Automatic tracking of grape clusters and early phenotyping from UAV video sequences

M. Ariza-Sentís, S. Vélez, H. Baja, J. Valente
Information Technology Group, Wageningen University & Research, 6708 PB Wageningen, Netherlands.
*Correspondence: mar.arizasentis@wur.nl

WAGENINGEN UNIVERSITY & RESEARCH

INTRODUCTION

Phenotyping is recognised as a powerful tool in plant breeding. The phenotype of a plant includes morphological and physiological characteristics, and it is a consequence of the interaction between genotype and environment. As in other crops, identifying phenotyping traits in viticulture is essential due to their impact on the plant adaptation to the environment and the quality of the grapes and wine. These phenotyping traits are related to several characteristics, such as **leaf characteristics, plant phenology or canopy architecture**. Some of them are related to cluster dimensions and are very important in grapevine breeding because they directly affect yield, as well as grape and wine quality. These traits have been studied for a long time in viticulture through ampelography. **Ampelography** is the science that studies the identification and classification of grapevines (*Vitis*). The description of these **ampelographic descriptors** is necessary to identify *Vitis vinifera* varieties and species. Therefore, most of these traits are recognised by the International Organisation of Vine (OIV) and Wine) and the International Plant Genetic Resources Institute (IPGRI) as descriptors for grape varieties and *Vitis* species. Bunch length and width are **OIV codes 202 and 203**, respectively.

Traditionally, this information is achieved manually through visual surveys, which are time-consuming, expensive, subjective and laborious. Therefore, UAVs (Unmanned Aerial Vehicles) could be a promising tool to obtain these phenotyping traits **faster, cheaper, and more objectively**. Computer vision based on deep learning can assist in this task. Object detection is a computer vision technique used to identify and locate the object in an image or video. Another functionality is object tracking, which consists of following the movement of an object through the video frames, allowing counting the number of objects in a video since it assigns an ID (unique identifier) to each object, while storing important information of each object (for instance, to count the number of visible grape clusters in a vineyard row). This work aims to **identify and track visible grapevine clusters** and **estimate their dimensions** (OIV codes 202 and 203) in an early development stage using **UAV RGB video sequences** and **Deep Learning**. For this purpose, the PointTrack algorithm has been implemented on 29 videos as it allows multi-object tracking and segmentation.

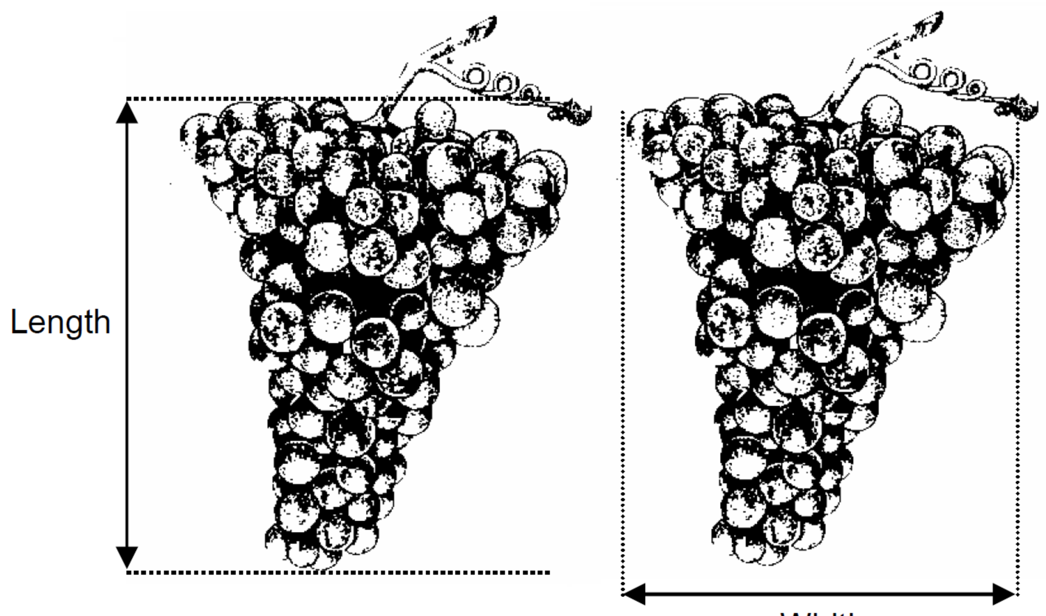


Fig. 1. Cluster length and width, according to OIV codes 202 and 203.

Fig. 2. Snapshot of the grape row (top) and grape cluster detection with PointTrack (bottom).

MATERIALS AND METHODS

Cultivar. *Vitis vinifera* L. cv Loureiro.

Dataset. 29 videos split into a total of 679 images were annotated using the software CVAT. 432 for training and 247 for testing (70% for training and 30% for testing).

UAS. DJI Matrice 210 and Zenmuse X5S.

Software. CVAT (annotation), Anaconda v.2020.11, Python v.3.9.X. **Framework.** PointTrack algorithm. **Network Architecture.** ERFNet. **CNN Library.** PyTorch. **Instance Segmentation Architecture.** Spatial Embedding (Proposal-free clustering loss function). **Layers.** Dense-prediction encoder-decoder layers (ConvNet). **Model configuration.** 400 Iterations over the training data, additional 800 iterations over the training data for fine-tuning, 100 iterations for instance association.

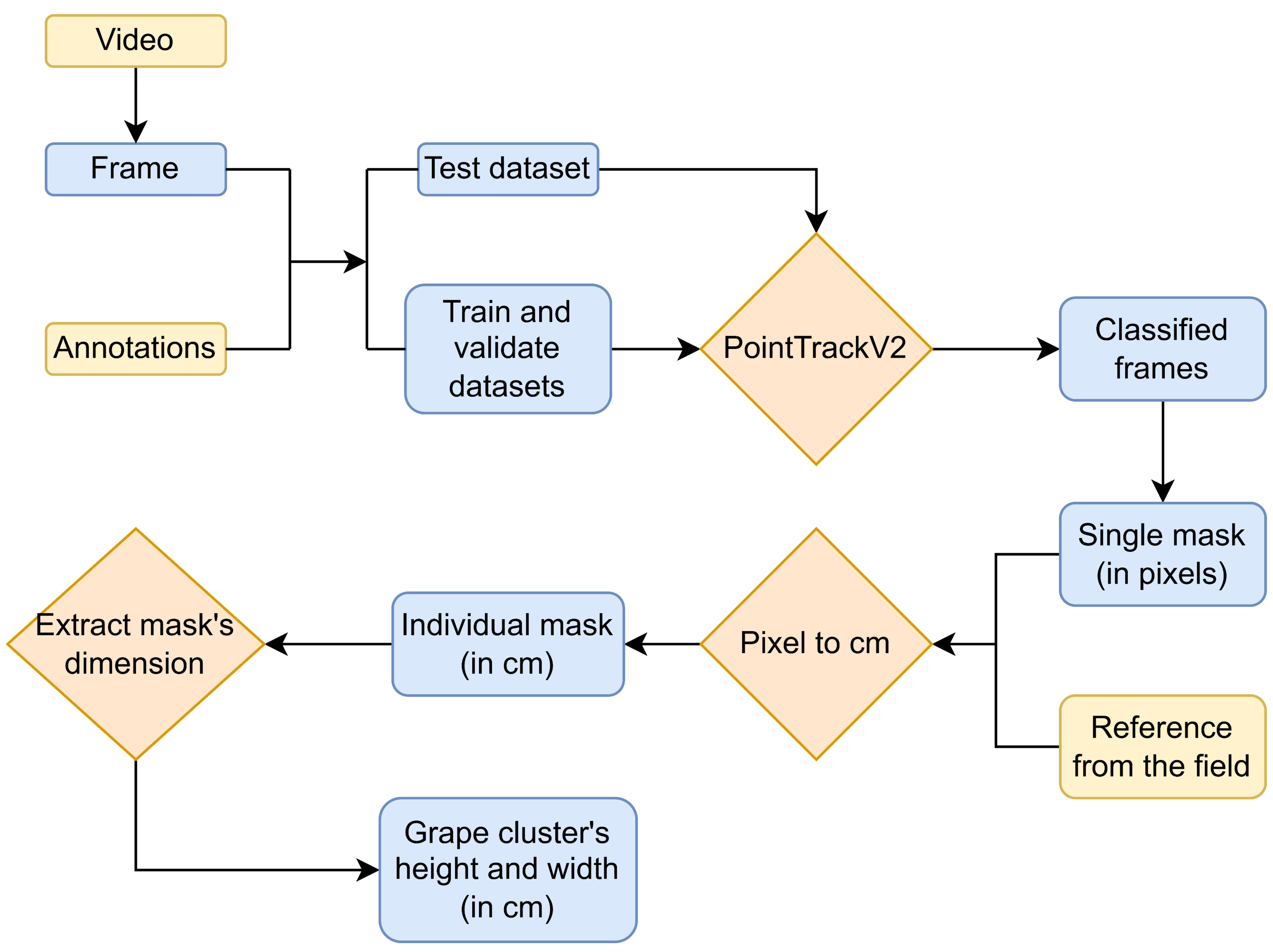


Fig. 3. Workflow of the study. The videos are split into frames and combined with the annotations are used to train, validate, and test the algorithm. The outputs are the classified frames, employed to extract single masks and derive width and height of the grape clusters.

RESULTS

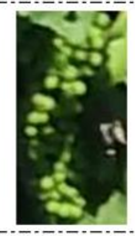
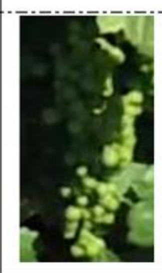
ID	Length (pixels)	Width (pixels)	Length (cm) OIV 202	Width (cm) OIV 203	Mask	Cluster
66	115	59	16.4	8.4		
2	122	65	15.7	8.4		
1	76	38	11.4	5.7		
4	144	77	21.6	11.6		

Fig. 4. OIV codes 202 and 203 extracted from the PointTrack algorithm, together with the grape cluster mask.

CONCLUSIONS

- ✓ PointTrack could learn grape clusters in the imagery by identify spatial features of patterns between the foreground and the background by means of convolutional layers.
- ✓ PointTrack extracted significant features from the input imagery, generating a successful model to identify clusters.
- ✓ The algorithm could detect shaded clusters.
- ✓ The methodology was able to extract cluster length and width, which are key descriptors according to OIV codes 202 and 203, respectively.

ACKNOWLEDGMENTS

This work has been carried out in the scope of the H2020 FlexiGroBots project, which has been funded by the European Commission in the scope of its H2020 programme (contract number 101017111). The authors acknowledge valuable help and contributions from all partners of the project.

FLEXIGROBOTS

7th International Plant Phenotyping Symposium
27.09.2022 - 30.09.2022 – Wageningen, The Netherlands

IPPS 2022 PLANT PHENOTYPING FOR A SUSTAINABLE FUTURE