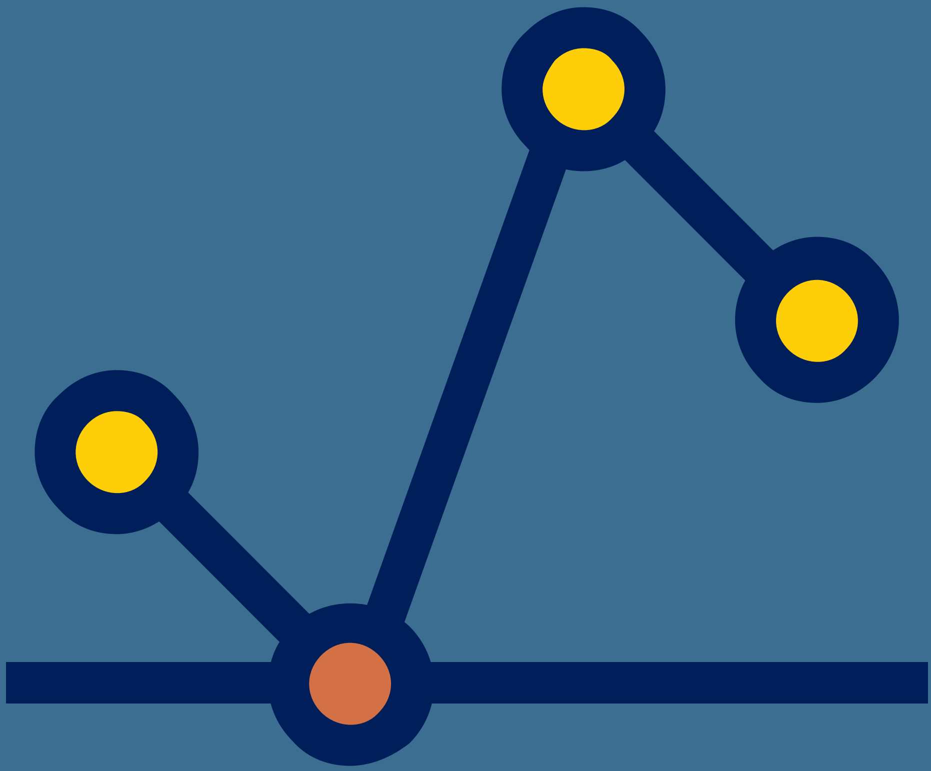
Edited by
Paola Cerchiello · Arianna Agosto
Silvia Osmetti · Alessandro Spelta

Proceedings of the Statistics and Data Science Conference



Pavia University Press

Copertina: Cristina Bernasconi, Milano

Via Salasco, 5 - 20136 Milano
Tel. 02/5836.5751 - Fax 02/5836.5753
egea.edizioni@unibocconi.it - www.egeaeditore.it

Pavia University Press
info@paviauniversitypress.it – www.paviauniversitypress.it

Prima edizione: maggio 2023
ISBN volume 978-88-6952-170-6

Increasing accuracy in classification models for the identification of plant species based on UAV images

Aumentare l'accuratezza nei modelli di classificaizone per l'identificazione delle species vegetali basati su immagini UAV

Anna Simonetto, Girma Tariku, Gianni Gilioli

Abstract The identification of plant species from RGB images is a challenge of growing importance in the field of biodiversity assessment. This study aims to develop an image pre-processing procedure that increases the accuracy of classification models applied to low-resolution plant images collected by RGB Unmanned aerial vehicles (UAVs). This procedure, based on contrast enhancement and super-resolution techniques, has been successfully tested on RGB images collected in agroecosystems.

Key words: Accuracy, Classification procedures, RGB images, Neural networks, K nearest neighbours

1 Introduction

Biodiversity assessment is a crucial aspect of the sustainable management of natural capital. In order to quickly and accurately analyse the biodiversity in wide areas, Unmanned aerial vehicles (UAVs) are increasingly being used due to their high mobility and ability to cover areas at different altitudes and locations with relatively lower costs [16]. The drawbacks is that at a high altitude, UAVs will have a low spatial resolution and it could make more difficult to detect features on plants [2].

Anna Simonetto
Department of Civil Engineering, Architecture, Land and Environment, and Mathematics (DICATAM), University of Brescia, via Branze 43, Brescia, e-mail: anna.simonetto@unibs.it

Girma Tariku
Department of Information Engineering (DII), University of Brescia, via Branze 38, Brescia e-mail: g.tariku@unibs.it

Gianni Gilioli
Department of Civil Engineering, Architecture, Land and Environment, and Mathematics (DICATAM), University of Brescia, via Branze 43, Brescia e-mail: gianni.gilioli@unibs.it

Furthermore, images taken by UAVs have shadows due to terrain factors [14]. The image pixel brightness of the shadow areas is compressed, and the information is deficient, which impacts the recognition of image information and thus limits the subsequent image application. Due to the development of different machine and deep learning topologies over the last few years, classification methods for plant identification from RGB images are focused on supervised learning techniques [7] by using machine learning tools [12] or transfer learning tools [13].

In our paper, we propose a pre-processing step based on image contrast enhancement [3] and super-resolution (SR) [15] image preprocessing technique for the training image dataset to improve the classification accuracy of low-resolution plant images of RGB UAVs. Although using high-quality images for image analysis would be ideal, this is not always possible in practice (e.g., attempting to identify relatively small objects in remote sensing imagery [10]). In these cases, performing transformations to increase image quality may prove useful in the attempt to identify and classify less salient objects in the imagery [3].

2 Procedure for plant image classification

The procedure used to generate the training image dataset is graphically illustrated in figure 1. The first step to produce an orthomosaic map of the studied area is collecting RGB images. We use DroneDeploy [9] to design an automated flight plan to take aerial pictures perfect for producing orthomosaic maps and 3D models. Then, we merge RGB images in the collection by looking for ”link points” and we use a digital surface model (DSM) to produce orthomosaic images from the drone photos. The orthophoto is segmented into useful single-picture objects. eCognition Developer v9.0.0 software [11] is used to segment the tree canopies applying the multi-resolution segmentation algorithm [1] to Orthomosaic image and DSM. Scale, shape, and compactness are the segmentation criteria. Following the segmentation process, plant species mapping is carried out for a limited set of pixels of image objects using the eCognition software, taking into account the spatial features of the image objects with regard to one another. A set of pixels would in this way act as a training sample for the classification technique known as supervised object-based classification [6]. We apply supervised learning using K nearest neighbours (KNN) to map plants. A class prediction for each group of pixels would then be generated by the KNN classification technique. Then, the training dataset plant pictures are extracted from the Orthomosaic image of the ground truth map with a class label. In the final step, we perform the classification using pre-trained transfer learning models.

The procedures proposed in this study focused on two steps (Image Contrast Enhancement and Super-Resolution preprocessing) as key elements to improve the classification accuracy of plant species. These steps are briefly described in the following sections.

2.1 Image Contrast Enhancement

Contrast enhancement is adjusting the dynamic range of pixel intensity distribution for good contrast enhancement in images facing low contrast concerns. we used an Image contrast enhancement technic for improving extracted UAV training dataset-images quality by decreasing the impact of distortions (i.e., blur, shadows, contrast issues, and noise) contained therein. The histogram of an image is an approximation of the pixel values distribution and can be calculated as:

$$h[i] = n_k/(M \times N) \qquad k \in [0,..,L-1]; \tag{1}$$

where I is the intensity of the pixel, L is the number of values (different intensities) that can be assumed by each pixel, n_k is the number of pixels with intensity k, $M \times N$ is the number of image pixels.

Starting from the consideration that good contrast images have a histogram close to uniform distribution U(0, L-1), we apply the following procedure:

- Compute the Image Histogram (probability distribution)
- Compute the Cumulative Distribution $F[k] = \sum_{i=0}^{k} h[i]$;
- Apply the point transformation $Ieq[i,j] = T[I[i,j]] = F[I[i,j]]$
- Rescale $Ieq[i,j]$ form $[0;1]$ to $[0;L-1]$.

where $Ieq[i,j]$ is the equalized image pixel intensity value at $(i = 0,...,M-1, j = 0,...,N-1)$, $T[I[i,j]]$ is point transformed image pixel intensity value at $(i = 0,...,M-1, j = 0,...,N-1)$ and $[I[i,j]]$ is the pixel intensity value K at $(i = 0,...,M-1, j = 0,...,N-1)$.

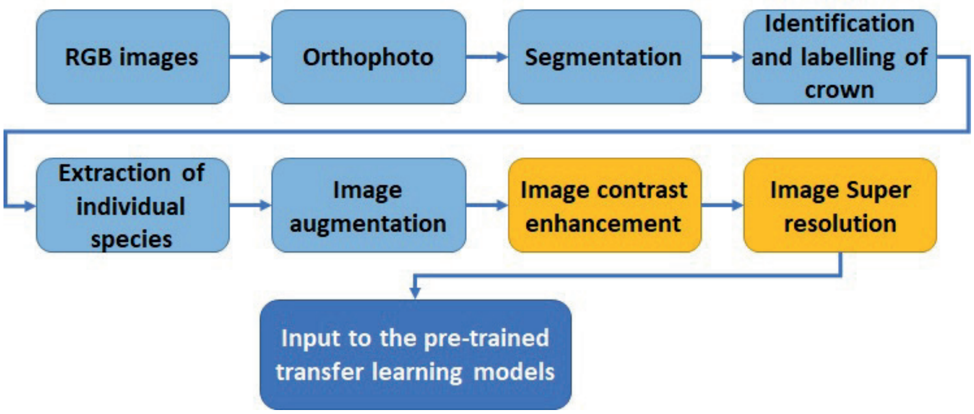


Fig. 1 Processing procedure to classify plant species starting from RGB images. The steps analysed in this study (contrast enhancement and super-resolution image preprocessing techniques) are highlighted in orange.

2.2 Super-Resolution preprocessing

When training a super-resolution network with a per-pixel loss function, the goal is to minimize the per-pixel difference between the output and the ground truth image. When using the perceptual loss function, high-resolution images are generated by minimizing the difference between high-level image features of the output and ground truth, which are extracted from a pre-trained convolutional neural network. We propose an algorithm of single image super-resolution using a generative adversarial network based on the work of Ledig et al [4] and graphically described in Figure 2.

For the Generator Network, the input is a LR image with 9*9 kernels with 64 filters and ReLU. Then B residual blocks are applied and each block has 3*3 kernel with 64 filters followed by batch normalization and ReLU. Then two sub-pixel convolution layers are applied to the up-sample image to 4x. In the Discriminator Network, a discriminator will also discriminate real HR images from generated SR images. It contains eight convolution layers with an increasing number of 3x3 filter kernels, increasing by a factor of 2 from 64 to 512 kernels. To reduce the image resolution, stride convolutions are applied each time the number of features is doubled. The resulting 512 kernel feature maps are followed by two dense layers and a final sigmoid activation function to obtain a probability for the real or fake image. For each layer of the neural network (orange blocks in Figure 2), we defined as activation function the Rectified Linear Unit (ReLu), and the value of each neuron needs to be calculated by the activation function to obtain a final value. In neural networks, the role of the activation function is to transform the neural network from linear to nonlinear, so that the neural network can better solve more complex problems.

Batch Normalization is applied to make neural networks faster and more stable adding extra layers in a deep neural network. The new layer performs the standardizing and normalizing operations on the input of a layer coming from a previous layer.

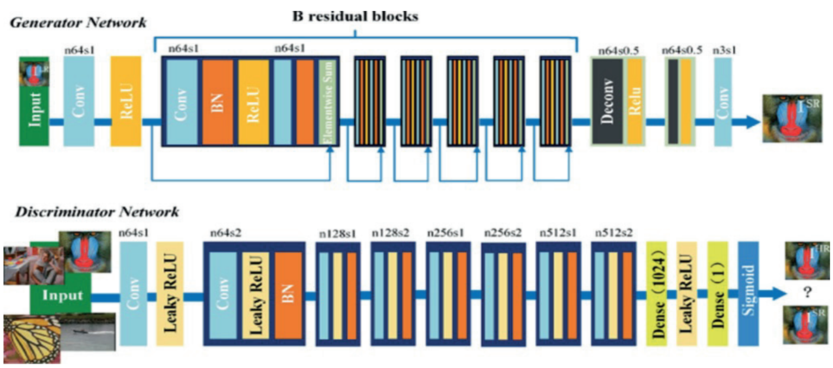


Fig. 2 Super resolution algorithm

We adopted mean squared error as loss function between content loss and adversarial loss. Adversarial loss tries to train the generator such that it produces natural-looking images which will be difficult for the discriminator to distinguish from real images.

3 Preliminary results

In Figure 3 we shown an application of the proposed algorithm to identify four plant species in an agroecosystem. We applied the ResNet50 pre-trained transfer learning model for four plant classes to 769 RGB UAV picture datasets. When we applied super-resolution and picture equalization, the classification accuracy jumped from 70.2% to 83.2%.

From our preliminary results, the proposed super-resolution and image contrast enhancement using histogram equalization methods improve the discriminative capacity of the applied classification model, helping to overcome the possible difficulties encountered when analysing low-resolution plant images collected by RGB UAVs.

References

1. Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I., Heynen, M.: Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS Journal of photogrammetry and remote sensing, **58(3-4)**, 239–258 (2004)
2. Chen, J., Chen, Z., Huang, R., You, H., Han, X., Yue, T., Zhou, G.: The Effects of Spatial Resolution and Resampling on the Classification Accuracy of Wetland Vegetation Species

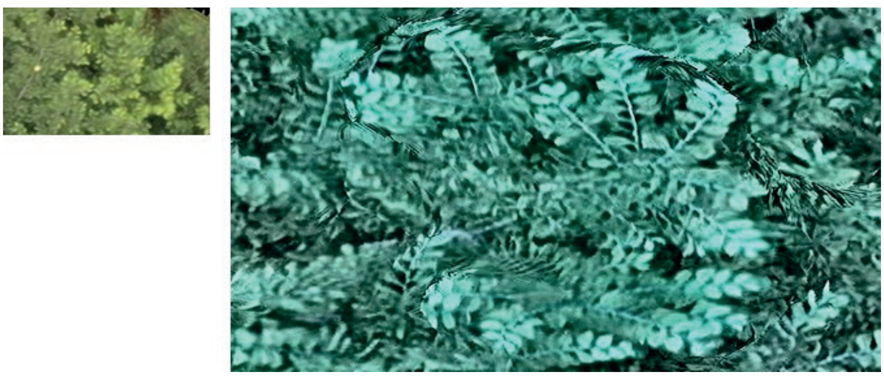


Fig. 3 Preliminary results of plant identification in an agroecosystem: on the left the original UAV image, on the right the equivalent super-resolution image using a generative adversarial network

and Ground Objects: A Study Based on High Spatial Resolution UAV Images. Drones. **7(1)**, 61 (2023)
3. González, D., Patricio, M. A., Berlanga, A., Molina, J. M.: A super-resolution enhancement of UAV images based on a convolutional neural network for mobile devices. Personal and Ubiquitous Computing. 1–12, (2019)
4. Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., ... Shi, W.: Photo-realistic single image super-resolution using a generative adversarial network. In Proceedings of the IEEE conference on computer vision and pattern recognition. 4681–4690 (2017)
5. Ma, L., Li, M., Ma, X., Cheng, L., Du, P., Liu, Y.: A review of supervised object-based land-cover image classification. ISPRS Journal of Photogrammetry and Remote Sensing. **130**, 277–293 (2017)
6. Ma, L., Li, M., Ma, X., Cheng, L., Du, P., Liu, Y.: A review of supervised object-based land-cover image classification. ISPRS Journal of Photogrammetry and Remote Sensing. **130**, 277–293 (2017)
7. Nasteski, V.: An overview of the supervised machine learning methods. Horizons. b. **4**, 51–62 (2017)
8. Natesan, S., Armenakis, C., Vepakomma, U.: Resnet-Based Tree Species Classification Using Uav Images. ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. **XLII-2/W13**, 475–481 (2019)
9. Putch, A.N.D.Y.: Linear measurement accuracy of DJI drone platforms and photogrammetry. San Francisco: DroneDeploy (2017)
10. Rabbi, J., Ray, N., Schubert, M., Chowdhury, S., Chao, D.: Small-object detection in remote sensing images with end-to-end edge-enhanced GAN and object detector network. Remote Sensing. **12(9)**, 1432 (2020)
11. Trimble Germany GmbH: Trimble Documentation eCognition Developer 10.1 Reference Book; Trimble Germany GmbH: Munich, Germany (2021)
12. Vercio, L. L., Amador, K., Bannister, J. J., Crites, S., Gutierrez, A., MacDonald, M. E., ... Forkert, N. D.: Supervised machine learning tools: a tutorial for clinicians. Journal of Neural Engineering. **17(6)**, 062001 (2020)
13. Weiss, K., Khoshgoftaar, T. M., Wang, D.: A survey of transfer learning. Journal of Big data. **3(1)**, 1–40 (2016)
14. Xi, W., Zuo, X., Sangaiah, A. K.: Enhancement of Unmanned Aerial Vehicle Image with Shadow Removal Based on Optimized Retinex Algorithm. Wireless Communications and Mobile Computing. (2022)
15. Yue, L., Shen, H., Li, J., Yuan, Q., Zhang, H., Zhang, L.: Image super-resolution: The techniques, applications, and future. Signal processing. **128**, 389–408. (2016)
16. Zhang, C., Kovacs, J. M.: The application of small unmanned aerial systems for precision agriculture: a review. Precision agriculture. **13**, 693–712 (2012)