that flower counting by image processing is development stage dependent, however once inflorescences reach approximately EL16, this dependency is reduced to a tolerable level and the same model can be used with the same variety within a range of about 2 EL stages.

In terms of a single estimation model for all cultivars, we have shown that this is not feasible unless there is similarity in both structure (density, spread and branching factor) and development stage, which is in agreement with prior work.

Future work will test a wider variety of cultivars. Although 18 datasets containing at least 4 cultivars in 4 growing regions and a total of 533 images with manual counts were used, they are not necessarily representative of the entire spectrum of inflorescence structures. Further analysis of the dependency of the estimation model with inflorescence structure is recommended. An automated method of determining the stage of inflorescence development from an image may also assist in improving the estimation models.

Acknowledgement

Many datasets were tediously collected and flowers counted manually, for which the authors acknowledge the following support. Martin Moran and Victor Sadras from the South Australian Research and Development Institute provided some of the Barossa datasets (CHA 7, SHI 5 and CAS 5). Catherine Wotton from Treasury Wine Estates provided the Padthaway datasets. Catherine Kidman from Treasury Wine Estates provided the Coonawarra dataset. Wine Australia project DPI1401 "Improved Yield Estimation for the Australian Wine Industry" provided funding for some of the authors and for collecting several of the datasets.

REFERENCES

Alahi, A., Ortiz, R., & Vandergheynst, P. (2012). Freak: Fast retina keypoint. In *Computer vision and pattern recognition (CVPR), 2012 IEEE conference on* (pp. 510–517). IEEE.

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.

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.

Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. (2008). Speeded-up robust features (surf). *Computer Vision and Image Understanding, 110*(3), 346–359.

Bennett, J., Jarvis, P., Creasy, G. L., & Trought, M. C. (2005). Influence of defoliation on overwintering carbohydrate reserves, return bloom, and yield of mature chardonnay grapevines. *American Journal of Enology and Viticulture, 56*(4), 386–393.

Coombe, B. (1995). Growth stages of the grapevine: Adoption of a system for identifying grapevine growth stages. *Australian Journal of Grape and Wine Research, 1*(2), 104–110.

Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. In *Computer vision and pattern recognition, 2005. CVPR 2005. IEEE Computer Society Conference on* (Vol. 1, pp. 886–893). IEEE.

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.

Dry, P., Longbottom, M., McLoughlin, S., Johnson, T., & Collins, C. (2010). Classification of reproductive performance of ten winegrape varieties. *Australian Journal of Grape and Wine Research, 16*(s1), 47–55.

Dunn, G. M., & Martin, S. R. (2007). A functional association in vitis vinifera l. cv. cabernet sauvignon between the extent of primary branching and the number of flowers formed per inflorescence. *Australian Journal of Grape and Wine Research, 13*(2), 95–100.

Fawcett, T. (2006). An introduction to roc analysis. *Pattern Recognition Letters, 27*(8), 861–874.

Gonzalez, R., & Wintz, P. (1977). *Digital image processing*.

Grossetete, M., Berthoumieu, Y., Da Costa, J.-P., Germain, C., Lavialle, O., Grenier, G., et al. (2012). Early estimation of vineyard yield: Site specific counting of berries by using a smartphone. In *International conference of agricultural engineering CIGR-AgEng*.

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, volume 1: Statistics* (pp. 281–297). Berkeley, Calif: University of California Press.

May, P. (1987). The grapevine as a perennial, plastic and productive plant. In T. H. Lee (Ed.), *Proceedings of the sixth Australian Wine Industry technical conference* (pp. 40–49).

May, P. (2000). From bud to berry, with special reference to inflorescence and bunch morphology in Vitis vinifera L. *Australian Journal of Grape and Wine Research, 6*(2), 82–98.

Miles, J. (2014). *R squared, adjusted R squared*. John Wiley & Sons, Ltd. URL https://doi.org/10.1002/9781118445112.stat06627.

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.

Otsu, N. (1975). A threshold selection method from gray-level histograms. *Automatica, 11*(285–296), 23–27.

Petrie, P. R., & Clingeleffer, P. R. (2005). Effects of temperature and light (before and after budburst) on inflorescence morphology and flower number of chardonnay grapevines (Vitis vinifera L.). *Australian Journal of Grape and Wine Research, 11*(1), 59–65.

Poni, S., Casalini, L., Bernizzoni, F., Civardi, S., & Intrieri, C. (2006). Effects of early defoliation on shoot photosynthesis, yield components, and grape composition. *American Journal of Enology and Viticulture, 57*(4), 397–407.

Shavrukov, Y. N., Dry, I. B., & Thomas, M. R. (2004). Inflorescence and bunch architecture development in Vitis vinifera L. *Australian Journal of Grape and Wine Research, 10*(2), 116–124.