

Data augmentation for pathogen segmentation in vinewood fluorescence microscopy images

Julie Munsch^{1,2}, Sonia Ouali¹, Jean-Baptiste Courbot¹, Romain Pierron³

and Olivier Haeberlé¹

¹IRIMAS, UR 7499, Université de Haute-Alsace, Mulhouse, France

²Eiffage Energie Systèmes - Centre Expertise IA, Mulhouse, France

³LVBE, UR 3991, Université de Haute-Alsace, Colmar, France

Abstract. In this paper, we address the problem of segmentation of vine pathogens within fluorescence microscopy images. To our knowledge, the quantification from such images is an original problem. As a consequence, there is no available database to rely upon in order to use supervised machine learning techniques. We provide a workaround by creating realistic images containing the desired filamentary pattern and variable blur effect. Numerical results show the interest of this data augmentation technique, especially on images corresponding to difficult segmentation.

Keywords: data augmentation, image segmentation, fluorescence microscopy deep learning, machine learning

1 Introduction

1.1 Context

Grapevine trunk diseases have become a serious problem for vinegrowers around the world. In France, approximately 13% of the vineyard is unproductive each year, leading to losses of 1 billion euros [1]. Esca is one of these diseases and the oldest described. It affects 4.4% of vines causing deterioration or even complete dieback of the wood. The behavior of pathogens believed to be responsible for this disease is until today poorly known, and no treatment exists once the wood is contaminated. Understanding the colonization process is key to the

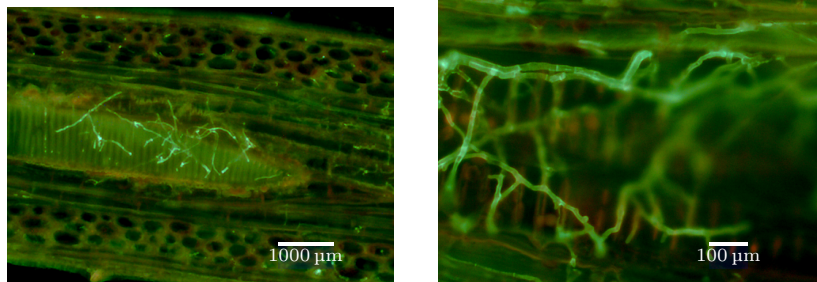


Fig. 1: Examples of pathogen in vinewood images observed with a fluorescence microscope, at x10 magnification on the left and x40 on the right. The pathogen fluorescence appears as green-yellow filaments, while wood auto-fluorescence appears as yellow-brown.

development of sustainable cures. To gain a better understanding of the pathogens' behavior, inoculation experiments are performed and the resulting plant is observed in fluorescence microscopy (see Fig. 1). Then, image segmentation techniques can help quantify the pathogen's presence in vinewood.

Obtaining real microscopic grapevine images, together with expert-segmented images (here by Dr. R. Pierron), is however a time-consuming process. The resulting images form a relatively limited database that might be too small to use robust supervised learning techniques. Thus, data augmentation is a key step to fully make use of these methods.

1.2 Related works

Data augmentation techniques are, in general, a helpful tool to enlarge databases. This is in particular the case for medical images, for which the patient set is often limited [2].

When handling images, basic augmentation relies on some simple transforms, such as rotation, flip, cropping, and dilation. These techniques might be improved using advanced deep learning approaches, such as adversarial training or neural style transfer (see [3] for a review). Those are mostly based on image textures, and do not incorporate a model for the image formation process. We can make a similar observation regarding segmentation in microscopy images, as highlighted in [4].

Hence, there is a lack for a data augmentation technique that specifically accounts for the image formation model. This is particularly striking when handling the case of fungi in vineyard images, which can be affected by a varying blur depending on the region of the image.

In this article, we propose to generate synthetic images that mimic real fluorescent microscopy grapevine images in order to train supervised algorithms. This article is organized as follows. First, we describe the image formation process of both the real fluorescence microscopy images and the synthetic ones (Section 2). Then, in Section 3 we present the segmentation methods and the results obtained on the synthetic and real images.

2 Image formation

2.1 Real images

Colored grapevine cuttings are observed using a wide-field fluorescent microscope. The resulting images are characterized by the filamentary light green fluorescent pathogen (see Fig. 1). However, the wood also exhibits auto-fluorescence, which makes it challenging to distinguish between the pathogen and the wood. The latter is a complex plant tissue, which results in heavily textured background images.

Furthermore, the images acquired using a fluorescence microscope suffer from a blurring effect. The blur is due to the convolution of the sample with the system's Point Spread Function (PSF). The latter is known at an instrumental level, so the blur is partially known. The remaining unknown is the sample's depth, which influence on the width of the PSF to take into account.

Given an un-blurred image \mathbf{I} containing S pixels, corresponding to an ideal sample, we formalize this model as

$$(\mathbf{H}(\mathbf{v})\mathbf{I})_s = \langle \mathbf{h}_{s,[v_s]}^{3D}, \mathbf{I} \rangle, \quad (1)$$

where $\mathbf{h}^{3D} \in \mathbb{R}^{S \times S \times P}$ is known only on a 3D grid, see Fig. 2. P represents the depth of the PSF, $\mathbf{h}_{s,[v_s]}^{3D} \in \mathbb{R}^S$ represents the PSF of the system at depth v_s in the pixel s , and \mathbf{v} is a depth map associated to the image. In other words, \mathbf{H} forms a depth-selection operator to model the blur. Thus, we consider that an image $\mathbf{y} \in \mathbb{R}^S$ is formed as:

$$\mathbf{y} = \mathbf{H}(\mathbf{v})\mathbf{I} + \mathbf{b}, \quad (2)$$

with \mathbf{b} an additional background term. Note that we consider \mathbf{I} the ideal image containing only the pathogen, including the colors. Later on, we note its binary counterpart \mathbf{x} to model the segmentation mask.

2.2 Synthetic images

In this section, we present how we mimic the real images to augment the available database. As part of the grapevine inoculation process, some plants were part of a control group, *i.e.* there were no inoculated fungi. These plants were also observed with fluorescence microscopy. The resulting images capture well what is considered as a background in our segmentation framework (\mathbf{b} in Eq. (2)).

Besides, there is a need to generate synthetic pathogens in the image, that have a filamentary texture and are also corrupted by a varying blur. Filamentary texture is well represented in retina images, which have been investigated for many years (see [6] for an overview). Several databases are publicly available, and we selected the DRIVE (Digital Retinal Images for Vessel Extraction) [7] database as filament models in our approach. We divided the masks of the segmented retina image into two sub-parts, first vertically and then horizontally, resulting in two sub-parts measuring 128x256 and 256x128, respectively. We reconstruct an image of size 256x256 by filling the missing spaces with a black (empty) rectangle of dimensions 128x256 or 256x128. Additionally, for data augmentation, we perform for each picture rotations of 90°, 180° and 270°. So we obtain images containing pieces of segmented retina. The synthetic images $\tilde{\mathbf{y}}$ are designed such that:

$$\tilde{\mathbf{y}} = \mathbf{H}(\mathbf{v})(\mathbf{r} + \sigma \mathbf{w}), \quad (3)$$

where \mathbf{w} represents images of healthy grapevine wood, \mathbf{r} are images containing pieces of segmented retina. σ allows to define the SNR, according to :

$$\text{SNR} = 10 \log_{10} \left(\frac{\|\mathbf{H}\mathbf{r}\|^2}{\sigma S} \right). \quad (4)$$

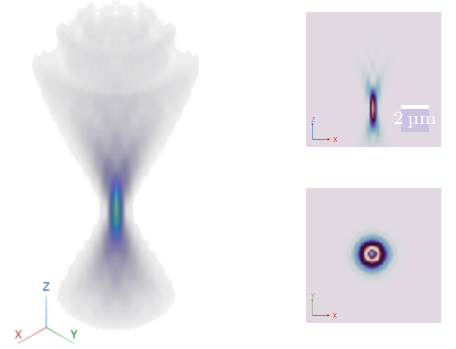


Fig. 2: Typical PSF \mathbf{h}^{3D} in fluorescence microscopy generated via PSF generator [5], with axial and longitudinal slices of the central plans.

$\mathbf{H}(\mathbf{v})$ is the 3D depth-selection operator for the 3D PSF, as described in (1). It relies on a continuous field $\mathbf{v} \in \mathbb{R}^S$, for which we make a Gaussian Markov Random Field assumption:

$$p(\mathbf{v}) \propto \exp\left(-\frac{1}{2}\mathbf{v}^T \boldsymbol{\Sigma}^{-1}\mathbf{v}\right) \quad (5)$$

We assume that $\boldsymbol{\Sigma} \in \mathbb{R}^{S \times S}$ is a circulant covariance matrix, whose basis is parameterized by a Gaussian correlation function.

Summarizing, the image generation is as follows:

- select a snippet \mathbf{w} from a healthy wood image.
- select a same-sized snippet \mathbf{r} from the DRIVE database, color and crop it.
- sample a GMRF realization \mathbf{v} (5).
- blend the images according to (3) and a given target SNR (4).

Fig. 3 illustrates the process.

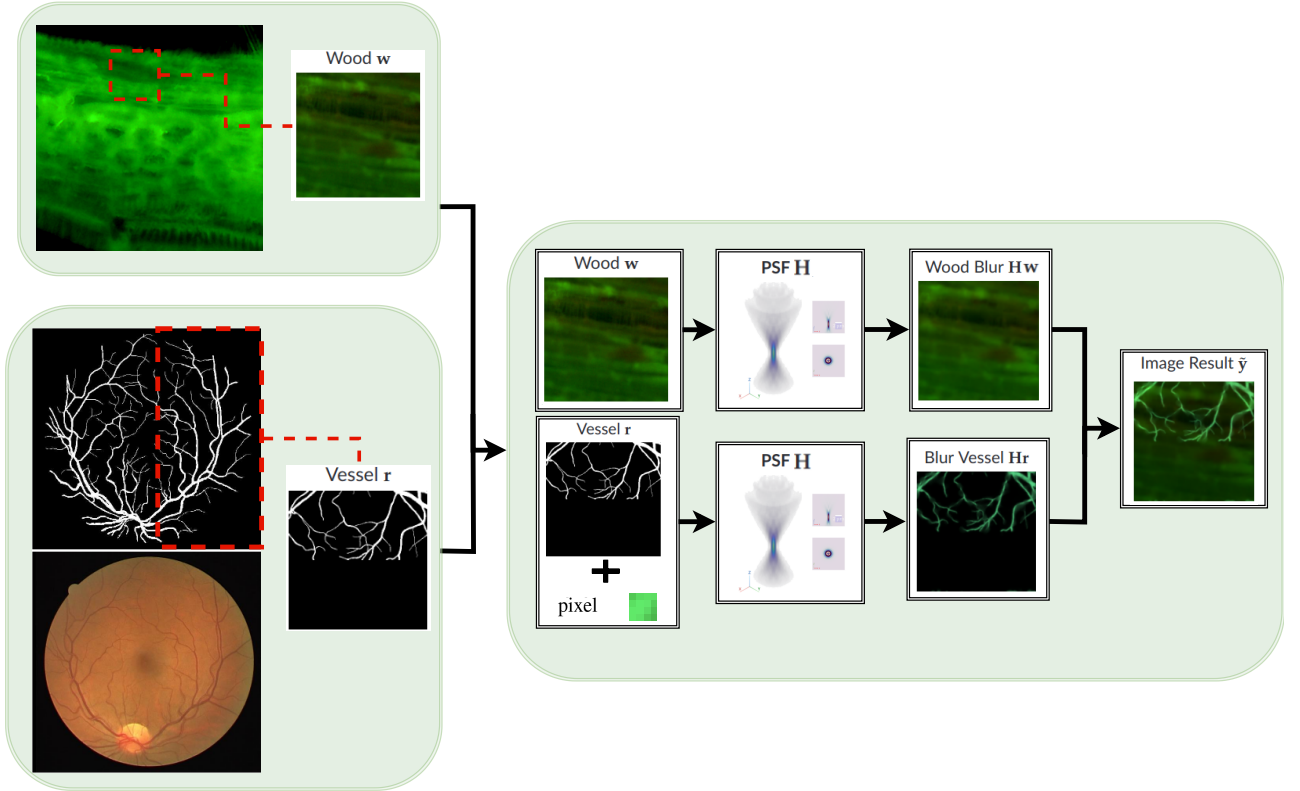


Fig. 3: Illustration of the synthetic image formation process. \mathbf{y} correspond to a SNR = -9dB.

3 Numerical results

To evaluate our approach, we test how its use impacts the segmentation of images, depending on the choice of the training and testing databases. To measure the quality of the segmentation, we will use the Accuracy metric, which allows to measure the total percentage of correctly predicted pixels. To show performance, we only present the accuracy score to keep the reading straightforward.

3.1 Segmentation methods

- *Pre-processing.* Prior to segmentation, we extract relevant features as a pre-processing. The latter were identified in a preliminary random forest-based study using the Ilastik software [8]. The features were sorted along their Gini importance [9], and the 14 first features were retained (see Table 1). In addition, preliminary experiments have shown that including classical image filters (namely Gabor, Sobel, Roberts, Scharr, and Prewitt filters) also improves segmentation. Then, for any 3-channel RGB input image, the pre-processed version contains 57 channels.
- *Random Forests.* The random forest algorithm [10] requires several parameters to be selected. The number of decision trees was selected from grid search and we set 80 trees in this model. The other RF parameters are set to have fully-grown and unpruned trees.

Identity	Hessian	Gaussian $\sigma = \{3; 5; 7\}$	Gradient Magnitude $\sigma = 2$
Difference of Gaussian (σ_i, σ_j) = $\{(1, 3.5); (1, 12); (1, 30)\}$	Laplacian of Gaussian $\sigma =$ $\{0.5; 1.6; 3\}$	Eigenvalues of Tensor Structure $\sigma = 0.7$	Hessian of Gaussian Eigenvalues $\sigma = 3.5$

Table 1: 14 features retained based on their Gini importance.

- *U-Net*. The U-Net model is widely recognized for its effectiveness in image segmentation and is commonly used for this task [11]. To implement a U-Net model, we selected from preliminary experiment 4 depth levels, from 256×256 input images to $16 \times 16 \times 256$ bottleneck, and a block convolution parameter α of 0.2, 0.2, 0.1 and 0.1 for each level. The architecture weights were estimated with the Adam optimizer, using the binary cross entropy criterion. The optimizer runs for at most 1000 epochs, stops when accuracy decreases, and handles each image one by one (the batch are one-sized). We also set the input layer depth as 16, *i.e.* 16 features are first generated. The former takes RGB images as input, while the latter takes the pre-processed images as input.

3.2 Synthetic and real image databases

We design and make use of several databases (Fig. 4) in order to validate the proposed approach:

- The first dataset contains only synthetic images, produced along the procedure detailed in Subsection 2.2. We name it *dataset A* in the following, and it contains 427 images.
- The second dataset exclusively contains real images together with expert-labeled ground truths. We label it as *dataset B*. We split this dataset depending on the content of the image:
 - with fungi in lower quality image, presenting notably a higher blur: *dataset B1*, 247 images.
 - without fungi: *dataset B2*, 312 images.
 - with fungi : *dataset B3*, 569 images
- *Dataset C* contains images from both datasets A, B1 and B3, with 128 images.

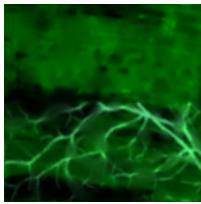
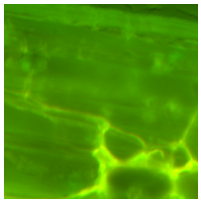
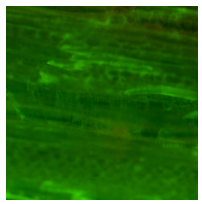
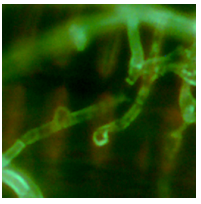
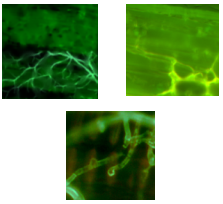




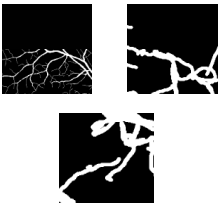
dataset name	A	B : real images			C
		B1	B2	B3	
description	synthetic images	with fungi in lower quality image	without fungi	with fungi	mix A, B1 and B3
number of images	427	247	312	569	128
images					
segmentation					

Fig. 4: Overview of our 5 datasets.

The database resulting from datasets A and B will be made available online. Note that we made the choice of relatively small images (256^2 pixels) to ensure diversity in learning, yielding a relatively high number of thumbnails. This is to be contrasted with the low number of acquisitions (15 images) used to form thumbnails in dataset B.

Then, we use datasets A, B1, and C for training. For testing, we target at first dataset B1 to cover difficult cases specifically, then dataset B to verify the assessment of true negatives, *i.e.* in the absence of fungi, and then dataset B3 to assess the presence of fungi. The dataset overlaps and use is depicted in Fig. 5.

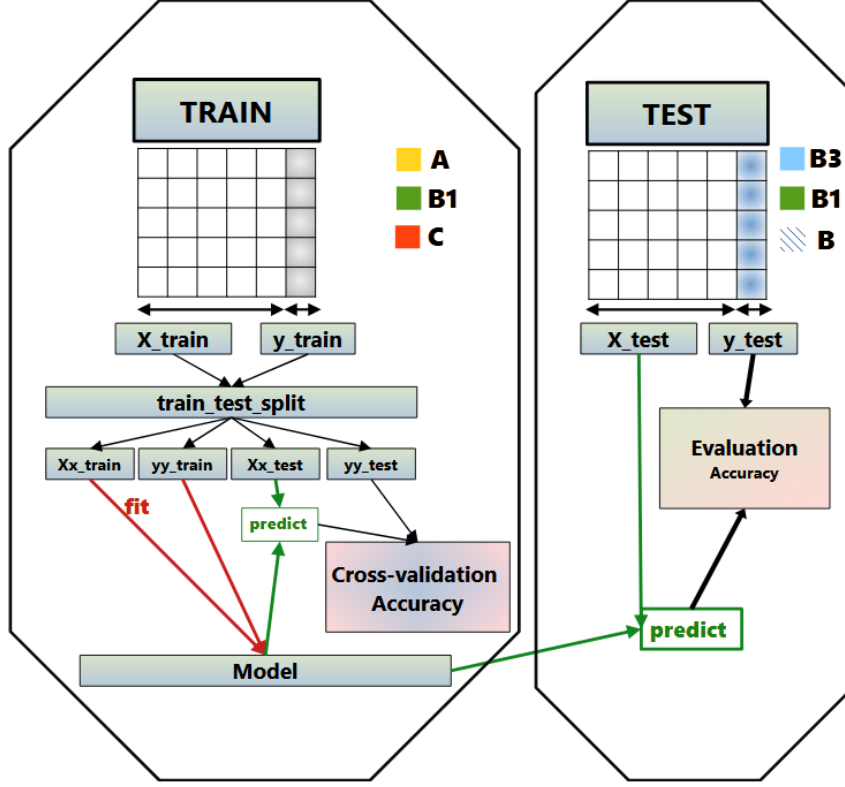


Fig. 5: Datasets split for the different evaluation processes run in this paper.

3.3 Results

Table 2 depicts the results (accuracy) we obtained for the three models of each learning method, using cross validation with 5 folds. For each dataset, the best method result is highlighted in bold font. Fig. 6 also depicts some results. The main outcomes can be summarized as follows:

- U-Net 57D yields better results than U-Net 3D for the segmentation of dataset B1 (93.5% instead of 90.1%). This shows that the pre-processing is helpful in the process, despite the change of dimensions.
- On B1 again, we observe an improvement of the results when incorporating synthetic images: the latter do allow an average 1.57% improvement of accuracy. An example of improvement can be seen in the first column of Fig. 6.
- Noteworthy, training on lower-quality images of B1 yields for the best segmentation results of the higher-quality datasets B and B3.
- Generally, the random forest remains a very powerful learning method, even when only training on synthetic image, as shown in columns 2–4 in Fig 6.
- Note: On retinal segmentation, it is common for the distribution of pixels between the two binary classes (e.g., "background" and "structures of interest," such as blood vessels) to be unbalanced. In other words, the majority of the pixels often belong to the background of the image, while a much smaller number corresponds to the specific structures we aim to segment.

		Random forest			U-Net 3D		
Test	Training database:	Dataset A	Dataset B1	Dataset C	Dataset A	Dataset B1	Dataset C
	CV accuracy	92.96%	96.16%	97.06%	89.68%	93.84%	90.96%
	Dataset B1	2.69%	94.13%	95.70%	18.98%	90.10%	83.25%
	Dataset B	41.00%	95.70%	92.45%	10.47%	94.17%	85.73%
	Dataset B3	6.58%	92.14%	88.08%	10.83%	93.79%	89.00%

Table 2: Segmentation accuracies, as a function of the segmentation method, the training database, and the testing database.

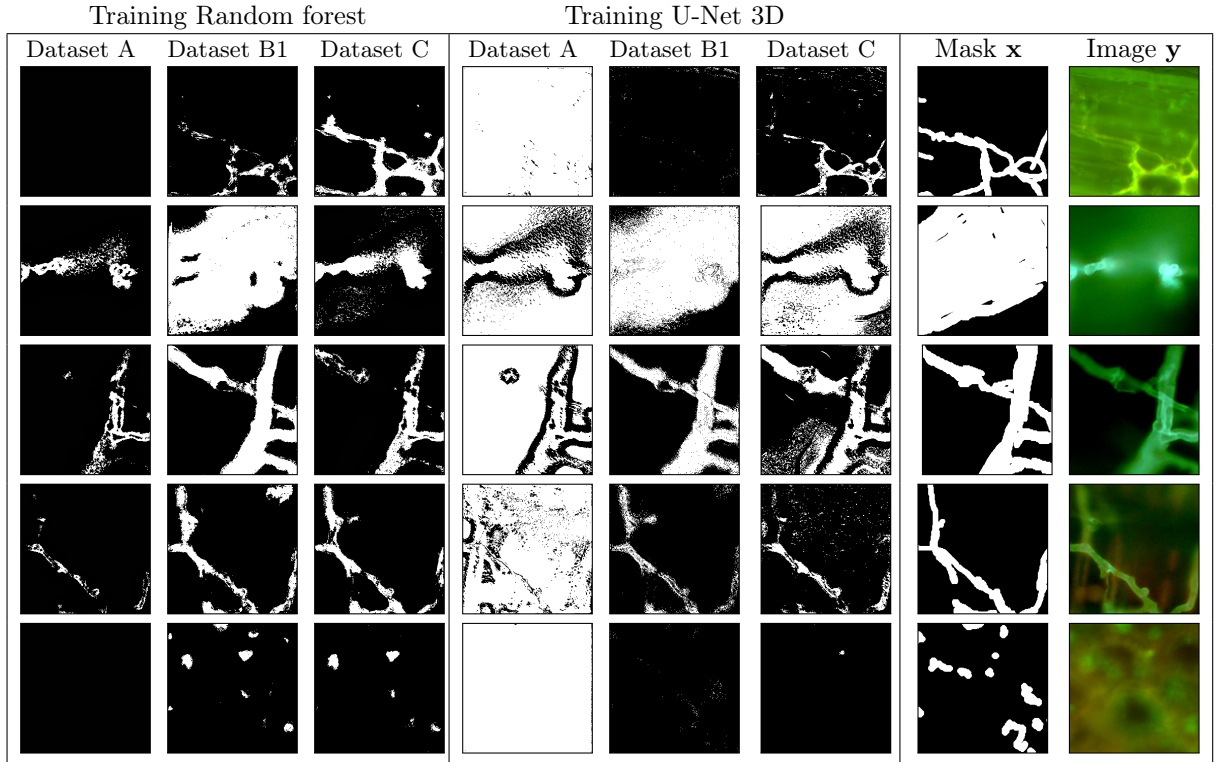


Fig. 6: Example of results on real vine fungi images.

4 Conclusion

In this work, we proposed a data augmentation technique dedicated to fluorescence microscopy images. We have shown that for the cleaner images, supervised segmentation is feasible even on a small database, *i.e.* without augmentation. However, we also showed that for lower-quality images, the addition of the synthetic images was indeed helpful, leading to noticeable accuracy improvements.

This work stems a few perspectives on the topic, as it could be generalized to other segmentation problems in fluorescence microscopy (such as tubulin networks in cells), as well as other imaging techniques (such as retina imaging).

References

1. Del Frari, G., Gobbi, A., Aggerbeck, M.R., Oliveira, H., Hansen, L.H., Ferreira, R.B.: Characterization of the wood mycobiome of vitis vinifera in a vineyard affected by esca. spatial distribution of fungal communities and their putative relation with leaf symptoms. *Frontiers in Plant Science* **10** (2019) 910
2. Chlap, P., Min, H., Vandenberg, N., Dowling, J., Holloway, L., Haworth, A.: A review of medical image data augmentation techniques for deep learning applications. *Journal of Medical Imaging and Radiation Oncology* **65**(5) (2021) 545–563
3. Shorten, C., Khoshgoftaar, T.M.: A survey on image data augmentation for deep learning. *Journal of big data* **6**(1) (2019) 1–48
4. Ma, J., Hu, C., Zhou, P., Jin, F., Wang, X., Huang, H.: Review of image augmentation used in deep learning-based material microscopic image segmentation. *Applied Sciences* **13**(11) (2023) 6478
5. Kirshner, H., Aguet, F., Sage, D., Unser, M.: 3-D PSF fitting for fluorescence microscopy: implementation and localization application. *Journal of Microscopy* **249**(1) (2013) 13–25
6. Schmidt-Erfurth, U., Sadeghipour, A., Gerendas, B.S., Waldstein, S.M., Bogunović, H.: Artificial intelligence in retina. *Progress in retinal and eye research* **67** (2018) 1–29
7. Challenge, G.: Drive: Digital retinal images for vessel extraction (Accessed on January 21st, 2024)
8. Berg, S., Kutra, D., Kroeger, T., Straehle, C.N., Kausler, B.X., Haubold, C., Schiegg, M., Ales, J., Beier, T., Rudy, M., et al.: Ilastik: interactive machine learning for (bio) image analysis. *Nature methods* **16**(12) (2019) 1226–1232
9. Hastie, T., Tibshirani, R., Friedman, J.: *The Elements of Statistical Learning*. Springer Series in Statistics. Springer New York Inc., New York, NY, USA (2001)
10. Breiman, L.: Random forests. *Machine learning* **45** (2001) 5–32
11. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III* 18, Springer (2015) 234–241