UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE CIÊNCIAS, TECNOLOGIAS E SAÚDE DO CAMPUS ARARANGUÁ
CURSO DE GRADUAÇÃO EM ENGENHARIA DE COMPUTAÇÃO

Alecsander Pasqualli Gesser

A Comparison of Supervised Segmentation Methods Based on Convolutional Neural Networks for Weed-Mapping Identification in UAV Images

Araranguá
2022

Alecsander Pasqualli Gesser

A Comparison of Supervised Segmentation Methods Based on Convolutional Neural Networks for Weed-Mapping Identification in UAV Images

Trabalho de Conclusão de Curso submetido ao curso de Graduação em Engenharia de Computação do Centro de Ciências, Tecnologias e Saúde do Campus Araranguá da Universidade Federal de Santa Catarina como requisito parcial para a obtenção do título de Bacharel em Engenharia de Computação.

Orientador: Prof. Antônio Carlos Sobieranski, Dr.

Araranguá

2022

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Gesser, Alecsander
A Comparison of Supervised Segmentation Methods based on Convolutional Neural Networks for Weed-Mapping Identification in UAV images / Alecsander Gesser ; orientador, Antonio Sobieranski, 2022.
p.

Trabalho de Conclusão de Curso (graduação) - Universidade Federal de Santa Catarina, Campus Araranguá, Graduação em Engenharia de Computação, Araranguá, 2022.

Inclui referências.

1. Engenharia de Computação. 2. CNN. 3. Computer Vision. 4. Machine Learning. I. Sobieranski, Antonio. II. Universidade Federal de Santa Catarina. Graduação em Engenharia de Computação. III. Título.

Alecsander Pasqualli Gesser

A Comparison of Supervised Segmentation Methods Based on Convolutional Neural Networks for Weed-Mapping Identification in UAV Images

Este Trabalho de Conclusão de Curso foi julgado adequado para obtenção do título de Bacharel em Engenharia de Computação e aprovado em sua forma final pelo Curso de Engenharia de Computação.

Araranguá, 19 de dezembro de 2022.

ICP-Edu

Documento assinado digitalmente
Analucia Schiaffino Morales
Data: 21/12/2022 16:11:47-0300
CPF: ***.256.420-**
Verifique as assinaturas em https://v.ufsc.br

Profª Analucia Schiaffino Morales, Drª
Coordenadora do Curso

Banca examinadora

Documento assinado digitalmente
Antonio Carlos Sobieranski
Data: 21/12/2022 15:34:14-0300
CPF: ***.305.809-**
Verifique as assinaturas em https://v.ufsc.br

Prof. Antônio Carlos Sobieranski, Dr.
Orientador
Universidade Federal de Santa Catarina

Documento assinado digitalmente
ELIANE POZZEBON
Data: 21/12/2022 15:35:18-0300
CPF: ***.213.139-**
Verifique as assinaturas em https://v.ufsc.br

Profª. Eliane Pozzebon, Dra.
Avaliadora
Universidade Federal de Santa Catarina

Otávio Corrêa

Otavio de Oliveira Corrêa, Dr.
Avaliador
Universidade Federal de Pelotas

Araranguá, 2022

A Comparison of Supervised Segmentation Methods for Weed-Mapping Identification in UAV images based on Convolutional Neural Networks

Alecsander Pasqualli Gesser *   Antonio Carlos Sobieranski †

2022, September

Abstract

Precision Agriculture is a very important field of application, which is mainly determined by the use of high technology in agriculture. Its main goal is to increase productivity and quality, while making use of good practices to preserve the environment and at the same time optimize the use of agricultural inputs. One of the tool used in precision agriculture is the UAV, where an unmanned aerial vehicle used to image a specific area, targeting a large sampling at reduced time and costs requirement. UAV can be embedded by a light visible or multiespectral camera, allowing to identify in image several interesting patterns. One particularly useful analysis is the identification of weed, a very common kind of grass coexisting in the dominant culture. The present work proposes an comparison between state of the art convolutional neural network for identification and segmentation of weed *Cynodon* sp. in UAV images. Due its similarity in the visible spectrum of light, segmentation methods based on classical linear color metrics fail to properly identity the areas affected by this kind of grass. On the other hand, the use of Convolutional Neural Networks have been employed in a series of computer vision applications with success. The main goal of this work is to implement and validate the use of such convolutional approaches as a general problem-solver for weed mapping identification. The proposed approaches achieve 0.93 accuracy levels, enabling

Key-words: Computer Vision. Convolutional Neural Network. Weed Mapping. Segmentation

*alecsander.gesser@grad.ufsc.br
†a.sobieranski@ufsc.br

1 Introduction

Precision agriculture (PA) is a relatively new field of application mainly characterized by the use of high technology to increase productivity and quality in agriculture, prioritizing the study and the use of good practices to achieve higher productivity while preserving the environment in a sustainable manner (MCBRATNEY *et al.*, 2005). Examples of PA involves farm and data management, optimization of pesticides and nutrient applications according to the nutritional information on demand, and the business intelligence area such as crop marketing and telematics services (LOUDJANI *et al.*, 2014).

With the advent and price reduction of remote sensing devices, monitoring system are becoming a very desirable tool for many areas. Multispectral sensors carried by Unmanned Aerial vehicles (UAV) offers wide-range monitoring capacity, enabling high accuracy data acquisitions for decision support systems (GARCIA-MARTINEZ *et al.*, 2020). Application of UAV in agriculture varies from geo-referencing and monitoring of huge farming areas, security, animal livestock management to nutrient deficiency identification in cultures (TANG *et al.*, 2020).

Agriculture are taking advantage with the integration between remote sensing solutions for variable rate(VR) applications methods in agriculture. Through the precise analysis of nutrient or chemical demands (fungicides, insecticides, herbicide), applications can be proportionated, resulting in significant decreasing costs since its not exceeds the application rate indicated for the properly diagnoses (TIMMERMANN; GERHARDS; KUHBAUCH, 2003). Studies suggest that 28% to 90% of cost savings can be obtained by using a VR application (TIMMERMANN; GERHARDS; KUHBAUCH, 2003). Environment is also benefited by VR applications, providing a decreasing in ground and water contamination, soil dampening and compacting, and biodiversity influence impact. Moreover, precision agriculture can be combined with UAV+VR for weed management and control in a culture area. One example of weed is the *Cynodon Sp.* family, a very common kind of grass that may coexists in a predominant culture area (e.g.: cotton, sugarcane, tobacco), implying difficulties for productivity (REYNS *et al.*, 2002). The manual inspection of UAV images looking for *Cynodon Sp.* presence is a laborious and an error-prone task done by humans. According to the professional employed to label the infested areas, *Cynodon Sp.* is very similar to the predominant culture in terms of color information when the visual spectrum is used solely, and the correct discrimination by visual inspection may be ambiguous and inaccurate.

The use of digital image processing and pattern recognition techniques can be used for automating the identification of *Cynodon Sp.*. The literature reveals two main trends: methods based on classical approaches and convolutional neural networks. However, for the methods found over the literature several drawbacks can be verified, such as susceptibility to invariance such as light variations, flight instability, flickering acquisition, as well as shadow, brightness, white balance parameters (YAO; QIN; CHEN, X., 2019). On the other spectrum related to the nature of the classification problem, classical approaches to be effective need to be calibrated using a very specific domain, and this manner provides high accuracy and specificity, with low generalization. Neural networks, however, need huge amount of input samples to train the model properly, and may present limited specificity since the model is restricted to the quality of the input sample (for example: model is trained taking into account a specific domains) (ARAI; KAPOOR, 2020).

In this paper we explore computational approaches for weed segmentation in UAV images. For this purpose, convolutional-based methods were implemented and compared against the validation step dataset and related works over the literature. The following approaches include the *Segnet* and *UNet* convolutional neural network (CNN) architectures, as preliminary results indicates that *Segnet* approach is the most suitable for the weed-mapping segmentation, obtaining 72% overall accuracy with any training step, using only default *ImageNet* model weights.

The remainder of this paper is organized as follows: Section 2, related works where the research is shown and a comparison between classical and modern approaches are analyzed. Section 3, theory fundamentals with all the basics definitions to the works better understanding. Section 4, proposed approach for the selected challenge, an image segmentation model for vegetation and invader discriminator.

2 Related Works

A systematic literature reviews (SLR) was conducted in order to identify the state-of-the-art methods. Our SLR considered a window from 2017-2022 using the following databases: Researchgate, ScienceDirect, IEEE Xplorer, for algorithms proposing a weed mapping or segmentation of aerial images. Keywords used for such results: “UAV OR Aerial”, “weed OR cynodon”, “detection OR segmentation OR mapping OR classification”, some works discarded based on image dataset scope (culture area density, similarity, and textures appearances).

2.1 Classical Approaches

Many methods were found using the criteria above, the related works were categorized into classical and modern approaches. For the classical approaches, summarized in the Table 1, the distribution of methods and dataset characteristics quantified as Ground to Soil distance (GSD) commonly described as pixel size. Many works makes use of low GSD to max out the information data of given culture, therefore limiting UAV coverage area in a working Day (JIMÉNEZ-BRENES *et al.*, 2019; CASTRO *et al.*, 2017; GIROLAMO-NETO *et al.*, 2019). For the present development work, an acceptable 3 cm GSD value was defined by the Team, including data processing time and UAV area coverage. Besides interesting overall accuracy can be seen with the use of classical methods, it lacks in terms of generalization, performing well for the set of input data they were designed.

Popular classifiers such as Random Forest and K-Means (GASPAROVIC *et al.*, 2020; GIROLAMO-NETO *et al.*, 2019; GAO *et al.*, 2018), needs extra dimensions to describe the pixel more descriptively, reviewed articles makes the use of specific extra descriptors surrounding spectral, geometric, texture and spatial features.

- Gao *et al.* (2018). Presents a hybrid manner of weed segmentation, conducted at agricultural region of Merelbeke located in East-Flanders Province, Belgium, from a total area of 0.015 ha, combining pixel and object based features, the proposed method started from soil to vegetation discrimination making use of Excess of Green Vegetation Index and otsu’s threshold method, texture(GLCM) and a set of Geometry features, achieving results of 0.945 of overall accuracy.

- Jiménez-Brenes *et al.* (2019) suggests a deep research about vegetation index along side with extra spectral and spacial/physical dimensions/channels making use o RGB+RGNIR and Digital Surface Model (DSM), study happened at Cabra, Souther

Spain at a total area of 0.05 ha, started by ranking 26 distinct vegetation indexes and multiples threshold values, end up achieving 0.977 of overall accuracy

- Girolamo-Neto *et al.* (2019) Research occurred at Iraema Mill, located in São Paulo State, Brazil in a total area of 3 ha, for a distinct but similar culture named: *bermuda-grass*, making use of some color descriptors as: Green-Red Vegetation Index (GRVI), texture descriptors: Grey level co occurrence matrix, and finally for classification Random forest algorithm achieving 0.925 precision results for its set of tests.

- Gasparovic *et al.* (2020) follow an overall simple research using only color descriptors as: normalized green red difference and brightness index, but having a large dataset.

Table 1 – Classical Approaches

Author	Sensors	Data points	Features	GSD	OA
Gao *et al.* (2018)	RGB	6.66×10^7	GLCM; Geometry	0.26	0.89
Jiménez-Brenes *et al.* (2019)	RGB;NIR	7.49×10^8	VIs	0.70	0.93
Girolamo-Neto *et al.* (2019)	RGB	1.74×10^{10}	VIs; Texture	2.00	0.80
Gasparovic *et al.* (2020)	RGB	3.40×10^8	VIs	3.50	0.85

2.2 Modern Approaches

For the modern counterpart, with the advent of the Convolutional Neural Networks, several applications restricted due computational limitations are being implemented. Below at Table 2, the same behavior of classical approaches happen to the GSD value, low values contributes to great overall accuracy, but limiting UAV area coverage. Geo spacial area diversity (GAD) are not so discussed, few works creates a well generalized data set: making the use of data augmentation (KERKECH; HAFIANE; CANALS, 2020)(RAMIREZ *et al.*, 2020), extended temporal data acquisition (HAMYLTON *et al.*, 2020), wide spectral range (RAMIREZ *et al.*, 2020).

Table 2 – CNN Approaches

Author	Sensors	Architecture	Data set	GSD	OA
Barrero *et al.* (2016)	RGB	*	1	1.84	0.99
Huang *et al.* (2018)	RGB	FCN	1	0.30	0.92
Buddha *et al.* (2019)	RGB	Faster R-CNN	1	0.46	0.93
Kerkech, Hafiane e Canals (2020)	RGB	Segnet	1	1.00	0.95
Ramirez *et al.* (2020)	RGB	DeepLabv3	1	1.00	0.89
Hamylton *et al.* (2020)	RGB	LeNet	1	3.00	0.85
Zou *et al.* (2021)	RGB	UNET	1	0.50	0.98
Reedha *et al.* (2022)	RGB	ViT	1	0.50	0.98

3 Fundamentals Theory

The analysis stays upon a comparison between traditional computer vision and convolutional neural networks techniques, using as metrics of best overall performance the confusion matrix and processing time speed. Both methods were exposed to the same dataset and same parameters and conditions.

3.1 Precision Agriculture and UAV

Precision Agriculture is a management strategy that gathers, processes and analyzes temporal, spatial and individual data and combines it with other information to support management decisions according to estimated variability for improved resource use efficiency, productivity, quality, profitability and sustainability of agricultural production (ISPAG, 2019).

Figure 1 – DJI Phanton UAV

Source: (DJI, s.d.)

Unmanned Aerial Vehicle (UAV), usually known as Drone, an aircraft with no human pilot, have becoming a excellent tool for precision agriculture due to its excellent spectral and temporal resolutions when compared with satellites imagery (MCBRATNEY *et al.*, 2005). Gathering great attention from many area due to its many benefits: data acquisition, obtaining excellent resolution at multiples spectral bands, capturing vegetation spectral reflectance and many other classes, minimizing time and resource for field campaign to investigate weed occurrence. Each class acquired from UAV sensor has its own reflectance signature for each wavelength of electromagnetic spectrum, its signature being a excellent tool for crop monitoring, providing various spatial and temporal resolution results, easily scalable implying for great crop coverage.

3.2 Digital Image Processing and Computer Vision for UAV

With the advent of computer capabilities recently, at both performance and efficiency scopes, many areas were benefited by, enabling the use of never even thought integration. Agriculture, health, education and energy sectors are being revolutionized, moreover for the first segment, machine learning and all its derivatives(branches) has been broadly employed for many purposes.

Growth rate prediction from climate data makes the use of artificial neural networks, regression algorithms and gene-expression programming to achieve more precise reports (LIU *et al.*, 2021). Classification approaches uses convolutional neural networks, to digest sensing data for agricultural crop management(JIANG *et al.*, 2020; HAO *et al.*, 2020; PARVATHI; TAMIL SELVI, 2021). UAV image dataset are utilized for all cited approaches yield to great advantage to inspect larger areas, providing high spacial and temporal resolution at multi spectral images.

Figure 2 – Drone Camera Sensor

Source: (DJI, s.d.)

Figure 3 – Crop Flight Plan

Source: (MACHADO *et al.*, 2015)

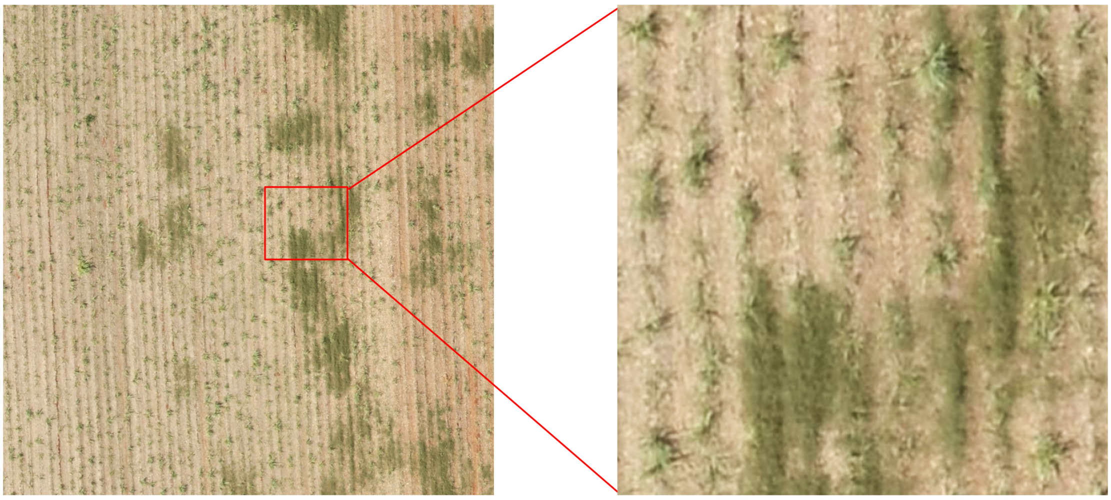
3.3 Machine Learning

Described as branch from artificial intelligence, the use of computer systems to reproduce human mind behavior, learning and adaptation capabilities from experience. The term refers to the automated detection of substantial patterns in data, has become a widely used tool in almost any application that requires large data sets extraction, being able to be found from cars to medicine complex systems (SHALEV-SHWARTZ; BEN-DAVID, 2014).

Machine learning can be categorized at two main subdivisions: supervised and unsupervised learning. At supervised learning, the objective is to predict an output measure based upon a given input, alternatively unsupervised learning objective is not to predict an output measure outcome but to describe the associations and patterns upon the given input (HASTIE; TIBSHIRANI; FRIEDMAN, 2009).

At current article supervised learning is applied, where the human-labeled training sets are composed by a pair of images and its associated area. From its fundamentals, supervised learning, have as veracious and correct all training set, consequently its quality

Figure 4 – Crop Result Zoom.



Source: created by author

and size are crucial to the success of the predictions made by the learner (MOHRI; ROSTAMIZADEH; TALWALKAR, 2018).

3.4 Convolutional Neural Network

A branch from machine learning, a kind of neural network for processing matrix structured data, has shown optimal results over the last decade to many areas, differing from ANN (Artificial neural networks) by the lower numbers of parameters, being able to be dismantle at main modules: Convolution, Pooling and Fully Connected Layers(Dense Layers).

The main advantage from classical models as: linear regressions, decision trees and random forest algorithm stays upon by the explicit instruction set absence, leading to better results as: generality performance, processing times and precision results.

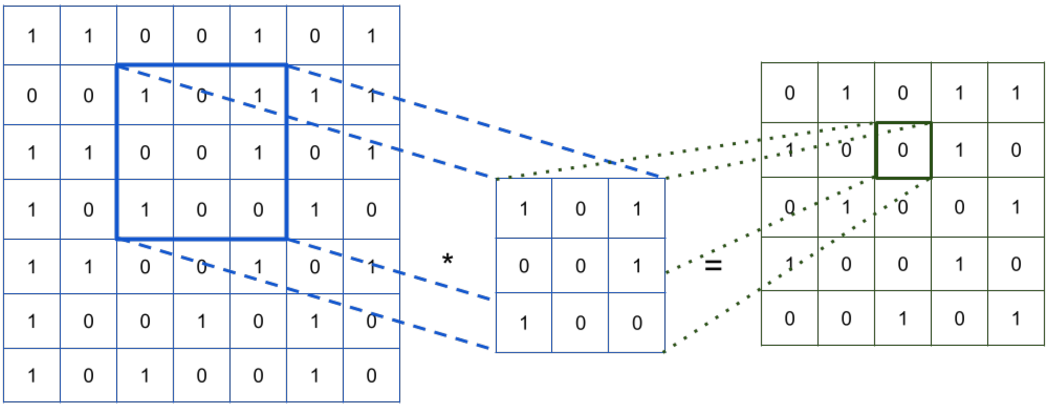
Dealing with algorithms and statistical models, large amount of data are needed to models infers data patterns. The algorithm employs a method called by *convolution* , and differs from ANN from its reduced sets of parameters and its capacity of extract abstract features as the input goes deeper into the layers (ALBAWI; MOHAMMED; AL-ZAWI, 2017; GOODFELLOW; BENGIO; COURVILLE, 2016). With the advent of computational developments and algorithmic optimizations, it has been more and more capable and easy to integrate at several systems.

3.4.1 Convolution

Specific kind of linear mathematical operation defined at Yamashita *et al.* (2018a) employed for feature extraction, where two dimensional array known as kernel (k) is applied throw the entire image input also being two dimensional array (I), following the expression below:

$$S(i,j) = (K * I)(x,y) = \sum_{i=-\frac{m}{2}}^{\frac{m}{2}} \sum_{j=-\frac{n}{2}}^{\frac{n}{2}} K(i,j)K(x+i, y+j) \tag{1}$$

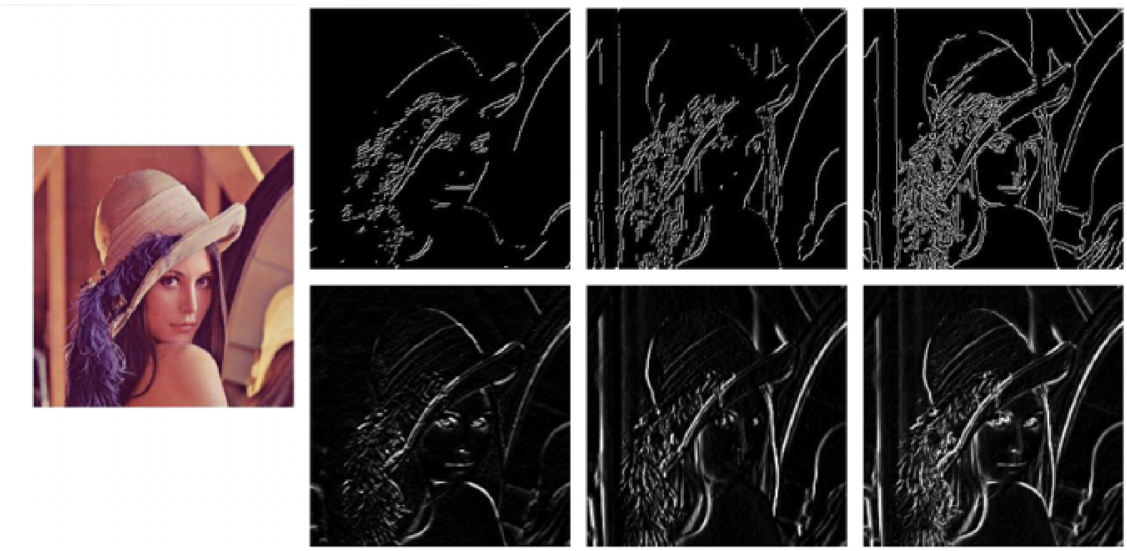
Figure 5 – Convolution between Image (I) and Kernel (K)



Source: created by author

Other use of this operation are in the use of filters for features image extractions as edge detection proposed for the Sobel and Canny extractors (SOBEL; FELDMAN, 1968; CANNY, 1986),

Figure 6 – Examples of the application of Canny (top row) and Sobel (bottom row) edge detection filters applied to computer graphics test image Lena.



Source: Jovanovic, Tuba e Simian (2012)

One of the most important roles of convolution for CNNs, is the reduction at the propagated information as it goes into the deeper layers, as exemplified at the figure 5, where one input of 7x7px convoluted with an kernel of 3x3px results in a 5x5px output.

3.4.2 Pooling

Despite convolution layer, polling state have as principal and single objective lowering the dimension from previous layers creating a summary statistic of the features maps (GOODFELLOW; BENGIO; COURVILLE, 2016). The most famous pooling algorithms are: *Max-pooling* and *Average-pooling*, the first selects the greater value from a region and the following selects the average value from a region as the below image (Figure 8) exemplifies.

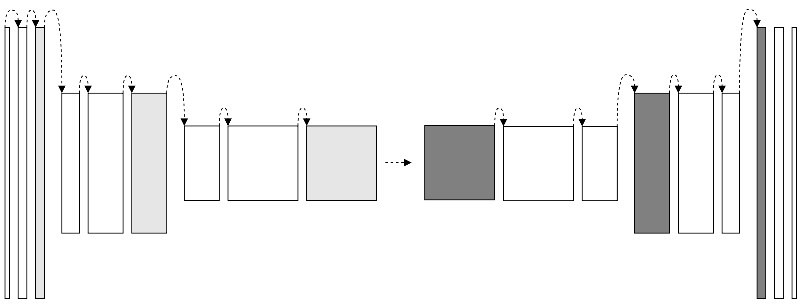
Figure 7 – Pooing

Source: created by author

3.4.3 Dense layer

Dense layer or deeply connected layer is where all inputs are connected to all outputs by a learnable weight (YAMASHITA *et al.*, 2018b). Once features extracted by convolution layers, down sampled by polling step, its features are mapped by subsets of dense layer the output of the network, where the classification occurs. Here is where the high level logic operations occurs, having high computational demand due to high number of parameters hence high complexity, a *dropout* layer are added previously to the dense layer, inactivating some neurons (weights) reducing overall complexity avoiding over fitting.

Figure 8 – pooling most used methods

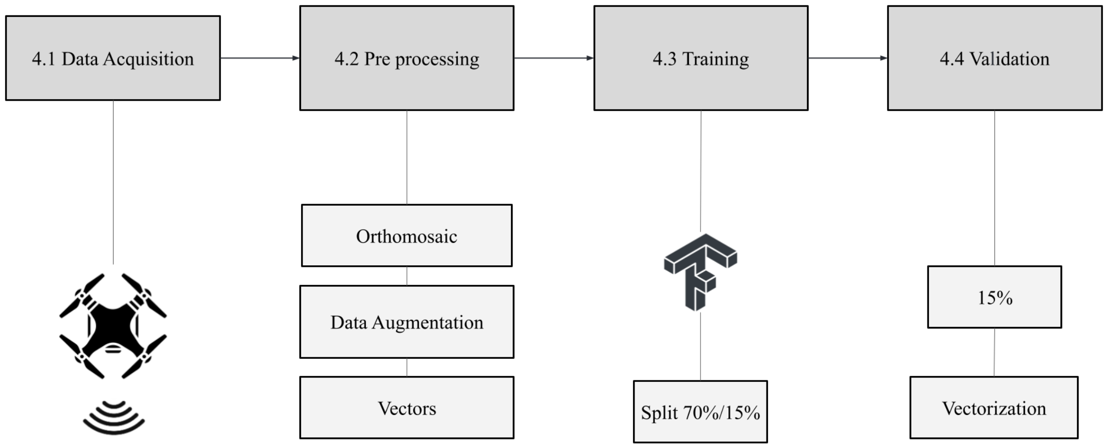


Source: created by author

4 Proposed Approach

The proposed approach is summarized in the flow chart presented in the Figure 9. In 4.1 the data acquisition procedure is described, where an UAV device is used to perform data acquisition. In 4.2, a preprocessing step is applied where the source orthomosaic are tilled and augmented. The labeling procedure is also performed in this step, where the region of interest is marked in order to train the computational model. Using the labeled vectors, the training step takes place as shown in 4.3. The training is performed using two convolutional neural networks architecture. The final step of the proposed approach is the validation procedure shown in 9 and explained in the Section 5.

Figure 9 – General Flow



Source: created by author

4.1 Data Acquisition

Data acquisition occurred over 4 distinct regions described below at table 3 performed by DJI Phanton 3 UAV equipped with 12.4 MP RGB sensor (4000x3000) , following flight plan created by flight professional using Drone deploy software ensuring 75%/75% overlaping and Lawnmower pattern for optimal image quality and minimal terrain distortions ensuring a 60m flight height. Sensors and flight height were normalized to the best relationship related to time, cost and battery viability, ending at an optimized value at 3cm/pixel. From a set of overlapping images with the corresponding referencing information, the orthomosaic generation workflow was performed by Agisoft Metashape PhotoScan (DUTTA *et al.*, 2021).

Local	State	Area(ha)
Guatapará	SP	32.5
Magalhães	SP	209.2
Santa Maria	SP	214.9
Porto Pinheiro	SP	93.3

Table 3 – Dataset Area

Source: created by author

4.2 Pre Processing and Data Augmentation

After data acquisition and georeferenced orthomosaic generation, the final result end up covering an total of 550 ha as presented above at Table 3. Before any step tilling procedure are needed to the employment of CNN approaches. Geospatial Data Abstraction Library (GDAL) was used for dataset transformations and for data augmentation, OpenCV module utilities performed the desired image replications for the dataset (WARMERDAM *et al.*, s.d.).

a) GDAL Retile: for CNN approach, the dataset must be resized for a more optimal size focusing at memory and performance balance. For resizing the dataset, *Gdal Retile* tilled the dataset, as shown at Figure 10, resizes the dataset to smaller sizes: eg.: Table 3 Guatapará region was translated from 1 image of 18982x51214 pixels to 969 images of 1024x1024 pixels.

b) GDAL Rasterize: having vectorized data files as source labels, the translation to ND Array structures are needed. *Gdal Rasterize* burns vector geometries (points, lines, and polygons) into the raster band(s) of a raster image. Vectors are read from OGR supported vector formats(Shapefile), and from these ones RGB images are created as input label, as it is the dataset are compose by pair of source RGB images and binary image label.

c) Data Augmentation: Observable at Table 13 shows the current not balanced dataset, for correction 8 data augmentation steps were applied, brightness(3 levels), blur (2 levels), rotation (3 levels) for the minority class of the dataset(Weed) multiplying its numbers for a more uniform class distribution.

Originating the following transformed dataset:

Local	Ortho	Chunks
Guatapará	1 x 18982 x 51214	1040 x 1024 x 1024
Magalhães	1 x 62977 x 71924	2464 x 1024 x 1024
Santa Maria	1 x 52831 x 64095	3368 x 1024 x 1024
Porto Pinheiro	1 x 45103 x 37353	1748 x 1024 x 1024
Total		**8620 x 1024 x 1024**

Table 4 – Dataset dimensions

Source: created by author

The dataset end up being composed by 2 classes: Weed (invader) and Non-Weed(anything that do not correspond to the invader eg.: soil, vegetation, cars, objects). At the figure 11 the dataset are shown, pair of images composed by source RBG and its corresponding labels, following one samples for each region, from up-down order being: Guaatapará-SP, Magalhães-SP, Porto Pinheiro-SP, Santa Maria-SP.

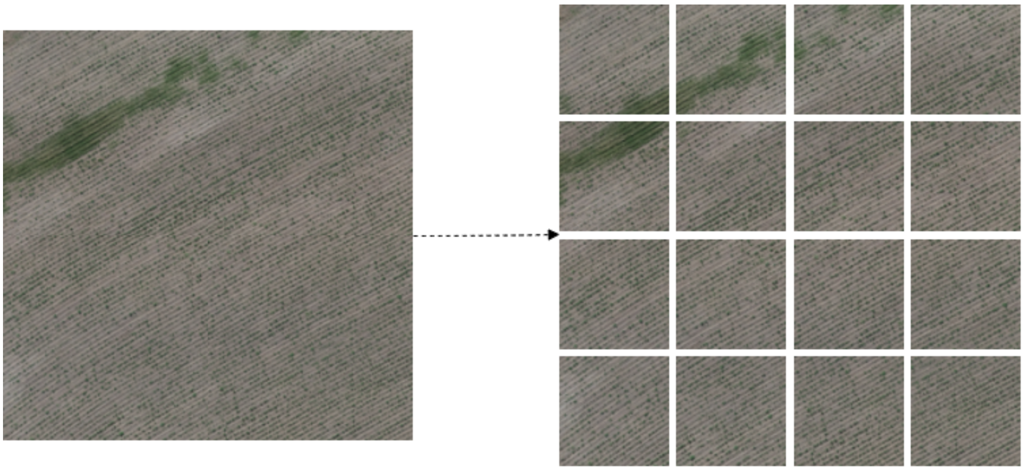


Figure 10 – Tilling algorithm
Source: created by author

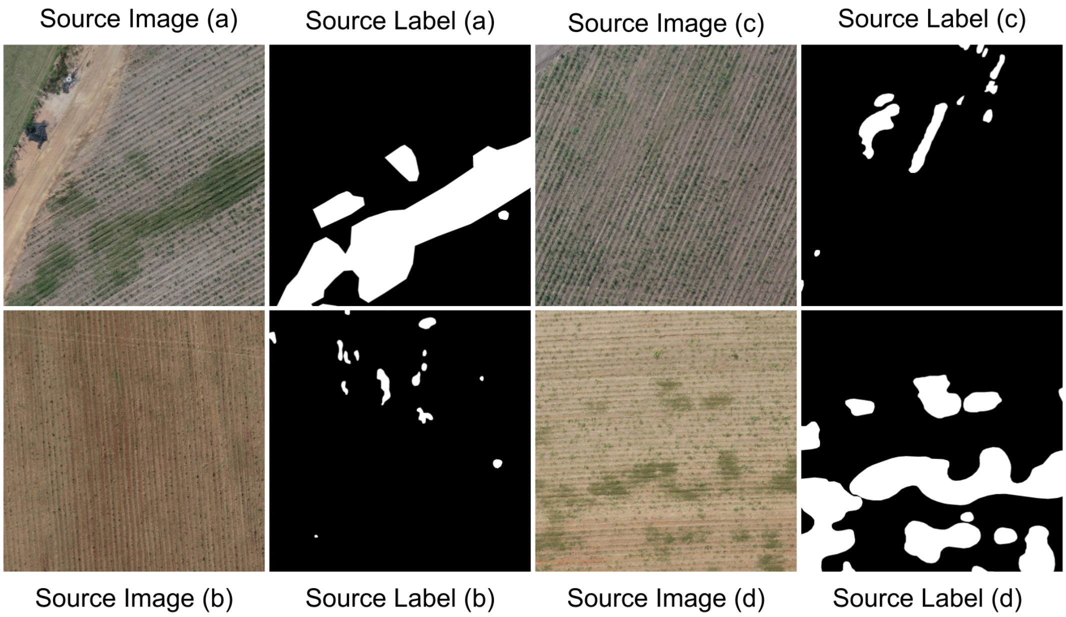


Figure 11 – Dataset Samples
Source: created by author

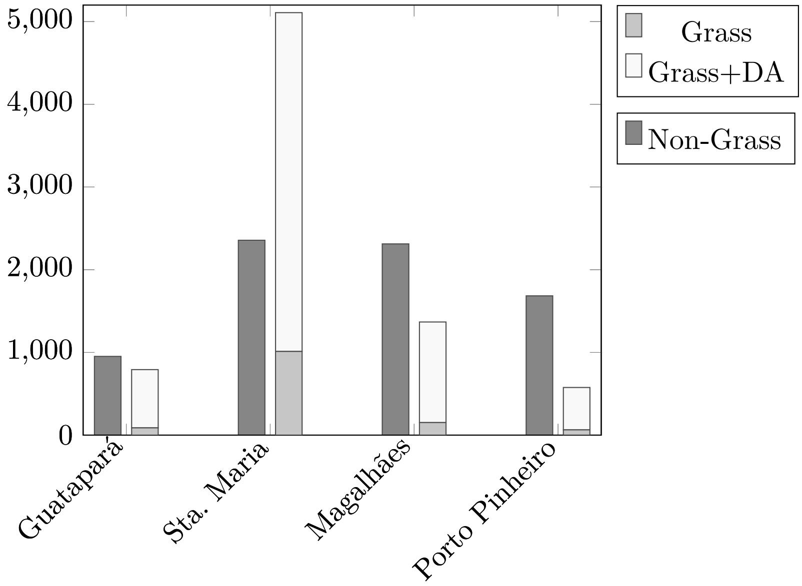


Figure 12 – Quantity of images by region
Source: created by author

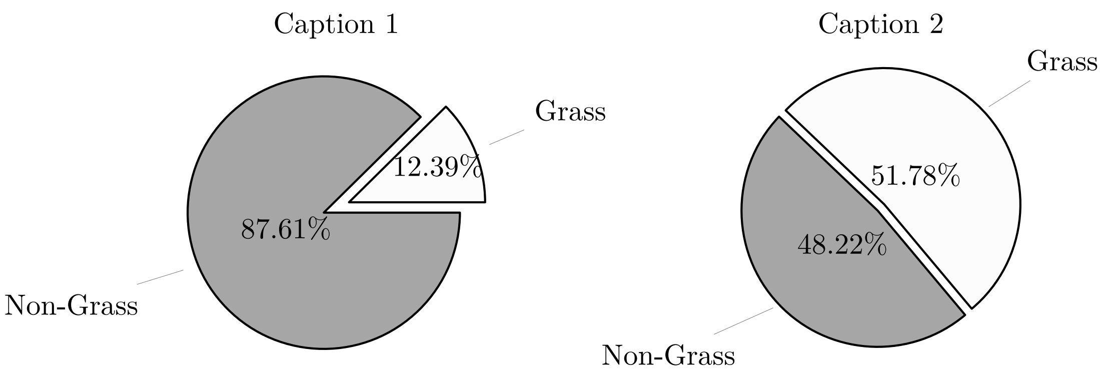


Figure 13 – Class Distribution
Source: created by author

4.3 Training and Recognize model Definition

The elected models for weed semantic segmentation from sugar cane crop images were, following literature most consolidated models.

a) UNet: created at Computer Science Department of the University of Freiburg, based upon fully connected neural network (FCN) family, successor from FCN where poolling steps are replace for upsampling operators. Defined as an encoder-decoder, it is characterized by it contract and expansive path, resulting at an U shape design as shown at its architecture at figure 14

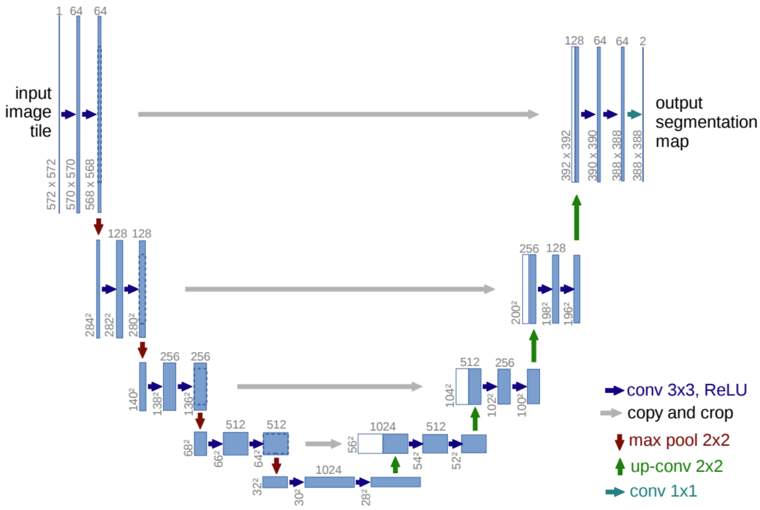


Figure 14 – UNet Architecture

Source: Ronneberger, P.Fischer e Brox (2015)

b) SegNet: defined as an encoder network, topologically identical to the 13 convolutional layers in the VGG16 network containing layers composed by Convolutions + Batch Normalization, and Activation ReLU functions. The decoder responsibility stays upon mapping the low resolution features maps to a full input resolution feature map as represented at figure 15

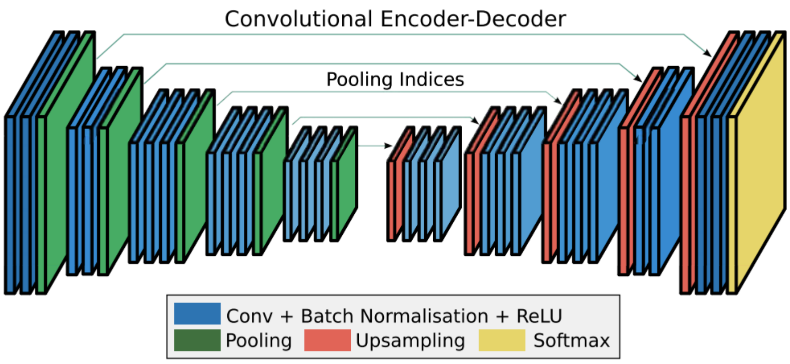


Figure 15 – SegNet Architecture

Source: Badrinarayanan, Kendall e Cipolla (2017)

After model definition and implementation, the Dataset were splitted in a 70/15/15 ratio, 70% for model training, 15% for validation step, and 15% for final testing purposes. All subsets were randomly selected following ratio commented above and also maintaining homogeneous classes distribution. At the training step iterations all over given dataset monitoring determined parameters for error minimization occurred, obtaining results of accuracy evaluation for proposed models discussed at the next section.

4.4 Validation and prediction

After training, with the desired response from the model, all desires metrics were calculated from confusion matrix from de 15% final dataset slice, shown at table 5. Finally, the output needs an extra step: translating the binary image to an vectorized file (shapefile), using Gdal for polygons creation and OpenCV module *findCountours* function for as the funciton mentions, find contours, the method end up resulting at an georeferenced shapefile output.

5 Experimental Results

Following training step, the proposed models presented accuracies up to 92%. Although this accuracy shows the overall quality of prediction results, it cannot completely evaluate the real performance of the models, more metrics are needed for a former discussion. *Precision* and *Recall* metrics are calculated from the Confusion Matrix presented at Table 5, *Precision* representing all correctly predicted results, and *Recall* all incorrectly predicted results. Finally, for a fair evaluation of *Precision* and *Recall*, *F1-Score* is denoted as:

$$F_1 = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{2}$$

The final overall quality of each method can be seen at tables 6. For each parameter, resolution, characteristic and method a private environment of test was created for the experiment.

Table 5 – Confusion Matrix

Weed	0.9220	0.0970	0.8720	0.0970
Non-Weed	0.0880	0.9130	0.2380	0.9130
	Weed	Non-Weed	Weed	Non-Weed
	SegNet		UNet	

Source: created by author

Table 6 – Model Results

Model	F1-Score	Precision	Recall	Accuracy
SegNet	0.9268	0.9420	0.9120	0.9280
Unet	0.8802	0.9180	0.8450	0.8850

Source: created by author

Table 7 – Model Comparison

Author	Model	Dataset Ratio	GSD	Accuracy	Precision
-	SegNet	1.00	3.00	0.92	0.942
-	UNet	1.00	3.00	0.88	0.918
Barrero *et al.* (2016)	*	0.37	1.84	0.99	NE
Huang *et al.* (2018)	FCN	0.10	0.30	0.92	NE
Buddha *et al.* (2019)	FR-CNN	0.03	0.46	0.93	NE
Kerkech, *et al.* (2020)	Segnet	1.58	1.00	0.93	NE
Ramirez *et al.* (2020)	DeepLabv3	0.47	0.01	0.89	NE
Hamylton *et al.* (2020)	LeNet	0.31	3.00	0.85	NE
Zou *et al.* (2021)	UNet	0.01	0.50	0.98	NE
Reedha *et al.* (2022)	ViT	0.11	0.50	0.98	NE

Source: Created by author

6 Conclusion and Discussions

This work proposed a solution for a weed mapping optimization problem, the two following models implementations presented great results achieving >92% precision metrics and successfully mapping infested areas, providing many savings for the current dataset owner and farmer.

Analyzing final results and comparing it to literature similar solutions, as detailed at table 7 the current dataset did not achieve comparable accuracy values, debatable justification could be mentioned as, all cited works shows lower ground to soil distance value hence having a better resolution it certainly contributed to the better results, an important point as comment earlier at this work, smaller GSD results in a total area coverage limitation.

Dataset difference at resolution and area coverage could result in a unfair results comparisons, as presented similar projects using the same models resulted and contrasting results, smaller train/validation set tends to result at a less generalization model. Applying the same related work proposed model to the current work dataset could result at an even lower accuracy results despite of lacking of pre processing data augmentation techniques.

Future remarks for the current proposed solution cloud be cited as, implementation of newer convolution neural network models as BEiT-3 Wenhui Wang *et al.* (2022) showing excellent quality and performance results at popular testing dataset over the literature. Another method becoming very popular, visual transformers (ViTs) are overcoming top notch CNN models, related works revels excellent result over differences dataset (COCO, ImageNet, ADE20K) (WANG, Wenhai *et al.*, 2022; CHEN, Z. *et al.*, 2022; REEDHA *et al.*, 2022).

References

ALBAWI, Saad; MOHAMMED, Tareq Abed; AL-ZAWI, Saad. Understanding of a convolutional neural network. In: IEEE. 2017 International Conference on Engineering and Technology (ICET). [S.l.: s.n.], 2017. P. 1–6.

ARAI, Kohei; KAPOOR, Supriya (Ed.). **Advances in Computer Vision**. [S.l.]: Springer International Publishing, 2020.

BADRINARAYANAN, Vijay; KENDALL, Alex; CIPOLLA, Roberto. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 39, n. 12, p. 2481–2495, 2017.

BARRERO, O.; ROJAS, D.; GONZALEZ, C.; PERDOMO, S. Weed detection in rice fields using aerial images and neural networks, p. 1–4, 2016.

BUDDHA, K.; NELSON, H. J.; ZERMAS, D.; PAPANIKOLOPOULOS, N. Weed Detection and Classification in High Altitude Aerial Images for Robot-Based Precision Agriculture, p. 280–285, 2019.

CANNY, John. A Computational Approach to Edge Detection. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, PAMI-8, n. 6, p. 679–698, 1986.

CASTRO, A. I. de; PEÑA, J. M.; TORRES-SÁNCHEZ, J.; JIMÉNEZ-BRENES, F.; LÓPEZ-GRANADOS, F. Mapping Cynodon dactylon in vineyards using UAV images for site-specific weed control. **Advances in Animal Biosciences**, Cambridge University Press, v. 8, n. 2, p. 267–271, 2017.

CHEN, Zhe; DUAN, Yuchen; WANG, Wenhai; HE, Junjun; LU, Tong; DAI, Jifeng; QIAO, Yu. Vision Transformer Adapter for Dense Predictions. **arXiv e-prints**, mai. 2022. arXiv: <2205.08534> `[cs.CV]`.

DJI. [S.l.]: DJI. Disponível em: <<https://www.dji.com/br/phantom-3-standard>>.

DUTTA; DUTTA, Partha; BARUAH, Arunjoy; KONWAR, Abhigyan; SEPAT, Vinod. A Technical Review of Lawn Mower Technology. **ADBU Journal of Engineering Technology (AJET)**, v. 4, p. 2016, ago. 2021.

GAO, Junfeng; LIAO, Wenzhi; NUYTTENS, David; LOOTENS, Peter; VANGEYTE, Jürgen; PIŽURICA, Aleksandra; HE, Yong; PIETERS, Jan G. Fusion of pixel and object-based features for weed mapping using unmanned aerial vehicle imagery. **International Journal of Applied Earth Observation and Geoinformation**, Elsevier BV, v. 67, p. 43–53, mai. 2018.

GARCIA-MARTINEZ, Hector; FLORES-MAGDALENO, Hector; ASCENCIO-HERNANDEZ, Roberto; KHALIL-GARDEZI, Abdul; TIJERINA-CHAVEZ, Leonardo; MANCILLA-VILLA, Oscar R; VAZQUEZ-PEÑA, Mario A. Corn grain yield estimation from vegetation indices, canopy cover, plant density, and a neural network using multispectral and RGB images acquired with unmanned aerial vehicles. **Agriculture**, MDPI, v. 10, n. 7, p. 277, 2020.

GASPAROVIC, Mateo; ZRINJSKI, Mladen; BARKOVIĆ, Đuro; RADOCAJ, Dorijan. An automatic method for weed mapping in oat fields based on UAV imagery. **Computers and Electronics in Agriculture**, v. 173, p. 105385, jun. 2020.

GIROLAMO-NETO, Cesare Di; SANCHES, Ieda Del'Arco; NEVES, Alana Kasahara; PRUDENTE, Victor Hugo Rohden; KÖRTING, Thales Sehn; PICOLI, Michelle Cristina Araujo; ARAGÃO, Luiz Eduardo Oliveira e Cruz de. Assessment of Texture Features for Bermudagrass (Cynodon dactylon) Detection in Sugarcane Plantations. **Drones**, v. 3, n. 2, 2019. ISSN 2504-446X.

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. **Deep Learning**. [S.l.]: MIT Press, 2016.

HAMYLTON, S.M.; MORRIS, R.H.; CARVALHO, R.C.; RODER, N.; BARLOW, P.; MILLS, K.; WANG, L. Evaluating techniques for mapping island vegetation from unmanned aerial vehicle (UAV) images: Pixel classification, visual interpretation and machine learning approaches. **International Journal of Applied Earth Observation and Geoinformation**, v. 89, p. 102085, 2020.

HAO, Xia; JIA, Jingdun; MATEEN KHATTAK, Abdul; ZHANG, Li; GUO, Xuchao; GAO, Wanlin; WANG, Minjuan. Growing period classification of Gynura bicolor DC using GL-CNN. **Computers and Electronics in Agriculture**, v. 174, p. 105497, 2020.

HASTIE, Trevor; TIBSHIRANI, Robert; FRIEDMAN, Jerome. **The elements of statistical learning: data mining, inference, and prediction**. [S.l.]: Springer Science & Business Media, 2009.

HUANG, Huasheng; DENG, Jizhong; LAN, Yubin; YANG, Aqing; DENG, Xiaoling; ZHANG, Lei. A fully convolutional network for weed mapping of unmanned aerial vehicle (UAV) imagery. Edição: Jose Luis Gonzalez-Andujar. **PLOS ONE**, Public Library of Science (PLoS), v. 13, n. 4, e0196302, abr. 2018.

ISPAG. **Precision Ag Definition | International Society of Precision Agriculture**. [S.l.: s.n.], 2019. Accessed: 2021-12-01.

JIANG, Honghua; ZHANG, Chuanyin; QIAO, Yongliang; ZHANG, Zhao; ZHANG, Wenjing; SONG, Changqing. CNN feature based graph convolutional network for weed and crop recognition in smart farming. **Computers and Electronics in Agriculture**, v. 174, p. 105450, 2020. ISSN 0168-1699.

JIMÉNEZ-BRENES, Francisco Manuel; LÓPEZ-GRANADOS, Francisca; TORRES-SÁNCHEZ, Jorge; PEÑA, José Manuel; RAMIREZ, Pilar; CASTILLEJO-GONZÁLEZ, Isabel Luisa; CASTRO, Ana Isabel de. Automatic UAV-based detection of Cynodon dactylon for site-specific vineyard management. Edição: Anil Shrestha. **PLOS ONE**, Public Library of Science (PLoS), v. 14, n. 6, e0218132, jun. 2019.

JOVANOVIC, Raka; TUBA, Milan; SIMIAN, Dana. Parallelization of the local threshold and Boolean function based edge detection algorithm using CUDA, p. 157–161, mai. 2012.

KERKECH, Mohamed; HAFIANE, Adel; CANALS, Raphael. Vine disease detection in UAV multispectral images using optimized image registration and deep learning segmentation approach. **Computers and Electronics in Agriculture**, Elsevier BV, v. 174, p. 105446, jul. 2020.

LIU, Li-Wei; MA, Xingmao; WANG, Yu-Min; LU, Chun-Tang; LIN, Wen-Shin. Using artificial intelligence algorithms to predict rice (Oryza sativa L.) growth rate for precision agriculture. **Computers and Electronics in Agriculture**, v. 187, p. 106286, 2021.

LOUDJANI, Philippe *et al.* Precision Agriculture: An Opportunity for EU-Farmers–Potential Support with the CAP 2014-2020. EPRS: European Parliamentary Research Service, 2014.

MACHADO, Penousal; ROMERO, Juan; NADAL, Marcos; SANTOS-DEL-RIEGO, Antonino; CORREIA, João; CARBALLAL, Adrián. Computerized measures of visual complexity. **Acta Psychologica**, v. 160, p. 43–57, set. 2015.

MCBRATNEY, Alex; WHELAN, Brett; ANCEV, Tihomir; BOUMA, Johan. Future Directions of Precision Agriculture. **Precision Agriculture**, Springer Science e Business Media LLC, v. 6, n. 1, p. 7–23, fev. 2005.

MOHRI, Mehryar; ROSTAMIZADEH, Afshin; TALWALKAR, Ameet. **Foundations of machine learning**. [S.l.]: MIT press, 2018.

PARVATHI, Subramanian; TAMIL SELVI, Sankar. Detection of maturity stages of coconuts in complex background using Faster R-CNN model. **Biosystems Engineering**, v. 202, p. 119–132, 2021. ISSN 1537-5110.

RAMIREZ, W.; ACHANCCARAY, P.; MENDOZA, L. F.; PACHECO, M. A. C. Deep Convolutional Neural Networks for Weed Detection in Agricultural Crops Using Optical Aerial Images, p. 133–137, 2020.

REEDHA, Reenul; DERICQUEBOURG, Eric; CANALS, R.; HAFIANE, Adel. Transformer Neural Network for Weed and Crop Classification of High Resolution UAV Images. **Remote Sensing**, v. 14, jan. 2022.

REYNS, P.; MISSOTTEN, B.; RAMON, H.; BAERDEMAEKER, J. De. A review of combine sensors for precision farming. **Precision Agriculture**, Springer Science e Business Media LLC, v. 3, n. 2, p. 169–182, 2002.

RONNEBERGER, O.; P.FISCHER; BROX, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Springer, v. 9351, p. 234–241, 2015.

SHALEV-SHWARTZ, Shai; BEN-DAVID, Shai. **Understanding machine learning: From theory to algorithms**. [S.l.]: Cambridge university press, 2014.

SOBEL, I.; FELDMAN, G. A 3×3 Isotropic Gradient Operator for Image Processing. **Presentation at Stanford A.I. Project 1968**, p. 271–272, 1968.

TANG, Qing *et al.* High-accuracy, high-resolution downwash flow field measurements of an unmanned helicopter for precision agriculture. **Computers and Electronics in Agriculture**, v. 173, p. 105390, 2020. ISSN 0168-1699.

TIMMERMANN, Clara; GERHARDS, Roland; KUHBAUCH, W. The Economic Impact of Site-Specific Weed Control. **Precision Agriculture**, Springer Science e Business Media LLC, v. 4, n. 3, p. 249–260, 2003.

WANG, Wenhai *et al.* **InternImage: Exploring Large-Scale Vision Foundation Models with Deformable Convolutions**. [S.l.]: arXiv, 2022.

WANG, Wenhui *et al.* **Image as a Foreign Language: BEiT Pretraining for All Vision and Vision-Language Tasks**. [S.l.]: arXiv, 2022.

WARMERDAM, Frank; BUTLER, Howard; SZEKERES, Tamas; ROUAULT, Even; RAHKONEN, Jukka; SCHWEHR, Kurt; BARKER, Norman; GILLIES, Sean; LOSKOT, Mateusz. **Gdal Rasterize**. [S.l.: s.n.]. Disponível em: <<https://gdal.org/programs/>>.

YAMASHITA, Rikiya; NISHIO, Mizuho; DO, Richard Kinh Gian; TOGASHI, Kaori. Convolutional neural networks: an overview and application in radiology. **Insights into imaging**, Springer, v. 9, n. 4, p. 611–629, 2018.

YAMASHITA, Rikiya; NISHIO, Mizuho; DO, Richard Kinh Gian; TOGASHI, Kaori. Convolutional neural networks: an overview and application in radiology. en. **Insights Imaging**, v. 9, n. 4, p. 611–629, ago. 2018.

YAO, Huang; QIN, Rongjun; CHEN, Xiaoyu. Unmanned Aerial Vehicle for Remote Sensing Applications—A Review. **Remote Sensing**, v. 11, n. 12, 2019. ISSN 2072-4292.

ZOU, Kunlin; CHEN, Xin; ZHANG, Fan; ZHOU, Hang; ZHANG, Chunlong. A Field Weed Density Evaluation Method Based on UAV Imaging and Modified U-Net. **Remote Sensing**, v. 13, n. 2, 2021.