



Data augmentation from RGB to chlorophyll fluorescence imaging Application to leaf segmentation of *Arabidopsis thaliana* from top view images

Natalia Sapoukhina, Salma Samiei, Pejman Rasti, David Rousseau

► To cite this version:

Natalia Sapoukhina, Salma Samiei, Pejman Rasti, David Rousseau. Data augmentation from RGB to chlorophyll fluorescence imaging Application to leaf segmentation of *Arabidopsis thaliana* from top view images. Conference on Computer Vision and Pattern Recognition (CVPR), Jun 2019, Long Beach, United States. pp.4321-4328, 10.1109/CVPRW.2019.00312 . hal-02307721

HAL Id: hal-02307721

<https://hal.science/hal-02307721>

Submitted on 13 Oct 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Data augmentation from RGB to chlorophyll fluorescence imaging Application to leaf segmentation of *Arabidopsis thaliana* from top view images

Natalia Sapoukhina

INRA, UMR1345 Institut de Recherche en Horticulture et Semences IRHS,
 SFR 4207, PRES UNAM,
 49071 Beaucouzé, France

natalia.sapoukhina@inra.fr

Salma Samiei

salma.samiei@univ-angers.fr

Pejman Rasti

spejman.rasti@univ-angers.fr

David Rousseau

david.rousseau@univ-angers.fr

Laboratoire Angevin de Recherche en Ingénierie des Système (LARIS),
 UMR INRA IRHS, Université d'Angers,
 62 avenue Notre Dame du Lac, 49000 Angers, France

Abstract

*In this report we investigate various strategies to boost the performance for leaf segmentation of *Arabidopsis thaliana* in chlorophyll fluorescent imaging without any manual annotation. Direct conversion of RGB images to gray levels picked from CVPPP challenge or from a virtual *Arabidopsis thaliana* simulator are tested together with synthetic noisy versions of these. Segmentation performed with a state of the art U-Net convolutional neural network is shown to benefit from these approaches with a Dice coefficient between 0.95 and 0.97 on the segmentation of the border of the leaves. A new annotated dataset of fluorescent images is made available.*

1. Introduction

Due to heavy occlusion, variability in terms of size and shape, leaf segmentation is a challenging task from the computer vision perspective [16]. One strategy to simplify the segmentation is to reduce the biological variability and focus on a limited amount of plant species of specific interest. This has been undertaken in the CVPPP challenge since 2014 with a focus on a few species including *Arabidopsis thaliana* which serves as a reference for a number of fundamental biological questions. The effort to provide finely annotated data [14] has enabled great improvement of the state

of the art on segmentation performance. An open question is now how to transfer this knowledge obtained from RGB images on annotated plants either to other species or other modalities of imaging. In this work, we focus on the transfer of the knowledge gained from annotated leaves of *Arabidopsis thaliana* in RGB to images of the same plant in chlorophyll fluorescence imaging.

2. Related Work

Segmentation of *Arabidopsis thaliana* leaves in RGB images has been highly studied since the introduction of the CVPPP challenge. If in 2014 and 2015 the contributions of this challenge proposed methods based on models [20, 27, 21], most of the participants have so far mainly tackled the challenge with deep neural network [29, 26, 31]. In this work we did not propose any innovation on this side and rather work on a standard neural network architecture but applied it for the first time on another imaging modality. We used the U-Net architecture [23] which had been mainly employed for the pixel-wise segmentation of separation boundaries in medical [34] and satellite images [13]. Here, we applied U-Net for the first time to the best of our knowledge on leaf segmentation of *Arabidopsis thaliana* in chlorophyll fluorescence imaging.

Chlorophyll fluorescence analysis is a non-destructive technique which has been developed to probe plant physiology [6]. Among all the chlorophyll fluorescence param-

eters that can be estimated, the maximum quantum yield of photosystem II (PSII) photochemistry ($F_v/F_m = (F_m - F_0)/F_m$) [9] is an indicator of plant stress [22]. Fluorescence chlorophyll by image analysis on whole plant has been widely studied [24, 4, 17]. So far, to the best of our knowledge analysis on individual leaves has not been tackled in top view images of *Arabidopsis thaliana*.

Image simulation to boost machine learning received an increasing interest in plant imaging [35, 10, 1, 8]. This can include standard data augmentation, sophisticated infography or generative models from convolutional networks. In this communication we generated the images from one imaging modality to learn on another imaging modality. This topic has been demonstrated possible for instance for life science applications in the medical domain [12] in a cross modal image synthesis and also in microscopy in a superresolution problem [19]. We considered for the first time data augmentation from the synthesis of images from RGB imaging modality to chlorophyll fluorescence imaging in plant sciences.

3. Method

3.1. Datasets

Three datasets coined *CVPPP*, *CSIRO* and *Real-Fluo* are considered in this study. They are described in the following lines.

CVPPP. We used the dataset provided in the Leaf Segmentation Challenge held as part of the computer vision problems in plant phenotyping *CVPPP* workshop [14]. *CVPPP* dataset consists in 27 RGB images of tobacco plants and 783 RGB images of *Arabidopsis* wild and mutant plants. We considered only the *Arabidopsis* dataset in this study. All images are hand labelled to obtain ground truth masks for each leaf in the scene (as described in [14]). These masks are image files encoded in PNG where each segmented leaf is identified with a unique integer value, starting from 1, where 0 is background.

CSIRO. To extend the *CVPPP* dataset we also used synthetic images of top down view renders of *Arabidopsis* generated with the simulator described in [30, 32]. The *CSIRO* dataset contains 10000 synthetic images (width x height: 550 x 550 pixels). Similarly to *CVPPP* dataset, each RGB image has a corresponding leaf instance segmentation annotation: each leaf in an image is uniquely identified by a single color value, starting from 1, where 0 is background. All images are stored in PNG format.

Real-Fluo. For model testing we used 38 real gray-scale fluorescent images of *Arabidopsis*. The PSI Open FluorCam FC 800-O (PSI, Brno, Czech Republic) was used to capture chlorophyll fluorescence images and to estimate the maximum quantum yield of PSII (F_v/F_m) on wild type control of *Arabidopsis thaliana*. The system sensor is a

CCD camera with a pixel resolution of 512 by 512 and a 12-bit dynamic range. The system includes 4 LED panels divided into 2 pairs. One pair provides an orange actinic light with a wavelength of around 618 nm, with an intensity that can vary from 200 to 400 $\mu\text{mol}/\text{m}^2/\text{s}$. It provides a 2s pulse that allows the measurement of the initial fluorescent state (F_0). The other pair provides a saturating pulse during 1s in blue wavelength, typically 455 nm, with an intensity of up to 3000 $\mu\text{mol}/\text{m}^2/\text{s}$. The saturating pulse allows collecting of the maximum fluorescence (F_m). Fluorescence chlorophyll imaging was used in a dark adapted mode after a dark period of 45 min to produce maps with the fluorescent quantum efficiency $F_v/F_m = (F_m - F_0)/F_m$. All these 38 images were manually annotated using Phenotiki image analysis software [15] and are made available to the reader (see the web link at the end of the article).

3.2. U-Net Model

The segmentation of the leaves was considered to be a pixel-wise classification where the pixel of the leaf contour should be detected among the other pixels of the image. Picking out leaf contours allowed separating leaves and thereby performing leaf segmentation, for example with help of a watershed transform. Each pixel was therefore classified among three mutually exclusive classes: mask without contours, leaf contours and background. It means that every pixel was labeled by a three-component one-hot vector.

The U-Net model [23] was used for the pixel-wise classification. As shown in Figure 1 U-Net architecture is separated in 3 parts: the contracting/downsampling path, the bottleneck, the expanding/upsampling path. The encoder-decoder type architecture with skipped connections allows combining low-level feature maps with higher-level ones, and enables precise pixel classification. A large number of feature channels in upsampling part allows propagating context information to higher resolution layers. The output of the model was a three-channel label that indicated the class of every pixel as shown in Figure 2. All activation functions in the convolutional layers were rectified linear units, ReLU [11]. The last layer before the prediction was a softmax activation with 3 classes. Images and labels from all datasets were resized to width x height: 128 x 128 pixels. Using ground truth (GT) labels, we created three-channel labels as shown in Figure 2. To reinforce the learning of the contour class, which was highly unbalanced, we replaced the encoder by a ResNet152 backbone pre-trained on ImageNet [33]. The decoder was not changed from the original description [23]. We empirically found that the best performances were obtained when all skipped connections were kept which was in accordance with the intrinsic multiscale nature of plants [25]. The resulting U-Net neural network had a total 1,942,275 trainable parameters.

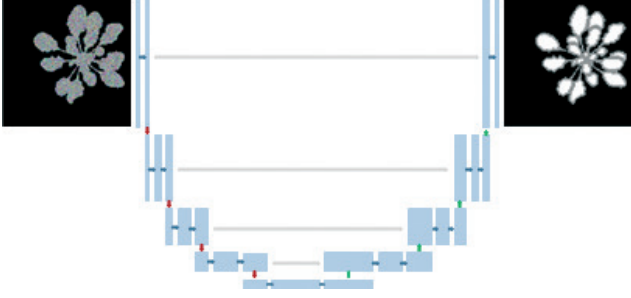


Figure 1: U-Net architecture. Each blue box corresponds to a multi-channel feature map. Gray arrows indicate the merging of the context and localization information that was done by concatenating the features from the contracting path with the corresponding ones in the expansion path. Input image has 128x128 pixels, the output of the model is a three-channel binary image: mask without contours, leaf contours and background.

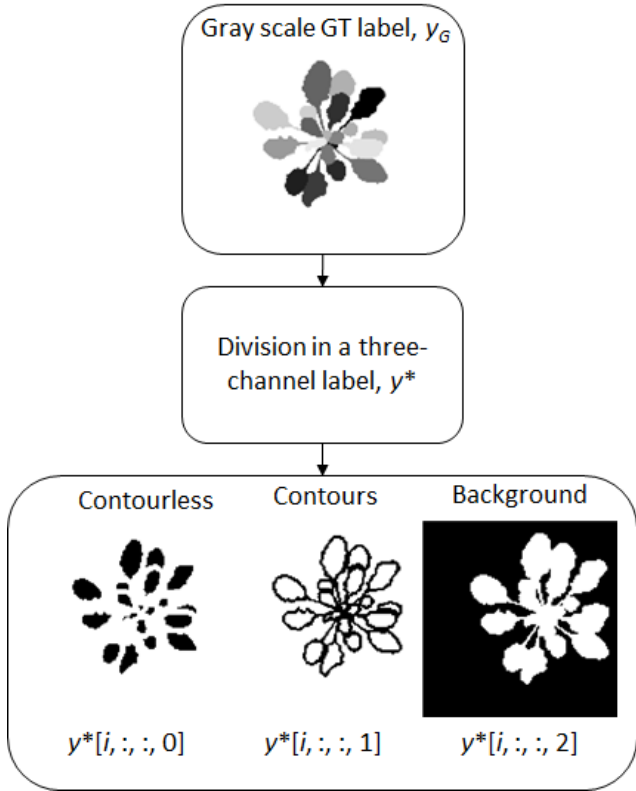


Figure 2: Production of the three-channel binary labels from ground truth (GT) labels: the first channel contains mask without leaf contours, the second channel - leaf contours and the third one - background.

3.3. Data augmentation

Several strategies of data augmentation were investigated from *CVPPP* and *CSIRO* datasets to train our U-Net in order to perform leaf segmentation on the *Real-Fluo* dataset.

In a first simplest strategy, we converted *CVPPP* and *CSIRO* directly from RGB to gray levels along the simple CIE formulae $Gray = .299 * Red + .587 * Green + .114 * Blue$. In a second strategy, we considered binary images such as the ones in Figure 3 column (b) and mapped on them a noisy texture learned from the real fluorescence images, *Real-Fluo*, shown in Figure 3 column (a). A copy of the original binary image for each plant was also kept so as to produce the associated GT. For a first trial of transfer from RGB images to fluorescence images, we propose to test an extremely simple model for the noisy texture which is estimated as an additive Gaussian white noise process independent and identically distributed for a given leaf. This choice was first driven by an Occam's Razor simplicity spirit. Indeed with such a model the simulated leaves have no spatial structures such as vascular veins. Leaves are therefore expected to be distinguished in real images only from their first order statistics. Also, as an additional motivation to test this simple fluorescence chlorophyll simulator, the noise in real fluorescence images is expected to be mostly thermal noise on the camera which will control the standard deviation of the noise. The leaves themselves, if considered to have homogeneous tissue, may have a variety of average values in fluorescence emission depending on their physiological state.

To estimate the parameters of these Gaussian processes, we analyzed the distribution of the gray levels among a small set of images of real plants. In order to ensure that this small set of chlorophyll fluorescence images was representative from the rest of the images we considered one image of plant at each developmental stage represented in the test dataset. Average value and standard deviation of the gray levels inside the plants for both considered chlorophyll fluorescence parameters F_0 and F_m are given in Table 1. The order of magnitude of the average value and standard deviation of the chlorophyll fluorescence parameters F_0 and F_m remained in the same range in our experiment.

Synthetic chlorophyll fluorescent images were then simply produced by adding Gaussian noises with mean μ and variance σ^2 for each fluorescence map $(\mu_{F_0}, \sigma_{F_0}^2)$, and $(\mu_{F_m}, \sigma_{F_m}^2)$, randomly sampled in Table 1. A different realization of these noises was applied for each individual leaf of gray scale GT labels in *CVPPP* and *CSIRO* datasets so as to produce a synthetic fluorescent example x_F given by

$$x_F = 1 - \frac{\sum_l (y_g^{(l)} + \mathcal{N}(\mu_{F_0}, \sigma_{F_0}^2))}{\sum_l (y_g^{(l)} + \mathcal{N}(\mu_{F_m}, \sigma_{F_m}^2))}, \quad (1)$$

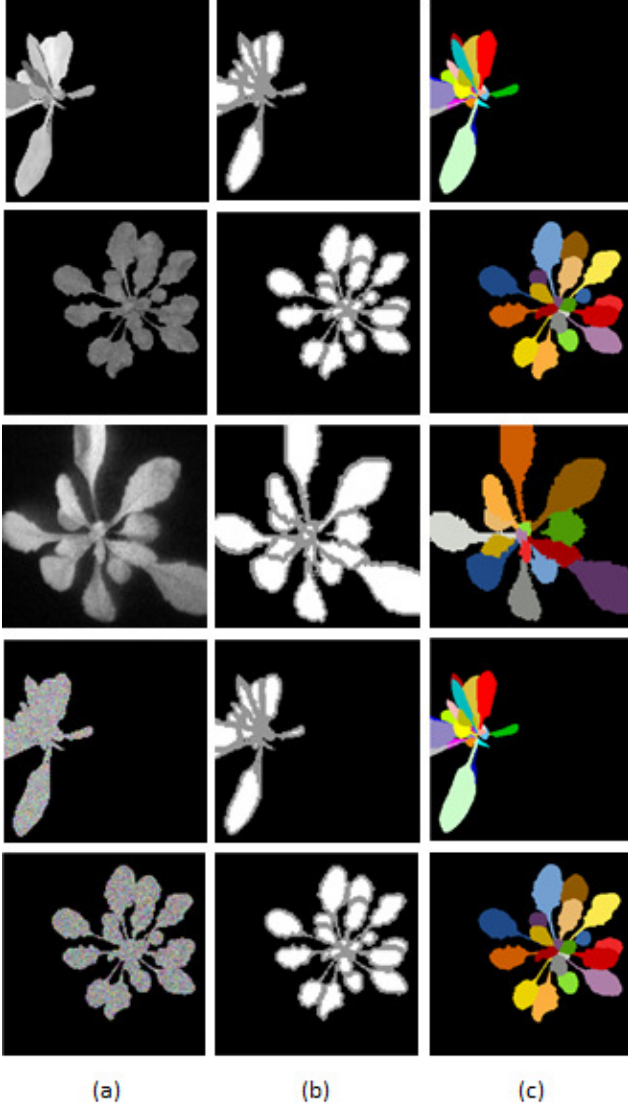


Figure 3: Examples from datasets used for model training and its evaluation. (a) Plant image examples. (b) three-channel labels for pixel-wise classification. (c) Ground truth labels with leaf segmentation. First line: *CSIRO* dataset, 783 examples. Second line: *CVPPP* dataset, 783 examples. Third line: *Real-Fluo* dataset, 38 examples. Forth line: *CSIRO-Fluo* dataset, 5481 examples. Fifth line: *CVPPP-Fluo* dataset, 5481 examples. Number of examples in datasets are given before application of the standard data augmentation.

where $y_g^{(l)}$ is l th binary leaf from a gray scale GT label and $\mathcal{N}(\mu_{F_0}, \sigma_{F_0}^2)$ is a Gaussian noise realization. For every GT label we produced 7 synthetic fluorescent examples x_F by drawing random values for μ_{F_0}, σ_{F_0} and μ_{F_m}, σ_{F_m} from Table 1. The pipeline of data augmentation is shown

in Figure 4.

As a result, in addition to *CVPPP* and *CSIRO*, we obtained new datasets, *CVPPP-Fluo* and *CSIRO-Fluo*, containing $5481 = 783 * 7$ and $70000 = 10000 * 7$ synthetic fluorescent images (width x height: 128×128 pixels), respectively. Now, our objective is to compare the added value of all these datasets for leaf segmentation in *Real-Fluo* dataset with the U-Net model presented in previous section.

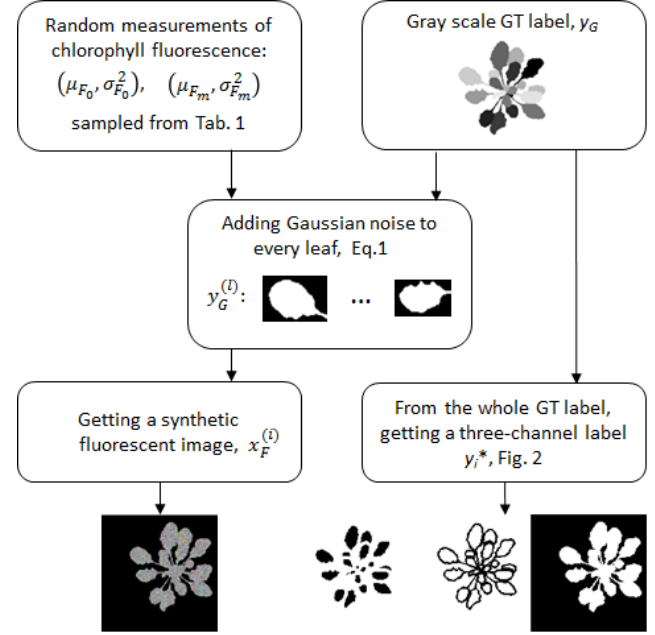


Figure 4: Data augmentation using synthetic fluorescent training data. For each gray-scale GT label from *CVPPP* or *CSIRO* datasets we produced fluorescent images and associated three-channel labels.

Time	μ_{F_0}	σ_{F_0}	μ_{F_m}	σ_{F_m}
Day 1	167.83	34.88	180.77	24.68
Day 5	165.81	33.1	180.00	22.36
Day 6	164.48	30.87	177.9	20.8
Day 7	158.16	31.45	174.73	21.1
Day 8	165.24	32.31	181.14	21.36
Day 9	168.3	28.03	184.36	17.86
Day 12	173.06	28.01	189.96	17.15

Table 1: Mean, μ , and standard deviation, σ for chlorophyll fluorescence F_0, F_m estimated on a single plant from *Real-Fluo* dataset at different dates after emergence of first leaves (cotyledons).

3.4. Watershed Post-Processing

To segment leaves with use of estimated three-channel labels, we applied the classical marker-controlled watershed segmentation [3, 2]. The markers were generated with a contourless mask from output three-channel label and then, to segment leaves, we flooded marked “basins” within the bounds of mask.

4. Experiment and Results

4.1. Training

On top of the data augmentation techniques that we generated from *CVPPP* and *CSIRO* datasets as described in section 3.3, we apply a standard data augmentation strategy in order to further reduce overfitting and improve generalization. For this data augmentation we used Albumentations library [5]. While the data augmentation strategies of section 3.3 focused on contrast and noise distribution, here we generated geometrical transformation such as horizontal flip, vertical flip, random rotate at 90 degree and random half-sized crop and applied them to shuffled training dataset.

It was shown that for high level of imbalance, loss functions based on overlap measures appeared to be more robust [28]. Through all of our experiments, we minimized weighted combination of multi-class cross entropy and dice losses

$$L(y, y^*) = w_0 C(y, y^*) + w_1 (1 - D(y[\dots, 0], y^*[\dots, 0])) + w_2 (1 - D(y[\dots, 1], y^*[\dots, 1])). \quad (2)$$

$C(y, y^*)$ is the categorical cross entropy defined as

$$C(y, y^*) = - \sum_{ij} y_{ij} \log y_{ij}^* \quad (3)$$

and $D(y, y^*)$ is the Dice coefficient

$$D(y, y^*) = \frac{2 \sum_{ij} y_{ij} y_{ij}^* + \epsilon}{\sum_{ij} y_{ij} + \sum_{ij} y_{ij}^* + \epsilon}, \quad (4)$$

where y is a model prediction with values y_{ij} , y^* is a ground truth label with values y_{ij}^* and $\epsilon = 0.001$ is used here to ensure the coefficient stability by avoiding the numerical issue of dividing by 0. The weight ratios (w_0, w_1, w_2) used to correct the class imbalance was respectively 0.4, 0.1, 0.5 for cross entropy, contourless mask and contours. Adam optimizer was used with default parameters $\eta = 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$. Our training procedure consisted of splitting the data into 80% and 20% training and cross validation respectively. We shuffled the dataset examples at the beginning of each epoch and used a batch size of 16 examples. We also implemented batch normalization before each activation.

Leaf segmentation in fluorescence images was done with different data augmentation strategies for the training based on the datasets of Figure 3 and their combinations. A base line consisted in training directly on the *CVPPP* or *CSIRO* RGB to gray images. The learning from the simulated fluorescence dataset either generated along Eq. (1) from *CVPPP* labels and/or *CSIRO* labels was tested for comparison. The previous strategies were tested also when small amount of real fluorescence images were added in the training. The eight different tested training strategies are summarized in Table 2.

4.2. Results

To assess the quality of segmentation, we used the soft Dice coefficient, Eq. (4), that was computed separately for all pixels and for leaf contours. Furthermore, the pixel-wise accuracy was evaluated in order to get a general idea of the model performance. It was computed as the ratio between correctly classified pixels and the total number of pixels in the test sample. To assess the performance of leaf contour detection we computed additional metrics. True positives (TP) are contour pixels present in both prediction and GT mask. False positives (FP) are contour pixels present in prediction but absent in GT mask. False negatives (FN) are contour pixels absent in prediction but present in GT mask. Knowing these numbers we can estimate true positive rate

$$TPR = \frac{TP}{TP + FN}, \quad (5)$$

that describes the fraction of correctly classified contour pixels in comparison of the total number of contour pixels in GT mask. Moreover, positive predictive value

$$PPV = \frac{TP}{TP + FP}, \quad (6)$$

gives us the fraction of correctly classified contour pixels among all predicted contour pixels.

Table 2 displays the model performance on the *Real-Fluo* dataset for eight model training experiments. A first global observation is that the performance of training on *CVPPP* alone was rather high. This demonstrates a high similarity of RGB reflectance images converted to gray levels and the fluorescence images despite the physical differences in the mechanism of their production. Training on *CVPPP-Fluo* and *CSIRO-Fluo* alone or combined did not provide better performances than *CVPPP* alone. The best model Dice score was 97% obtained for extended *CVPPP* and *CVPPP-Fluo* datasets with 10 examples from *Real Fluo* dataset. The use of small quantity of real fluorescent images among images with modeled fluorescence resulted in Dice score gain of 2-3% in comparison with *CVPPP* and *CVPPP-Fluo* datasets. The same positive effect of the injection of 10 real fluorescent images on the model performance

Training dataset	Accuracy	L_{train}	D_{train}	L_{test}	D_{test}	D_c	TPR	PPV
<i>CVPPP</i>	0.96	0.03	0.98	0.19	0.95	0.67	0.74	0.62
<i>CVPPP-Fluo</i>	0.96	0.01	0.99	0.22	0.94	0.68	0.82	0.58
<i>CSIRO-Fluo</i>	0.94	0.02	0.99	0.27	0.92	0.46	0.48	0.46
<i>CVPPP-Fluo + CSIRO Fluo</i>	0.95	0.01	0.99	0.3	0.93	0.62	0.63	0.62
<i>CVPPP + 10ex</i>	0.98	0.04	0.98	0.05	0.97	0.84	0.8	0.87
<i>CVPPP-Fluo + 10ex</i>	0.98	0.02	0.99	0.05	0.97	0.84	0.83	0.85
<i>CSIRO-Fluo + 10ex</i>	0.98	0.03	0.98	0.06	0.96	0.84	0.85	0.82
<i>CVPPP-Fluo + CSIRO Fluo + 10ex</i>	0.98	0.04	0.97	0.06	0.96	0.8	0.8	0.82

Table 2: Performance metrics on *Real-Fluo* samples for training strategies with the various strategies tested of data augmentation. To estimate the overall model performance, we used pixel-wise Accuracy. To assess overfitting, we calculated loss function, Eq. (2), and Dice coefficient, Eq. (4), for the training dataset, L_{train} , D_{train} , and for the test dataset, L_{test} , D_{test} , respectively. The accuracy of contour pixels detection was evaluated by means of Dice coefficient, D_c , true positive rate, TPR , and positive predictive value, PPV . See Eq. (5, 6) for the last two estimates.

was observed for the other datasets as well, *CSIRO-Fluo + 10ex* and *CVPPP-Fluo + CSIRO Fluo + 10ex*. It increased overall pixel-wise accuracy to 0.98 and the quality of contour detection became quite high as well: $D_c \in [0.8, 0.84]$, $TPR \in [0.8, 0.83]$, $PPV \in [0.82, 0.87]$. In more details, the comparison of model performance from training on *CVPPP* and *CVPPP-Fluo* showed that the imitation of fluorescence by modelling increased TRP by 9% and decreased PPV by 6%. It means that the modelled fluorescence allowed us to detect a little bit better leaf contours on real fluorescent plant images but at the same time it had the tendency to classify surplus pixels as contour pixels. Training on *CSIRO-Fluo* had the lowest values of metrics in comparison with the other training strategies. However these can be considered as interesting results if one keeps in mind that in this case the network was trained only on purely synthetic datasets. Probably, there is a deficiency of important information of leaf texture in synthetic plants from *CSIRO-Fluo* that prevents the simulation of fluorescence in a sufficient realistic way.

As shown in Figure 5, with the worst example from the best training strategy, most errors of pixel classification occurred for occluded leaves, i.e. for really difficult cases. Another source of discrepancies was an inaccurate annotation of some contour pixels. It means that some pixels were correctly classified as contour but since they were present in GT label with displacement they were not counted in true positive rate. However, this type of errors did not prevent the correct segmentation as it is shown in the upper line of Figure 6. Only two cases of occlusive leaves were not segmented. These kind of discrepancies could potentially be solved using a more advanced post-processing method. Overall, the segmentation performance was higher for young small plants where there was not a lot of leaf occlusion as it is shown in the lower line of Figure 6.

5. Conclusion and Discussion

In this paper, we studied the transfer of knowledge for leaf segmentation learned from RGB imaging to fluorescence imaging. Various data augmentation strategies were tested with real images of plants or on pure synthetic plants and from RGB to gray conversion up to a physical modelling of noise in fluorescence.

This was illustrated on *Arabidopsis thaliana* which is one of the most studied plant for fundamental biology and with the U-Net neural network architecture applied for the first time in this context. We have demonstrated that existing annotated datasets in RGB could be used to learn to segment leaves in fluorescence images by a simple RGB to gray conversion. Also, good performances (although not the best) of segmentation could be obtained by learning on purely synthetic datasets automatically annotated and mapped with a first order statistics physical modelling of noise in fluorescence. Segmentation performance were found higher when some real images were also introduced in the training process.

These results could be extended in various promising directions. First, one could try to improve the segmentation result presented here. Other neural network architectures could for instance be tested such as the one recently introduced to consider segmentation as a regression [18]. Also performances on training from simulated datasets could benefit from domain adaptation [7] to compensate for the necessarily non perfect match between simulation and reality. Other plant imaging modalities could finally be also investigated in the same way as in this communication. One could for instance think to thermal imaging or Tera hertz imaging which are also used to assess the physiological state of leaves. There are currently no annotated datasets for these images and it would therefore be interesting to explore if data augmentation from other imaging in which annotated

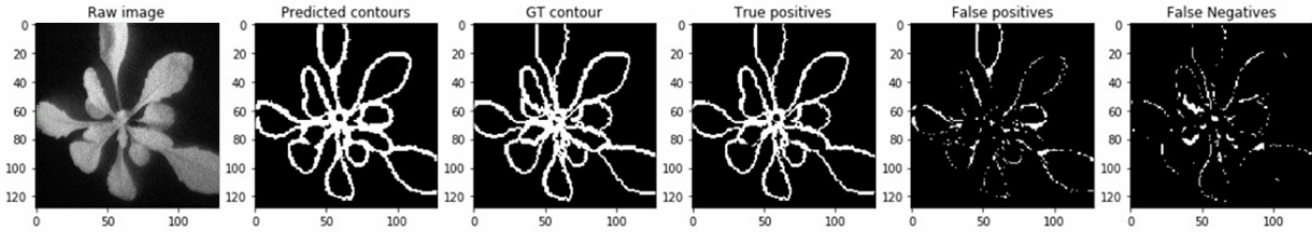


Figure 5: Example of leaf contours detection in a *Arabidopsis* fluorescent image. In this case the model was trained on *CVPPP-Fluo + 10ex* dataset including 5481 images with imitated fluorescence and 10 real fluorescent images. True positives show well predicted contour pixels existing in GT label. False positives show surplus pixels that were classified as the contour but did not exist in GT label. False negatives are pixels that had to be classified as contour pixels since they were presented in GT label.

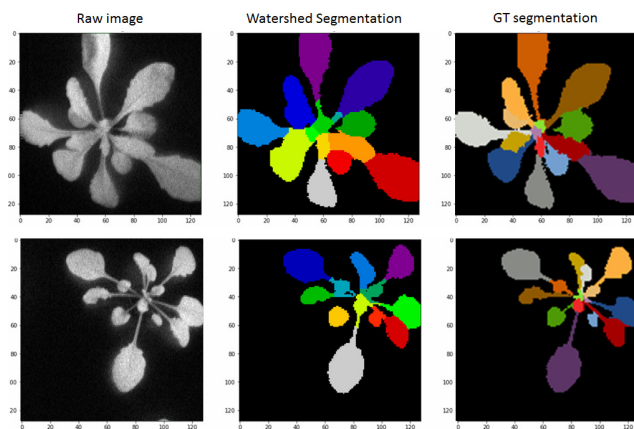


Figure 6: Examples of watershed segmentation produced by the model trained on *CVPPP-Fluo + 10ex*. Upper line: $D_c = 0.8$, $Accuracy = 0.96$, $TPR = 0.82$, $PPV = 0.78$. Lower line: $D_c = 0.84$, $Accuracy = 0.98$, $TPR = 0.83$, $PPV = 0.86$.

datasets are available could be helpful.

To contribute to reproducible science, we have opened, as an additional result from our study, access to our annotated dataset of *Arabidopsis thaliana* in fluorescence imaging (<https://uabox.univ-angers.fr/index.php/s/BglUZgoE5EWK4MM>).

6. Acknowledgements

Authors thank Etienne Belin and Tristan Boureau from Platform PHENOTIC part of PHENOME phenotyping network for the production of *Arabidopsis* and associated fluorescence images. Salma Samiei gratefully acknowledges Angers Loir Metropole for the funding of her PhD Grant.

References

- [1] R. Barth, J. IJsselmuiden, J. Hemming, and E. J. Van Henten. Data synthesis methods for semantic segmentation in agriculture: A capsicum annum dataset. *Computers and Electronics in Agriculture*, 144:284–296, 2018.
- [2] S. Beucher. The watershed transformation applied to image segmentation. *Scanning Microscopy International*, 6(1):299–314, 1991.
- [3] S. Beucher and C. Lantuejoul. Use of watersheds in contour detection. In *Proc. Int. Workshop Image Processing, Real-Time edge and Motion Detection/Estimation*, 1976.
- [4] J. Bresson, F. Vasseur, M. Dauzat, G. Koch, C. Granier, and D. Vile. Quantifying spatial heterogeneity of chlorophyll fluorescence during plant growth and in response to water stress. *Plant Methods*, 11(1):23, 2015.
- [5] A. Buslaev, A. Parinov, E. Khvedchenya, V. I. Iglovikov, and A. A. Kalinin. Albumentations: fast and flexible image augmentations. *ArXiv e-prints*, 2018.
- [6] W. L. Butler. Energy distribution in the photochemical apparatus of photosynthesis. *Annual Review of Plant Physiology*, 29(1):345–378, 1978.
- [7] N. Courty, R. Flamary, D. Tuia, and A. Rakotomamonjy. Optimal transport for domain adaptation. *IEEE transactions on pattern analysis and machine intelligence*, 39(9):1853–1865, 2017.
- [8] M. Di Cicco, C. Potena, G. Grisetti, and A. Pretto. Automatic model based dataset generation for fast and accurate crop and weeds detection. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5188–5195. IEEE, 2017.
- [9] B. Genty and S. Meyer. Quantitative mapping of leaf photosynthesis using chlorophyll fluorescence imaging. *Functional Plant Biology*, 22(2):277–284, 1995.
- [10] M. V. Giuffrida, H. Scharr, and S. A. Tsafaris. Arigan: Synthetic arabidopsis plants using generative adversarial network. In *Proceedings of the 2017 IEEE International Conference on Computer Vision Workshop (ICCVW), Venice, Italy*, pages 22–29, 2017.
- [11] K. He, X. Zhang, S. Ren, and J. Sun. Delving deep into rectifiers: surpassing human-level performance on imagenet classification. In *Proc. IEEE*, 2015.

- [12] Y. Hiasa, Y. Otake, M. Takao, T. Matsuoka, K. Takashima, A. Carass, J. L. Prince, N. Sugano, and Y. Sato. Cross-modality image synthesis from unpaired data using cycle-gan. In *International Workshop on Simulation and Synthesis in Medical Imaging*, pages 31–41. Springer, 2018.
- [13] V. Iglovikov, S. Seferbekov, A. Buslaev, and A. Shvets. Terausnetv2: Fully convolutional network for instance segmentation. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2018.
- [14] M. Minervini, A. Fischbach, H. Scharr, and S. A. Tsaftaris. Finely-grained annotated datasets for image-based plant phenotyping. pages 1–10, 2015.
- [15] M. Minervini, M. V. Giuffrida, and S. A. Tsaftaris. An interactive tool for semi-automated leaf annotation. *Proceedings of the Computer Vision Problems in Plant Phenotyping (CVPPP) Workshop*.
- [16] M. Minervini, H. Scharr, and S. A. Tsaftaris. Image analysis: the new bottleneck in plant phenotyping [applications corner]. volume 32, pages 126–131. IEEE, 2015.
- [17] A. M. Mutka, S. J. Fentress, J. W. Sher, J. C. Berry, C. Pretz, D. A. Nusinow, and R. Bart. Quantitative, image-based phenotyping methods provide insight into spatial and temporal dimensions of plant disease. *Plant physiology*, 172(2):650–660, 2016.
- [18] P. Naylor, M. Laé, F. Reyat, and T. Walter. Segmentation of nuclei in histopathology images by deep regression of the distance map. *IEEE transactions on medical imaging*, 38(2):448–459, 2019.
- [19] C. Ounkomol, S. Seshamani, M. M. Maleckar, F. Collman, and G. R. Johnson. Label-free prediction of three-dimensional fluorescence images from transmitted-light microscopy. *Nature methods*, 15(11):917, 2018.
- [20] J.-M. Pape and C. Klukas. 3-d histogram-based segmentation and leaf detection for rosette plants. In *European Conference on Computer Vision*, pages 61–74. Springer, 2014.
- [21] J.-M. Pape and C. Klukas. Utilizing machine learning approaches to improve the prediction of leaf counts and individual leaf segmentation of rosette plant images. *Proceedings of the Computer Vision Problems in Plant Phenotyping (CVPPP)*, pages 1–12, 2015.
- [22] S. A. Rolfe and J. D. Scholes. Chlorophyll fluorescence imaging of plant–pathogen interactions. *Protoplasma*, 247(3-4):163–175, 2010.
- [23] O. Ronneberger, P. Fischer, and T. Brox. U-net: convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241, 2015.
- [24] C. Rousseau, E. Belin, E. Bove, D. Rousseau, F. Fabre, R. Berruyer, J. Guillaumès, C. Manceau, M.-A. Jacques, and T. Boureau. High throughput quantitative phenotyping of plant resistance using chlorophyll fluorescence image analysis. *Plant methods*, 9(1):17, 2013.
- [25] D. Rousseau, Y. Chéné, E. Belin, G. Semaan, G. Trigui, K. Boudehri, F. Franconi, and F. Chapeau-Blondeau. Multi-scale imaging of plants: current approaches and challenges. *Plant methods*, 11(1):6, 2015.
- [26] H. Scharr, T. P. Pridmore, and S. A. Tsaftaris. Computer vision problems in plant phenotyping, cvppp 2017–introduction to the cvppp 2017 workshop papers. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2020–2021, 2017.
- [27] K. Simek and K. Barnard. Gaussian process shape models for bayesian segmentation of plant leaves. *Proceedings of the Computer Vision Problems in Plant Phenotyping (CVPPP)*, pages, pages 4–1, 2015.
- [28] H. C. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. J. Cardoso. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. *ArXiv e-prints*, 2017.
- [29] S. A. Tsaftaris, M. Minervini, and H. Scharr. Machine learning for plant phenotyping needs image processing. *Trends in plant science*, 21(12):989–991, 2016.
- [30] D. Ward and P. Moghadam. Synthetic arabidopsis dataset. *CSIRO. Data Collection.*, 2018.
- [31] D. Ward, P. Moghadam, and N. Hudson. Deep leaf segmentation using synthetic data. *arXiv preprint arXiv:1807.10931*, 2018.
- [32] D. Ward, P. Moghadam, and N. Hudson. Deep leaf segmentation using synthetic data. In *Proceedings of the British Machine Vision Conference (BMVC) Workshop on Computer Vision Problems in Plant Phenotyping (CVPPP)*, 2018.
- [33] P. Yakubovskiy. Segmentation models. https://github.com/qubvel/segmentation_models, 2019.
- [34] A. Zaimi, M. Wabartha, P. L. A. V. Herman, C. S. Perone, and J. Cohen-Adad. Axondeepseg: automatic axon and myelin segmentation from microscopy data using convolutional neural networks. *Scientific Reports*, 8(1):38216, 2018.
- [35] Y. Zhu, M. Aoun, M. Krijn, J. Vanschoren, and H. T. Campus. Data augmentation using conditional generative adversarial networks for leaf counting in arabidopsis plants. *Computer Vision Problems in Plant Phenotyping (CVPPP2018)*, 2018.