Fake it till you make it: generating NDVI from RGB UAV imagery in viticulture

Jurrian Doornbos[1], Önder Babur[1,2] and João Valente[1]

***Abstract*— Unmanned Aerial Vehicles (UAVs) have emerged as flexible and efficient tools for diverse monitoring activities, owing to their high spatial resolution and cloud-cover independence. Applications in agriculture, forestry, and wildlife observation have widely embraced UAVs, catalyzing new management practices such as precision agriculture. However, these applications often rely on costly multispectral sensors that capture essential near-infrared (NIR) data, which standard RGB cameras lack. To address this limitation, Generative Adversarial Networks (GANs), specifically the pix2pix model, offer a promising solution by generating synthetic NDVI images from UAV RGB data. Despite minor underestimations in certain edge cases, the GAN demonstrates strong accuracy overall, and its potential for translating RGB images to other outputs holds exciting research prospects for broader access to complex sensor insights.**

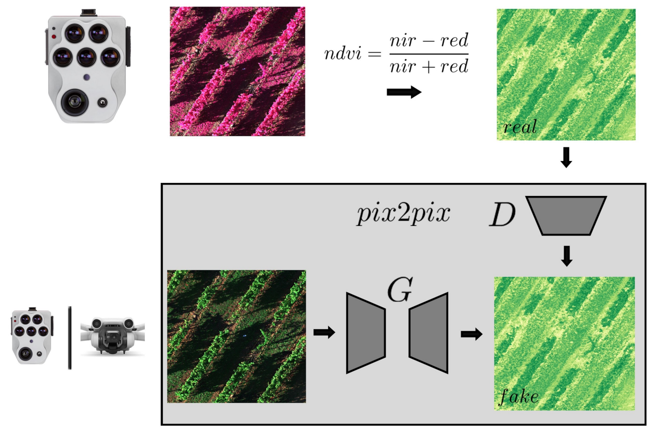

Fig. 1. pix2pix workflow for RGB-to-NDVI conversion

I. INTRODUCTION

The flexibility that the Unmanned Aerial Vehicle (UAV) sensing platform offers, in combination with the spatial resolution, which is also unaffected by cloud cover, makes it an ideal tool for a variety of monitoring activities [1], [2]. Activities such as leaf-level analysis in agriculture [3], and treetop analysis in forestry [4], to wildlife observations [5], all have embraced the UAV as an indispensable analytical tool. Powered by these possibilities new management practices are envisioned, practices such as precision agriculture [6]. These applications often make use of multispectral sensors [7]. A multispectral sensor captures the light intensity in reflectance at specific wavelengths of light, instead of the color representation of a 'standard' camera. This method of sensing light makes multispectral sensors useful for measuring features such as plant health, disease, water stress, etc. Standard RGB cameras lack the near-infrared band that is used by these applications. Most commonly, the near-infrared (NIR) band is used to calculate NDVI (Normalized Difference Vegetation Index) [8], which was developed in the 1970s for analyzing satellite imagery and estimating biomass, leaf area index, and potential yield.

In vineyard management, disease risk mapping has the potential to reduce pesticide use, but these assessment methods currently rely on NDVI maps, next to other input variables [9]. To achieve wider adoption, these approaches require more data for model development and evaluation. Which in turn requires more datasets. However, the multispectral sensor requires radiometric calibration; an understanding of how sunlight affects the reflectance; and is more expensive than the ever-popular RGB camera. Making data acquisition a resource-intensive activity.

A method of enabling a larger group to acquire data is through GANs (Generative Adversarial Networks). These networks generate synthetic (fake) imagery based on an input image. Using these approaches, existing datasets can be synthetically augmented with additional data without any fieldwork required. A similar work using GANs to convert RGB -¿ NDVI imagery reported positive assessment by human evaluation, although model accuracy was not clearly reported [10]. Furthermore, it focused on field crops and grasslands, which is a distinctly different environment from the woody crops in vineyards. This study presents the use of a GAN, namely pix2pix [11], trained on UAV RGB imagery to generate NDVI images in a vineyard.

II. METHODS

The pix2pix architecture consists of a generator and a discriminator, both are convolutional neural networks with 54.4M and 2.8M parameters respectively (noted as G and D in Fig. 1). The generator side takes in an input image and generates an output. The output is judged by the discriminator, which has the real output image. The judgment is given back to the discriminator which, with feedback from the discriminator generates a new image [11], this process is repeated thousands of times during training. For this study, an existing, open, dataset from the Bodegas Terras Gaudas

This work was supported by the European Commission under the project ICAERUS, grant number 101060643

[1]All authors are with the Information Technology Group, Wageningen University & Research, NL6707 KN Wageningen, The Netherlands

[2] Önder Babur is with the Department of Mathematics and Computer Science, Eindhoven University of Technology, NL5600 MB Eindhoven, The Netherlands

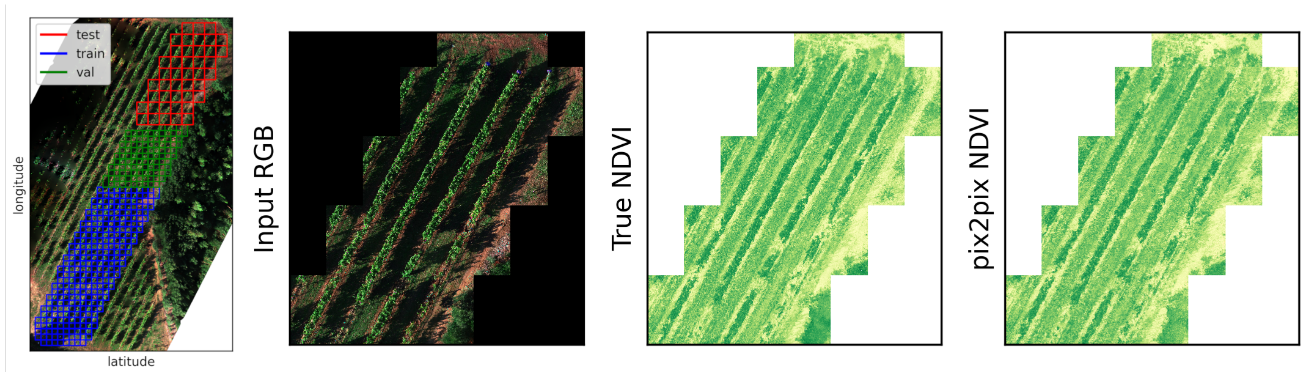

Fig. 2. Left: dataset chips for train (n=750), validation (n=258), with 128px overlap, and test (n=43) without overlap. Right: visual comparison of true NDVI and pix2pix NDVI test subset.

vineyard [12], with RGB and NDVI observations from a multispectral sensor [13] was used.

The resolution of the dataset is too large to be used directly for pix2pix, the large image was chipped into many smaller images. These chips were split into the training, test, and validation set, by selecting spatial invariant chips (see Fig. 2). The training and validation set have a 50% overlap between chips. Additionally, these chips were rotated and flipped to further increase the number of examples the pix2pix model sees during training. For the test set, there was no overlap in chips, and no augmentation was used.

The pix2pix model was trained for 200 epochs with the Adam optimizer, with a learning rate of 0.0002 during the first 100 epochs, which linearly decayed to 0 in the last 100 epochs. The model checkpoint after 200 epochs was used to generate NDVI images of the test set. All data and code is available online at: `github.com/jurriandoornbos/ndvi-pix2pix`.

III. RESULTS

The visual comarison between the true NDVI and the pix2pix-generated NDVI (Fig. 2) indicates a high degree of similarity between the two NDVI maps. The distinction between grass, soil, and vineyard crops is clearly discernible, while the impact of shadows on the NDVI value appears minimal.

Notably, the treetops at the edge of the vineyard (bottom right) exhibit lower NDVI values in the pix2pix estimation. Additionally, upon closer examination of a smaller subsection, some noise in the NDVI values is apparent. Furthermore, there are certain edge cases where the model underestimates the NDVI values. Particularly, NDVI values below 0, which have limited representation in the training set, are not adequately estimated by the pix2pix model.

Overall, the performance evaluation of the pix2pix model yields an R-squared value of 0.58 and a Root Mean Squared Error of 0.14, indicating a reasonable level of estimation accuracy, whilst not dealing too well with unseen data.

IV. CONCLUSIONS

This study shows the ability of GANs to generate NDVI maps from RGB data. Despite some minor underestimation in certain edge cases, the model shows promising accuracy in most NDVI estimations. However, there are a few clear next steps: 1. the generalization of NDVI estimation, 2. generating other outputs, and 3. exploring applications in existing pipelines.

1. Generalization in this case means the ability to translate the existing pix2pix model to a different vineyard, crop, region, and sensor, whilst reducing noise and increasing accuracy. This could potentially be achieved by increasing the dataset and training the pix2pix model for a longer period.

2. The pix2pix approach can be used for various translations, such as RGB to temperature and RGB to Water Index or RGB to Canopy Height (CHM). This enables a single RGB imagery to create additional outputs, synthetically increasing the input data variables that a model can use, without additional field missions.

3. Perhaps the most important step is to validate the output NDVI map. As NDVI is rarely the final result. This means using the generated NDVI as input for estimating botrytis risk in the vineyard, and comparing this approach to a fully multispectral one.

A generative approach to UAV imagery is a promising research direction and merits more attention, due to the potential of more affordable datasets and more robust models.

REFERENCES

[1] H. T. Berie and I. Burud, “Application of unmanned aerial vehicles in earth resources monitoring: Focus on evaluating potentials for forest monitoring in Ethiopia,” English, *European Journal of Remote Sensing*, vol. 51, no. 1, pp. 326–335, 2018, ISSN: 22797254. DOI: 10.1080/22797254.2018.1432993.

[2] N. Amarasingam, A. Ashan Salgadoe, K. Powell, L. Gonzalez, and S. Natarajan, “A review of UAV platforms, sensors, and applications for monitoring of sugarcane crops,” *Remote Sensing Applications: Society and Environment*, vol. 26, 2022. DOI: 10.1016/j.rsase.2022.100712.

[3] D. Tsouros, S. Bibi, and P. Sarigiannidis, "A review on UAV-based applications for precision agriculture," *Information (Switzerland)*, vol. 10, no. 11, 2019. DOI: 10.3390/info10110349.

[4] Y. Diez, S. Kentsch, M. Fukuda, M. Caceres, K. Moritake, and M. Cabezas, "Deep learning in forestry using uav-acquired rgb data: A practical review," *Remote Sensing*, vol. 13, no. 14, 2021. DOI: 10.3390/rs13142837.

[5] C.-U. Hyun, M. Park, and W. Y. Lee, "Remotely Piloted Aircraft System (RPAS)-Based Wildlife Detection: A Review and Case Studies in Maritime Antarctica," *Animals*, vol. 10, no. 12, 2020, ISSN: 2076-2615. DOI: 10.3390/ani10122387. [Online]. Available: https://www.mdpi.com/2076-2615/10/12/2387.

[6] P. Radoglou-Grammatikis, P. Sarigiannidis, T. Lagkas, and I. Moscholios, "A compilation of UAV applications for precision agriculture," *Computer Networks*, vol. 172, p. 107 148, May 2020, ISSN: 1389-1286. DOI: 10.1016/J.COMNET.2020.107148.

[7] L. Deng, Z. Mao, X. Li, Z. Hu, F. Duan, and Y. Yan, "UAV-based multispectral remote sensing for precision agriculture: A comparison between different cameras," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 146, pp. 124–136, Dec. 2018, ISSN: 0924-2716. DOI: 10.1016/J.ISPRSJPRS.2018.09.008.

[8] J. Rouse J. W., R. H. Haas, J. A. Schell, and D. W. Deering, "Monitoring vegetation systems in the Great Plains with ERTS," *NASA. Goddard Space Flight Center 3d ERTS-1 Symp., Vol. 1, Sect. A*, 1974.

[9] S. Vélez, M. Ariza-Sentís, and J. Valente, "Mapping the spatial variability of Botrytis bunch rot risk in vineyards using UAV multispectral imagery," *European Journal of Agronomy*, vol. 142, p. 126 691, Jan. 2023, ISSN: 1161-0301. DOI: 10.1016/J.EJA.2022.126691.

[10] C. Davidson, V. Jaganathan, A. N. Sivakumar, J. M. Czarnecki, and G. Chowdhary, "NDVI/NDRE prediction from standard RGB aerial imagery using deep learning," *Computers and Electronics in Agriculture*, vol. 203, p. 107 396, Dec. 2022, ISSN: 0168-1699. DOI: 10.1016/J.COMPAG.2022.107396.

[11] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks," *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, vol. 2017-January, pp. 5967–5976, Nov. 2016. DOI: 10.1109/CVPR.2017.632. [Online]. Available: https://arxiv.org/abs/1611.07004v3.

[12] S. Vélez, M. Ariza-Sentís, and J. Valente, "Dataset on unmanned aerial vehicle multispectral images acquired over a vineyard affected by Botrytis cinerea in northern Spain," *Data in Brief*, vol. 46, Feb. 2023, ISSN: 23523409. DOI: 10.1016/j.dib.2022.108876.

[13] MicaSense, *Altum-PT Integration Guide – MicaSense Knowledge Base*. [Online]. Available: https://support.micasense.com/hc/en-us/articles/4419868608407-Altum-PT-Integration-Guide.