GAN meets chemometrics: Segmenting spectral images with pixel2pixel image translation with conditional generative adversarial networks

Puneet Mishra^{a,*}, Ittai Herrmann^b

^a Wageningen Food and Biobased Research, Bornse Weiland 9, P.O. Box 17, 6700AA, Wageningen, the Netherlands

^b The Robert H. Smith Institute for Plant Sciences and Genetics in Agriculture, The Hebrew University of Jerusalem, P.O. Box 12, Rehovot, 7610001, Israel

ARTICLE INFO

Keywords:

Artificial intelligence
Semantic segmentation
Hyperspectral imagery
Image context

ABSTRACT

In analytical chemistry, spectral imaging of complex analytical systems is commonly performed. A major task in spectral imaging analysis is to extract signals related to important analytes present in the imaged scene. Hence, the first task of spectral image analysis is to perform an image segmentation to extract the relevant signals to analyze. However, in the chemometric domain, the traditional image segmentation methods are limited either to threshold-based or pixel-wise classification, therefore, no approach uses the contextual information present in the imaging scene. This study presents for the first time a pixel2pixel (p2p) image translation using conditional generative adversarial networks (cGAN) for the segmentation of spectral images. The p2p cGAN trains two neural models simultaneously where one model (generator) learns to segment and the other model learns to detect if the segmentation performed by the generator model is correct. During the process of generation and detection, the model automatically learns to segment the spectral images accurately. As an application of the p2p cGAN, a case of segmenting visible and near-infrared spectral images of plants was presented. Furthermore, as a comparison, threshold-based and pixel-wise image classification based on partial least-square discriminant analysis were also presented. The results showed that the p2p cGAN based image translation performed the best segmentation task with an intersection over union score of 0.95 ± 0.04 . The advanced new DL based image processing approaches can complement spectral image processing.

1. Introduction

In analytical chemistry domain, spatially distributed spectral properties of complex analytical systems are widely explored with spectral imaging. For example, detection of microplastics [1], analysis of fish [2, 3] and meats [4,5], pharmaceutical formulations [6,7], and many more [8–11]. Spectral imaging captures information in two modes, i.e., imaging as well as spectroscopy, where the spatial information captured by imaging carries the context information of the imaged scene, while the spectral information captures the physicochemical properties of the samples [8,12].

The rich spatial and spectral information captured by spectral imaging is alone of no use and requires extensive data processing to derive conclusions from the data [8,13]. The major processing steps of spectral images are radiometric correction, pre-processing, segmentation, and data modelling [14,15]. The two primary steps, radiometric correction and pre-processing are standard procedures. For radiometric correction, the same white and dark references are used for all images, while the

pre-processing actions are versatile as they are image specific. For example, if the image is noisy then denoising can be performed [16] and in the case of pixels having additive and multiplicative effects, spectral normalization and derivatives can be deployed [14,17]. After radiometric correction and pre-processing resulted in high-quality spectral imagery, the main task is to extract the relevant spectral data for model training [12,14]. Hence, the segmentation of spectral images is of crucial importance as the next step of data modelling is highly dependent on the quality of the segmentation. Poor segmentation might include pixels in the training stage which were not related to the analytes of interest, hence, can complicate or deteriorate further data modelling and model use [15].

In the domain of chemometrics, several approaches are used for the segmentation of spectral images, the easiest approach is to set a threshold on some specific contrasting band images or use classical Otsu based image segmentation methods [12]. When the number of spectral images to analyze is low then a manual region of interest selection with polygon drawing can also be performed [12,14]. In more complex scenarios,

* Corresponding author.

E-mail address: puneet.mishra@wur.nl (P. Mishra).

<https://doi.org/10.1016/j.chemolab.2021.104362>

Received 8 April 2021; Received in revised form 6 May 2021; Accepted 31 May 2021

Available online 4 June 2021

0169-7439/© 2021 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

pixel-based classifiers can be trained and used for pixel-based segmentation of spectral images [18]. The main challenge with threshold, Otsu and pixel classifier-based segmentation methods is that in many cases there are multiple analytes present in the imaged scene that may have similar spectral signatures, hence, in that case, the classical approaches can lead to poor segmentation of spectral images. However, in such a case, using the spatial contextual information present in the imaged scene can play a key role to achieve an efficient segmentation [19], as different analytes may have similar spectral signal and may appear in a different context in the imaged scene.

Although the widely used approaches to spectral image segmentation and modelling are limited to pixel-based modelling without considering the spatial context in the images [8,14,17], some recent works have realized the value of spatial information and have proposed approaches to combine the spatial and spectral information to achieve improved results which were otherwise unachievable with modelling solely on the spectral data. For example, a work proposed butterfly approach to learn from both topological and latent space of the spectral information to perform the segmentation of the spectral images [20]. Another recent work proposed to extract the textural features from the averaged (in the bands) spectral images and stacked them with the spectral planes to train a new pixel-based classifier [21]. Similarly, another recent work proposed to use 2D-wavelet decomposition where the spectral band images were decomposed to horizontal, vertical, diagonal, and approximate component and later stacked with the spectral planes to train pixel-based classifiers [22]. Recently, another work has realized the importance of spatial context and has proposed a method for posterior analysis of predicted spectral maps to evaluate their quality [23]. The main challenge with the currently used spatial-spectral analysis (segmentation/modelling) for spectral imaging is that they extract the spatial and spectral information in separate steps, further some methods only aims to extract the spatial information but later requires a new modelling step to use the extract information [21,22]. However, with the recent advancement in the computer vision domain and deep learning (DL) algorithms, several approaches are now available that jointly learns the spatial context and spectral patterns. For examples, the 2D and 3D convolution-based DL models are widely used to extract deep contextual and band-related information to achieve efficient modelling tasks [24].

Concerning image segmentation with joint spatial context and bands information, the DL domain provides a variety of flexible methods. Furthermore, approaches to both instance and semantic segmentation are available [25,26]. The first and the easiest to implement approach is the fully connected convolutional neural network that allows training the convolutional filters to learn the segmentation task on a set of training images and later segment new test images with the trained model weights [26]. Recently, a new DL architecture based on contracting and expansive path was proposed for the segmentation of biomedical images [27]. The new architecture was called the U-net as it was based on a set of down-sampling and up sampling convolutions, which make the network architecture appears like a U shape architecture. With further advancements, new segmentation approaches based on atrous and dilated convolutions were proposed, such as DeepLab [28,29]. Apart from the standard convolution-based networks, recently generative models were proposed for images translation tasks. In an easy language, image translation can be understood as tasks such as transforming satellite photographs to street maps, coloring black and white photographs, and transforming sketches of products to product photographs. An example of images translation can be understood as the image segmentation where a spectral image can be translated to either binary or multi-instance segmentation maps. A recently developed image translation approach is the pixel2pixel (p2p) conditional generative adversarial network (cGAN) [30], where two separate models are trained simultaneously in an adversarial process i.e., one generator model for synthesizing images, and a discriminator model that classifies images as from the dataset or synthesized by the generator model. With training, the model learns to discriminate the synthesized images with the use of the discriminator

model and eventually the generator model learns the image translation to fool the discriminator model [30]. Once trained, the generator model can be used to perform the p2p translation of images. A key point to note that, unlike the traditional GAN based image translation which can synthesize output of any given random input, the p2p cGAN is a conditional framework [30], where the output to the generator is conditioned on the input image. Several applications of p2p cGAN for image translation can be found elsewhere [30]. To the best of our literature search, this is the first work proposing the concept of image translation for segmentation of the spectral images to the analytical chemistry and chemometrics community.

The study aims to present p2p image translation using cGAN for segmenting spectral images. The p2p cGAN trains two neural models simultaneously where one model (generator) learns to segment the spectral images and the other model learns to detect if the segmentation performed by the generator model is correct. Along the process of generation and detection, the model automatically learns to segment the spectral images accurately. A case study of segmenting visible and near-infrared spectral images of plants was presented and compared with threshold-based segmentation and pixel-wise image classification with partial least-square discriminant analysis (PLS-DA).

2. Materials and methods

2.1. Data set

To explore the potential of p2p cGAN for segmenting spectral images, a spectral image of 36 *Arabidopsis thaliana* (ecotype Columbia) plants was used. The plants were sown inside a growth chamber with 16h: 8h light: dark cycle, air temperature of 25 °C and light intensity of 300 $\mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$. The image was acquired with a HySpex VNIR-1800 (Norsk Elektro Optikk, Oslo, Norway) from a 1 m distance from the top view. The spectral range was 407–997 nm in a sampling interval of 3.26 nm making a total of 186 bands. The raw data were converted into radiance using the HyperspecRad software (Norsk Elektro Optikk, Oslo, Norway) and then converted into relative reflectance by the HyperspecRef software (Norsk Elektro Optikk, Oslo, Norway) based on a reference panel imaged before plant scan. The total imaged data was $3000 \times 2796 \times 186$ (first two dimensions were spatial and the third dimension was spectral). The imaged vegetation scene consists of 36 individual potted plants imaged from top view. The size of the complete image data was $3000 \times 2796 \times 186$ where, the first two dimensions were the spatial dimensions and the third was the spectral dimension. We divided the imaged data into a model training part which was the image scene of $1976 \times 2796 \times 186$ consisting of 24 plants in total, and an independent test part which was an image scene of $1024 \times 2796 \times 186$ consisting of 12 plants. All model training and validation were performed on the calibration set and the final model testing were performed on the independent test set. Since, the training and validation of the semantic segmentation model require a ground truth segmentation mask against which the model performance can be judged, hence, the ground truth labels were generated manually in the MATLAB 2018b (The MathWorks, Natick, MA, USA). The ground truth masks were generated using the 'roipoly' function in MATLAB that allows to draw-free hand polygons to generate segmentation masks.

The images were pre-processed with variable sorting for normalization (VSN) [31] to correct for the illumination effects caused by the local curvature of plant leaves and geometries. The need for such a pre-processing is recently highlighted in the scientific literature [11,32]. Further, a principal component analysis (PCA) [33] was performed to reduce the total number of variables from 186 bands to 4 components. The PCA compression was required as otherwise, the computer can run out of memory during computation. After compression, the sizes of the calibration and independent test set were $1976 \times 2796 \times 4$ and $1024 \times 2796 \times 4$, respectively. Further, to train and test the model, random images patches of size $512 \times 512 \times 4$, were extracted from calibration

Fig. 1. A schematic of conditional generative adversarial network to translate PCA transformed spectral images to segmentation maps. The discriminator, D, learns to classify between fake (segmentation synthesized by the generator) and real PCA transformed images and manually segmented ground truths. The generator, G, learns to fool the discriminator. Eventually, after training, the G model learn to synthesize plausible segmentation mask.

(~500) and independent test data set (~1000), respectively. All model performances were accessed by estimating the average intersection over union (IoU) scores achieved for the 1000 image patches in the independent test set.

2.2. Pixel2Pixel image translation with conditional generative adversarial network

Pixel2Pixel image translation [30] is a type of conditional generative adversarial networks (cGANs) that learn a mapping from observed PCA transformed spectra image x and random noise vector z , to segmentation mask y , $G: (x, z) \rightarrow y$, where G is the generator model trained to produce segmentation masks that cannot be distinguished from “real” images by an adversarial trained discriminator, D . Model D is also trained to detect the generator's fake segmentation masks. The schematic view of the cGANs is provided in Fig. 1. The G model is provided with only the PCA transformed spectral images and G model tries to generate the segmentation mask for it. The D model is provided with both the PCA transformed spectral image, and the manually labelled target segmentation maps. D model identifies if the G model synthesized segmentation masks are fake or real, along the training process the G model learns how to generate plausible segmentation masks of the PCA transformed spectral images and the D model learns to detect if the G model synthesizes segmentation masks are real or fake. After some training, an equilibrium is reached, after which G model can synthesize plausible segmentation masks. The objective function of the cGANs [30] can be understood as Eq. (1):

$$\mathcal{L}_{cGAN}(G, D) = \mathbb{E}_{x,y}[\log D(x, y)] + \mathbb{E}_{x,y}[\log(1 - D(x, G(x, z)))] \quad (1)$$

where G tries to minimize this objective against an adversarial D that tries to maximize it [30] as Eq. (2).

$$G^* = \underset{G}{\operatorname{argminmax}}_D \mathcal{L}_{cGAN}(G, D) \quad (2)$$

In previous research, mixing the GAN objective with a more traditional loss i.e., L1 was found to be of high use for G model to achieve the near ground truth output in L1 sense [30]. L1 was used as it encourages less blurring during image synthesis compared to L2. The L1 on the G model can be imposed as Eq. (3):

$$\mathcal{L}_{L1}(G) = \mathbb{E}_{x,y,z}[|y - G(x, z)|] \quad (3)$$

Combining the objective (2) and (3) leads to the final objective Eq. (4):

$$G^* = \underset{G}{\operatorname{argminmax}}_D \mathcal{L}_{cGAN}(G, D) + \lambda \mathcal{L}_{L1}(G) \quad (4)$$

The G model in this study was a “U-net” architecture [27] and the D model was a convolutional “PatchGAN” classifier. The “PatchGAN” classifier was used to capture local style statistics [30]. The G model is

trained via adversarial loss, which encourages the G model to generate plausible segmentation maps. The G model is also updated via L1 loss measured between the generated segmentation and the manually labelled segmentation maps. The “U-net” based G model has an encoder part: C64-C128-C256-C512-C512-C512-C512 and a decoder part: CD512-CD512-CD512-C512-C256-C128-C64, where C means convolution, CD means transpose convolution and the number means convolutional filters. The D model architecture was: C64-C128-C256-C512, where C stands for convolution and the number stands for convolutional filters. For D model, the convolutional layers were connected with LeakyReLU with $\alpha = 0.2$, while the final layer was fed to a sigmoid activation function for binary classification related to detecting real or fake. The D model was compiled with adaptive moment (ADAM) optimizer [34], with a learning rate (LR) = 0.0002 and the loss function as ‘binary_crossentropy’ as the task was a binary classification of real or fake. The final cGAN model combines the G and the D model and uses the ‘binary_crossentropy’ to update the D model and the ‘mean absolute error’ to update the G model. The final cGAN model was compiled with ADAM optimizer with a LR = 0.0002. The total number of epochs were 100 and the batches per epoch were equal to the number of training samples, hence, the total number of training iterations were $100 \times$ number of training samples. Once the model was trained, the sole G model was used for synthesizing the segmentation masks. The performance of the segmentation was accessed with the intersection over union (IoU) scores calculated based on the synthesized segmentation mask and the ground truth segmentation masks. The final IoU score was reported as the mean and standard deviation of IoUs estimated on 1000 randomly extracted image patches from the test set image and the ground truth segmentation masks. The codes of the p2p cGAN model implemented in this study will be made available at: <https://github.com/puneetmishra2>.

2.3. Baseline comparison

As a baseline comparison to the p2p cGAN model, a commonly used threshold-based approach called normalized difference vegetation index (NDVI) was used [35,36]. The NDVI approach was used as the spectral images used for the demonstration were of plants that exhibit an extreme contrast after NDVI estimation. With extreme contrast between plant and background, a threshold can be decided to segment the plants from the background. The second baseline comparison was performed with a pixel-wise binary partial least-square discriminant analysis (PLS-DA) [37, 38]. The PLS-DA model was calibrated on manually extracted plant and background spectra, and later, the PLS-DA model was used to generate the binary segmentation mask for the complete spectral image pixel-wise. A key point to note is that unlike the cGAN model presented in this study, a key drawback of the pixel-wise PLS-DA modelling was that it does not take the spatial context of the images into account and its training was limited only to the spectral information present in the scene. The p2p cGAN modelling was performed in the Python (3.6) language, while the

Fig. 2. A sample plant image with soil background to demonstrate the drawback of threshold-based segmentation. (A) Normalized difference vegetation index image, (B) histogram plot for NDVI pixel intensities, and (C) segmentation based on threshold.

Fig. 3. A summary of pixel based PLS-DA segmentation analysis. (A) Mean spectra of plant and background, (B) Explained variance plot from PLS decomposition where 6 LVs were selected as optimal, (C) the PLS regression vector for binary classification, (D) PLS scores corresponding to first two principal components, and (E) confusion matrix for a argmax based pixel classification based on PLS prediction.

NDVI estimation and PLS-DA analysis were performed in MATLAB 2018b (The Mathworks, Natick, MA, USA).

3. Results and discussion

3.1. Baseline analysis with threshold and pixel-wise PLS-DA classification

As a baseline, two approaches were used for comparing the performance of the p2p cGAN based image translation. The first approach was the threshold-based segmentation of NDVI images, and the second approach was the pixel-wise segmentation based on the PLS-DA analysis. The results of the threshold-based segmentation are shown in Fig. 2. For exemplary purpose, a single plant was cropped from the large image scene and shown in Fig. 2. The NDVI estimated image is shown in Fig. 2A. The NDVI images attained high-intensity pixels for the plants while for the background the pixel has lower NDVI values. Although some pixels in the background seem to have high-intensity values which could be due to the presence of materials having similar spectral properties as leaves (e.g., moss). Overall, it can be noted that the plant attained a higher NDVI value which agrees with the NDVI range of -1 to 1 , where the healthy

green plants attain values near 1 . The histogram of pixel intensities of the NDVI image shows separate peaks that can be related to the plant and the background part of the images (Fig. 2B). For example, the peak near 0.8 is related to the plant (Fig. 2A) as the NDVI values are higher for the healthy green plant, while the peak near 0.4 can be related to the background which contains pixels also for material having similar spectral properties as plants. Based on two peaks in Fig. 2B, a threshold at 0.69 was set to segment the image (Fig. 2C). Several background pixels were segmented as plant pixels and several plant pixels especially related to the plant stem were segmented as background pixels. The inferior performance of the threshold-based segmentation was due to two reasons. First was the presence of materials having a similar spectral properties as the plant which made it difficult to set a suitable threshold. The second reason was that the threshold-based approach was a pixel-based approach and do not use the spatial context of the image, for example, in layperson terms, how a plant looks like.

After showing the drawbacks of pixel-based image segmentation, the results of pixels-based classification based on PLS-DA analysis are shown in Fig. 3. Fig. 3A shows the mean spectra for plant and the background material manually extracted with the region of interest selection. The

Fig. 4. An example segmentation performed by pixel-based application of the PLS-DA model. (A) NDVI image, (B) ground truth segmentation mask, and (C) PLS-DA predicted segmentation mask. The red circle in the figure presents the region where the PLS-DA based segmentation misclassified pixels. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

mean spectra of the plant (green) resemble the spectra of healthy green vegetation with a peak at 550 nm and a valley at 670 nm corresponding to the plant pigments such as chlorophylls [39]. Further, changes in the sharp rise of the reflectance between 670 and 720 nm (in the red-edge region), were related to plant photosynthetic activity [11]. Changes in reflectance in the region of >720 nm, were related to the leaf structure that was affected by moisture [36,40]. Then, to develop a pixel-based classifier, PLS analysis was carried out on a set of random 3000 spectra, where 1500/1500 belonged to plants/background pixels. The explained variance plot from PLS decomposition is shown in Fig. 3B. Visualizing the explained variance curve for classes (red), it can be noted that around 6 latent variables (LVs) the explained variable stabilizes, hence, 6 LVs were chosen to develop the final PLS model. The regression vector for the 6 LVs PLS model is shown in Fig. 3C. The regression vector received higher weights in the visible and the red-edge regions which can be related to the pigments present in the plants but absent in the background (excluding the moss). Further, the scores corresponding to the first two principal components of the PLS model are shown in Fig. 3D, where several background pixels achieved similar scores as the plants (black over green), thus, confirming that the background indeed has pixels with similar spectral properties as the plants. The performance of the PLS-DA model is presented as the confusion matrix in Fig. 3E. In the confusion matrix, it can be noted that several of the plant and background pixels were misclassified. Further, the portion of pixels of plants misclassified as the background was higher compared to the portion of background pixels misclassified as the plants. Overall classification

accuracy of 98.4% was achieved where 1.6% of pixels were misclassified. To have a visual understanding of the performance of the PLS-DA approach the model was applied to the independent test image and the segmentation mask was estimated. An exemplary plant from the segmented test image is shown in Fig. 4. To have a comparative statistical measure for the segmentation performance of the PLS-DA based pixel classification, the intersection over union (IoU) scores for the 1000 sub-samples images (512×512) from the independent test set was estimated. The IoU score for the PLS-DA based classification was 0.77 ± 0.10 . A key point to note was that the same set of 1000 samples images will be used for accessing the performance of the p2p cGAN based image segmentation, hence, the IoU score for the PLS-DA analysis can be directly compared with the IoU score of the p2p cGAN in the following section.

3.2. Pixel2pixel conditional generative adversarial network-based image translation for segmentation

The training performance of the p2p GAN model for segmenting plants from the background is shown in Fig. 5. Inside p2p GAN, there were two main models trained, i.e., discriminator model which detect if the synthesized images are real or fake, and the generator model which learns to synthesize the segmentation masks during training. In Fig. 5, it can be noted that the discriminator model attained a low loss compared to the generator model after some initial runs, however, the generator model loss was initially higher and slowly decreased as it learned to synthesize the segmentation masks. Finally, after 50,000 runs, the discriminator and generator model achieved an equilibrium, after which the generator model showed no further improvement. Hence, the final generator model was used to synthesize the segmentation mask for the independent test set image. Four exemplary images subsamples from the independent test set image are shown in Fig. 6. The p2p cGAN performed well and the results were remarkably like the ground truth segmentation masks. The IoU score estimated on 1000 subsamples images from the independent test set image was 0.95 ± 0.04 , higher compared to the performance of the pixels based PLS-DA classification.

4. Conclusions

In the domain of chemometrics, during spectral image processing, contextual information is minimally used and most of the times images are unfolded to matrices and pixel-wise data processing is performed. After performing the pixel-wise analysis, the matrices are usually reshaped to images and presented as spatial maps. However, unfolding the spectral images into matrices and processing data pixel-wise has a major drawback that it does not take the spatial context of the spectral

Fig. 5. Evolution of loss for discriminator and generator model during training of the pixel2pixel condition generative adversarial network.

Fig. 6. The performance of pixel2pixel conditional generative adversarial network (cGAN) for segmenting spectral images. From left to right, the 1st column represents false color spectral images, the 2nd column represents the ground truth segmentation masks, and the 3rd column represents the segmentation synthesized by the pixel2pixel cGAN. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

images into account and the operation of data unfolding loses the spatial context information present in the imaged scene. This study for the first time showed the novel implementation of pixel2pixel conditional generative adversarial networks (p2p cGAN) for processing spectral images using the spatial context and spectral information. A novel use case of translation spectral images to segmentation maps was presented. The results showed that the advanced method p2p cGAN using both the spatial context and spectral information outperformed the pixel-wise data processing approaches to spectral image processing. In the presented case of plants segmentation, the two pixel-based approaches i.e., threshold-based and pixels-based classification segmentation approaches suffered when the background has materials with similar spectral properties as the object to be segmented. However, p2p cGAN method using both the spatial context and spectral information was unaffected by the presence of unrelated pixels with similar spectral properties. Thus, it is concluded that the advanced image processing methods based on deep learning have brilliant potential to improve the processing of spectral images and particularly by using both the spatial context and spectral information.

Author statement

Puneet Mishra: Conceptualization; Methodology; Software; Writing - Original Draft; Data Curation; Ittai Herrmann: Methodology; Writing - Review & Editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] J. Shan, J. Zhao, Y. Zhang, L. Liu, F. Wu, X. Wang, Simple and rapid detection of microplastics in seawater using hyperspectral imaging technology, *Anal. Chim. Acta* 1050 (2019) 161–168.
- [2] J.-L. Xu, C. Riccioli, D.-W. Sun, Development of an alternative technique for rapid and accurate determination of fish caloric density based on hyperspectral imaging, *J. Food Eng.* 190 (2016) 185–194.

- [3] D. Wu, H. Shi, S. Wang, Y. He, Y. Bao, K. Liu, Rapid prediction of moisture content of dehydrated prawns using online hyperspectral imaging system, *Anal. Chim. Acta* 726 (2012) 57–66.
- [4] M. Kamruzzaman, G. ElMasry, D.-W. Sun, P. Allen, Prediction of some quality attributes of lamb meat using near-infrared hyperspectral imaging and multivariate analysis, *Anal. Chim. Acta* 714 (2012) 57–67.
- [5] D.F. Barbin, G. ElMasry, D.-W. Sun, P. Allen, Predicting quality and sensory attributes of pork using near-infrared hyperspectral imaging, *Anal. Chim. Acta* 719 (2012) 30–42.
- [6] L. Coic, P.-Y. Sacré, A. Dispas, C. De Bleye, M. Fillet, C. Ruckebusch, P. Hubert, E. Ziemons, Pixel-based Raman hyperspectral identification of complex pharmaceutical formulations, *Anal. Chim. Acta* 1155 (2021) 338361.
- [7] P.-Y. Sacré, P. Lebrun, P.-F. Chavez, C.D. Bleye, L. Netchacovitch, E. Rozet, R. Klinkenberg, B. Streel, P. Hubert, E. Ziemons, A new criterion to assess distributional homogeneity in hyperspectral images of solid pharmaceutical dosage forms, *Anal. Chim. Acta* 818 (2014) 7–14.
- [8] J.M. Amigo, I. Martí, A. Gowen, F. Marini, Chapter 9 - Hyperspectral Imaging and Chemometrics: A Perfect Combination for the Analysis of Food Structure, Composition and Quality, *Data Handling in Science and Technology*, Elsevier, 2013, pp. 343–370.
- [9] A.A. Gowen, C. O'Sullivan, C.P. O'Donnell, Terahertz time domain spectroscopy and imaging: emerging techniques for food process monitoring and quality control, *Trends Food Sci. Technol.* 25 (2012) 40–46.
- [10] A.A. Gowen, C.P. O'Donnell, P.J. Cullen, G. Downey, J.M. Frias, Hyperspectral imaging – an emerging process analytical tool for food quality and safety control, *Trends Food Sci. Technol.* 18 (2007) 590–598.
- [11] P. Mishra, S. Lohumi, H. Ahmad Khan, A. Nordon, Close-range hyperspectral imaging of whole plants for digital phenotyping: recent applications and illumination correction approaches, *Comput. Electron. Agric.* 178 (2020) 105780.
- [12] N. Mobaraki, J.M. Amigo, HYPER-Tools. A graphical user-friendly interface for hyperspectral image analysis, *Chemometr. Intell. Lab. Syst.* 172 (2018) 174–187.
- [13] L. Duponchel, Exploring hyperspectral imaging data sets with topological data analysis, *Anal. Chim. Acta* 1000 (2018) 123–131.
- [14] J.M. Amigo, H. Babamoradi, S. Elcoroaristizabal, Hyperspectral image analysis. A tutorial, *Anal. Chim. Acta* 896 (2015) 34–51.
- [15] M. Mäkelä, P. Geladi, M. Rissanen, L. Rautkari, O. Dahl, Hyperspectral near infrared image calibration and regression, *Anal. Chim. Acta* 1105 (2020) 56–63.
- [16] P. Mishra, A. Karami, A. Nordon, D.N. Rutledge, J.M. Roger, Automatic de-noising of close-range hyperspectral images with a wavelength-specific shearlet-based image noise reduction method, *Sensor. Actuator. B Chem.* 281 (2019) 1034–1044.
- [17] M. Vidal, J.M. Amigo, Pre-processing of hyperspectral images. Essential steps before image analysis, *Chemometr. Intell. Lab. Syst.* 117 (2012) 138–148.
- [18] P. Mishra, M. Schmuck, S. Roth, A. Nicol, A. Nordon, Homogenising and Segmenting Hyperspectral Images of Plants and Testing Chemicals in a High-Throughput Plant Phenotyping Setup, 2019 10th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing, WHISPERS), 2019, pp. 1–5.
- [19] L. Yilong, P.M. Panos, S. Eli, Spatial-spectral segmentation of hyperspectral images for subpixel target detection, *J. Appl. Remote Sens.* 13 (2019) 1–16.
- [20] N. Gorretta, J. Roger, G. Rabatel, V. Bellon-Maurel, C. Fiorio, C. Lelong, Hyperspectral Image Segmentation: the Butterfly Approach, 2009 First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, 2009, pp. 1–4.
- [21] P. Mishra, A. Nordon, M.S. Mohd Asaari, G. Lian, S. Redfern, Fusing spectral and textural information in near-infrared hyperspectral imaging to improve green tea classification modelling, *J. Food Eng.* 249 (2019) 40–47.
- [22] A. Nardecchia, R. Vitale, L. Duponchel, Fusing spectral and spatial information with 2-D stationary wavelet transform (SWT 2-D) for a deeper exploration of spectroscopic images, *Talanta* 224 (2021) 121835.
- [23] A. Herrero-Langreo, N. Gorretta, B. Tisseyre, A. Gowen, J.-L. Xu, G. Chaix, J.-M. Roger, Using spatial information for evaluating the quality of prediction maps from hyperspectral images: a geostatistical approach, *Anal. Chim. Acta* 1077 (2019) 116–128.
- [24] M.E. Paoletti, J.M. Haut, J. Plaza, A. Plaza, Deep learning classifiers for hyperspectral imaging: a review, *ISPRS J. Photogrammetry Remote Sens.* 158 (2019) 279–317.
- [25] L. Chen, G. Papandreou, I. Kokinos, K. Murphy, A.L. Yuille, DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs, *IEEE Trans. Pattern Anal. Mach. Intell.* 40 (2018) 834–848.
- [26] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: *IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 2015, pp. 3431–3440.
- [27] O. Ronneberger, P. Fischer, T. Brox, U-Net, *Convolutional Networks for Biomedical Image Segmentation*, Springer International Publishing, Cham, 2015, pp. 234–241.
- [28] L.-C. Chen, G. Papandreou, F. Schroff, H. Adam, Rethinking Atrous Convolution for Semantic Image Segmentation, 2017 arXiv preprint arXiv:1706.05587.
- [29] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, H. Adam, Encoder-decoder with Atrous Separable Convolution for Semantic Image Segmentation, pp. 801–818.
- [30] P. Isola, J.-Y. Zhu, T. Zhou, A.A. Efros, Image-to-image Translation with Conditional Adversarial Networks, pp. 1125–1134.
- [31] G. Rabatel, F. Marini, B. Walczak, J.-M. Roger, VSN: variable sorting for normalization, *J. Chemometr.* 34 (2020) e3164.
- [32] P. Mishra, G. Polder, A. Gowen, D.N. Rutledge, J.-M. Roger, Utilising variable sorting for normalisation to correct illumination effects in close-range spectral images of potato plants, *Biosyst. Eng.* 197 (2020) 318–323.
- [33] R. Bro, A.K. Smilde, Principal component analysis, *Anal. Methods* 6 (2014) 2812–2831.
- [34] D.P. Kingma, J. Ba, Adam, A Method for Stochastic Optimization, 2014 arXiv preprint arXiv:1412.6980.
- [35] M.S.M. Asaari, S. Mertens, S. Dhondt, D. Inze, N. Wuyts, P. Scheunders, Analysis of hyperspectral images for detection of drought stress and recovery in maize plants in a high-throughput phenotyping platform, *Comput. Electron. Agric.* 162 (2019) 749–758.
- [36] M.S.M. Asaari, P. Mishra, S. Mertens, S. Dhondt, D. Inze, N. Wuyts, P. Scheunders, Close-range hyperspectral image analysis for the early detection of stress responses in individual plants in a high-throughput phenotyping platform, *ISPRS J. Photogrammetry Remote Sens.* 138 (2018) 121–138.
- [37] P. Geladi, B.R. Kowalski, Partial least-squares regression: a tutorial, *Anal. Chim. Acta* 185 (1986) 1–17.
- [38] R.D. Cramer, Partial least squares (PLS): its strengths and limitations, *Perspect. Drug Discov. Des.* 1 (1993) 269–278.
- [39] P. Mishra, M.S.M. Asaari, A. Herrero-Langreo, S. Lohumi, B. Diezma, P. Scheunders, Close range hyperspectral imaging of plants: a review, *Biosyst. Eng.* 164 (2017) 49–67.
- [40] H.W. Gausman, W.A. Allen, Optical parameters of leaves of 30 plant species, *Plant Physiol.* 52 (1973) 57.