This is to certify that

EDGAR STEVEN CORREA

Has presented a paper titled

GFkuts: a novel multispectral image segmentation method applied to precision agriculture
at the **Virtual Symposium in Plant Omics Science**

Held between November 23rd - 27th, 2020 at Santiago de Cali, Colombia

Andrés Jaramillo Botero
Scientific Director

Jaime Aguilar
Institutional Strengthening Sub-director

Support:

GFkuts: a novel multispectral image segmentation method applied to precision agriculture

Edgar S. Correa, Francisco Calderon*, Julian D. Colorado
Department of Electronics Engineering, Pontificia Universidad Javeriana, Cr 7 No. 40-62, Bogotá, Colombia.
*Corresponding author: calderonf@javeriana.edu.co

***Abstract*—Image segmentation enables the precise extraction of several crop traits from multispectral aerial imagery. This paper presents a novel segmentation technique called GFKuts. The method integrates a graph-based optimization algorithm with a k-means Monte Carlo approach. Here, we evaluate the performance of the proposed method against other approaches for image segmentation found in the specialized literature. Results report an improvement on the F1-score accuracy in terms of crop canopy segmentation. These findings are promising for the precise calculation of vegetative indices and other crop trait features.**

***Index Terms*—Multispectral imagery, image segmentation, precision agriculture.**

I. Introduction

Image segmentation is a needed compliance to process aerial multispectral imagery of a crop, in particular, for the extraction of relevant canopy features associated with physico-chemical variables that are highly related to plant light reflectance variations [1], [2], [3], [4]. The proper characterization of these changes allows for the identification of plant varieties with nitrogen deficiency [5] [6] [7], and other crop yield performance, such as the accumulated biomass [8].

Most works in precision agriculture use segmentation techniques based on image thresholding, clustering, edge detection or machine-learning approaches [11]. In this regard, segmentation algorithms can be divided in two: hard or soft, depending on the output mask. Hard segmentation algorithms create a binary output with only two levels for the background and foreground, while soft segmentation algorithms create a set of levels between those two. The approach of interactive segmentation tools such as Magic Wand, Intelligent Scissors, Graph Cut and Level Sets, were the object of study to develop the well-known GrabCut technique [11], [12]. Our proposed technique, namely GFKuts, is based on the former GrabCut method.

II. Methods

The GFKuts segmentation proposal consists of three stages. The first strategy is based on a binary classification with Montecarlo sampling on the image [13]. In this paper, it has been implemented through a k-means approach in order to generate two masks: a foreground and a background mask. The second strategy is an optimization stage based on the original GrabCut algorithm [11]. The third strategy consists of a refinement stage using a guided filtering [18]. The integration of the aforementioned methods enables the proposed GFKuts soft segmentation technique presented herein. GFKuts can operate in the standard RGB color space, with single channel images, or images of a custom composite channel. This feature is important since it is desirable to segment multispectral images.

A. Montecarlo sampled K-means

The Monte Carlo method is a stochastic numerical method based on random sampling, it has been commonly used to approximate complex mathematical expressions that are expensive to evaluate accurately. This random distribution is ideal for a clustering algorithm. k-means aims to partition a set of **N** observations into **K** groups, each observation belongs to the group whose mean value is closest. This algorithm has been used as an image segmentation technique [15] The main drawback of using K-means for binary segmentation is the indeterminacy of regions of maximum likelihood. K-means can be classified as a local algorithm and can be highly affected by underexposed or overexposed regions, shadows or noise.

The K-means binary classification strategy based on Montecarlo sampling selects pixels from the image using a uniform random distribution in two spatial axes. The result is a subset of sRGB values for each selected pixel. These values are related to the refractive bands of the spectrum of light captured by the sensor-camera. The way to implement it can be seen in algorithm 1, a random selection of pixels is made in the image according to the spatial position x, y. The classification characteristics of the subset of pixels I_n is denoted as $Feature$ and is based on the sRGB multispectral information. The method classifies the samples-characteristics into two groups, these clusters follow all the desired properties for an initialization trim (TB, TF). Each characteristic of the sample I_n is associated with the respective position I_x, y, x to generate foreground and background masks. The process of binary classification can be seen in algorithm 1.

Figure 1 shows the distribution of samples over the reconstruction of an RGN image (Red, Green, and IR).

Figure 2 shows the development of the classification stage, for $k = 2$ groups and $n = 245760$ samples. Each cluster represents one of the T_B and T_F masks.

B. GrabCut

Image segmentation arises as a discrete energy minimization problem. It consists in defining an energy function whose minimum provides the desired segmentation. Once the energy function is defined, an absolute minimum is found. GrabCut is an iterative and semi-manual method that proposes interactive

Algorithm 1 Montecarlo Sampled K-means.
The input is the image **I** and the number of samples **n**

for Each pixel in range (1 ... n) **do**
  Select a random pixel from **I** to P
  Store sRGB value from P to $\mathbf{Feature}_n$
  Store pixel coordinates $\mathbf{I}_{x,y}$
end for
Run a binary K-means over $Feature$ to get the labels $\mathbf{I0}_{n,0}$ and $\mathbf{I}_{n,1}$
if length of $(\mathbf{I}_{n,0})$ > length of $(\mathbf{I}_{n,1})$ **then**
  Create a mask T_F and set the coordinates in $\mathbf{I}_{x,y}$ of each pixel in $\mathbf{I}_{l,0}$ as the foreground (in our case, the canopy).
  Create a mask T_B and set the coordinates in $\mathbf{I}_{x,y}$ of each pixel in $\mathbf{I}_{l,1}$ as the background (in our case, the ground).
else
  Create a mask T_F and set the coordinates in $\mathbf{I}_{x,y}$ of each pixel in $\mathbf{I}_{l,1}$ as the foreground (in our case, the canopy).
  Create a mask T_B and set the coordinates in $\mathbf{I}_{x,y}$ of each pixel in $\mathbf{I}_{l,2}$ as the background (in our case, the ground).
end if

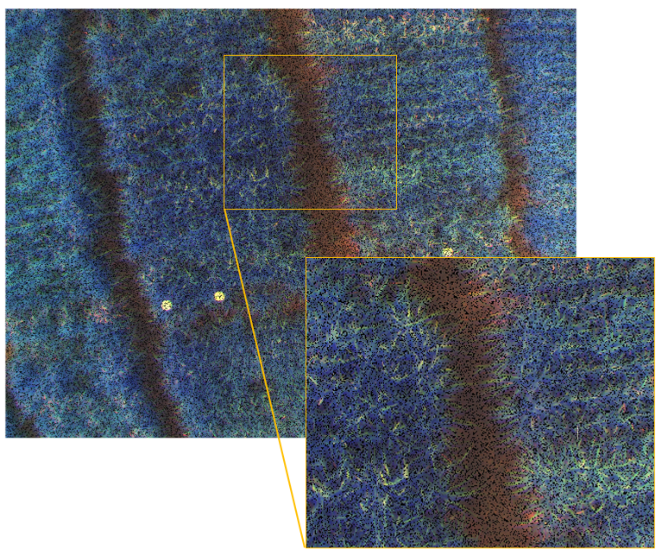

Fig. 1. A sampling of 20% of pixels on RGN image with size [960 * 1280].

foreground extraction. The interaction consists of dragging a rectangle around the desired object. Grabcut requires the creation of three image masks: a binary mask for the background T_B, a binary mask for the foreground T_F, and a final mask with uncertainty pixels T_U, which can be binary or have more levels quantification. The treatment of the image consists of taking pixels in the RGB space, it is denoted as I_n. Gaussian Mixture Model (GMM) is used, one for the background and one for the foreground, it is taken as a Gaussian mixture of complete covariance with K components. To treat MMG_s in a manageable way in the optimization framework, a vector $K = k_1, ..., k_n$, assigning each pixel a unique GMM component, according to α = 0 or 1 [12]. The optimization developed in GrabCut has two components: (i) $'U'$ evaluates the fit of the opacity distribution α_n and (ii) a smoothness function $'V'$. The optimization is based by the Gibbs energy function Eq. 1.

$$\mathbf{E}(\underline{\alpha}, \mathbf{k}, \underline{\theta}, \mathbf{I}) = U(\underline{\alpha}, \mathbf{k}, \underline{\theta}, \mathbf{z}) + V(\underline{\alpha}, \mathbf{z}). \tag{1}$$

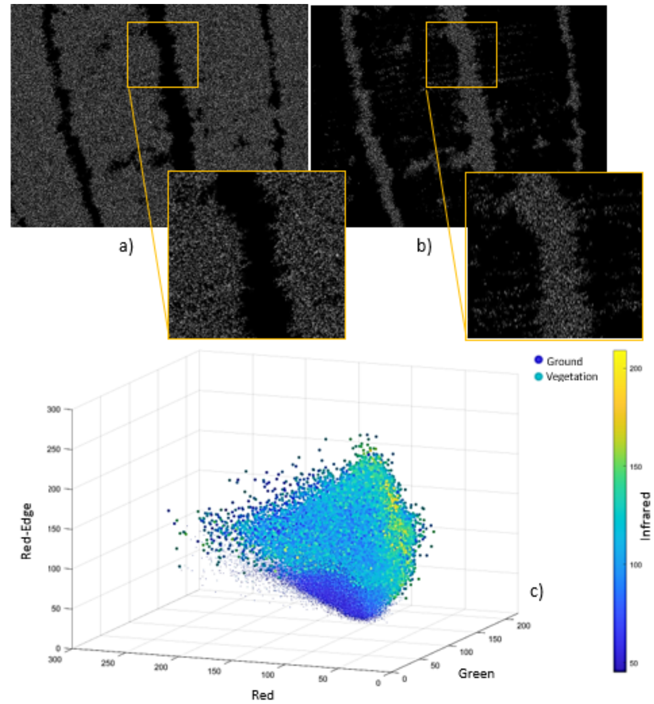

Fig. 2. a) Foreground mask, b) Background mask, c) Cluster Classification.

The term U defines the aptitude of the opacity distribution taking into account the GMM color models Eq. 2, of which p (·) is a Gaussian probability distribution and (·) are mixed weights.

$$U(\underline{\alpha}, \mathbf{k}, \underline{\theta}, \mathbf{I}) = \sum_{k=n} -logp(z_n|\alpha, k_n, \theta) - log\pi(\alpha, k_n) \tag{2}$$

GrabCut is a method that focuses on developing global optimization. The main advantage of GrabCut is Gaussian mix modeling and min-cut optimization, resulting in smooth image segmentation and fast-growing convergence.

In figure 3 the optimization result can be observed using masks a) and b) of figure 2. This result in a conventional way with GraCut would require generating the manual backgraund and foregraund masks, this is a laborious and non-repeatable task.

C. Guided Filter Refinement

Reduce noise and extract useful structures from images is used by Image filtering, for example, in image blurring/sharpening, edge detection, and feature extraction. [16].

LTI filter cores are usually used. In some cases is desirable to incorporate additional information from a particular orientation image during the filtering process. A useful approach consists of optimizing a quadratic function that directly imposes some restrictions on the unknown output. The solution is obtained by solving a dispersion matrix encoded with the information from the guide image. The output of each pixel is a weighted average of the nearby pixels, where the weights depend on the intensity/color similarities in the guide image

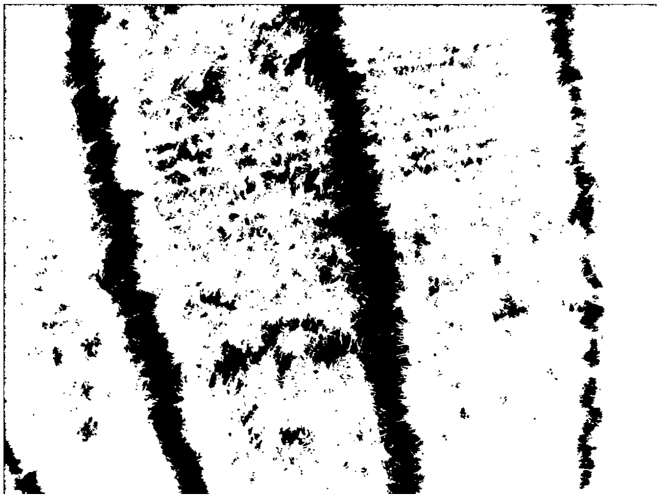

Fig. 3. Binary mask obtained from graph-based optimization.

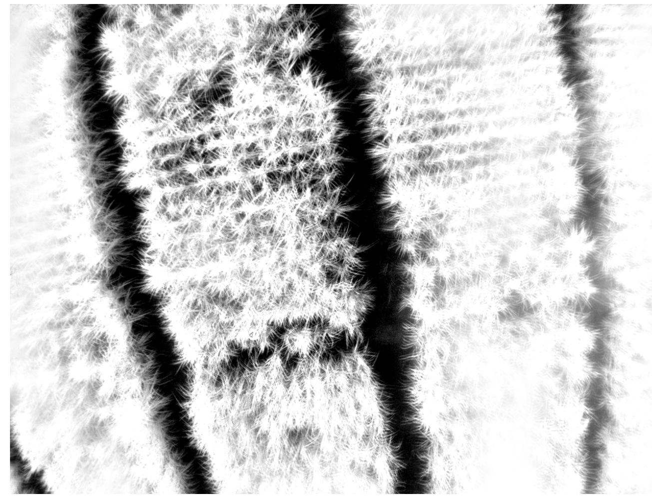

Fig. 4. GFKuts mask. Weight density mask takes values between 0 and 1.

[17]. This approach is initially developed in the bilateral filter and is the basis of the guided filter. This filter can smooth out small fluctuations and preserve the edges [18].

The guided filter approach is based on explicitly constructing of the cores using an orientation image. The output is a linear transform of the guide image. This filter has the edge-preserving and anti-aliasing property, like a bilateral filter, but does not suffer from gradient inversion problems [18].

The output of the filter is expressed as a weighted average 'q' having two inputs, a guide image 'I' and an input image 'p'. Both 'I' and 'p' can be identical. Where i and j are pixel indexes as seen in Eq.3

$$q_i = \sum_j W_{ij}(I) p_j \tag{3}$$

The filter kernel W_{ij} is a function of the guidance image 'I' and independent of 'p'. The kernel weights can be explicitly expressed by Eq. 4

$$W_{ij}^{GF}(I) = \frac{1}{|\omega|^2} \sum_{k;(i,j)\in\omega_k} \left(1 + \frac{(I_i - \mu_k)(I_j - \mu_k)}{\sigma_k^2 + \epsilon}\right) \tag{4}$$

The parameters μ_k and σ_k^2 are the mean and variance of w_k in image 'I' respectively, ϵ is a regularization parameter and $\|w\|$ is the number of pixels in w_k. The guided filter has an exact $O_{(N)}$ time algorithm (in the number of pixels N) [18].

Figure 4 shows the final result of the segmentation strategy. The method begins with the binary classification of a group of initial samples to develop a graph-based optimization and finally integrating a refinement stage.

D. GFKuts

The GFKuts algorithm is divided into three stages, (i) a stage of uniform random sampling and binary classification of some pixels in the image, (ii) an optimization stage based on graph cut, and (iii) a refinement stage based on guided filter GF. Finally and as a complement to the latter, is done a substage of binarization through an adaptive thresholding. The diagram in figure 5 shows the GFKuts segmentation strategy

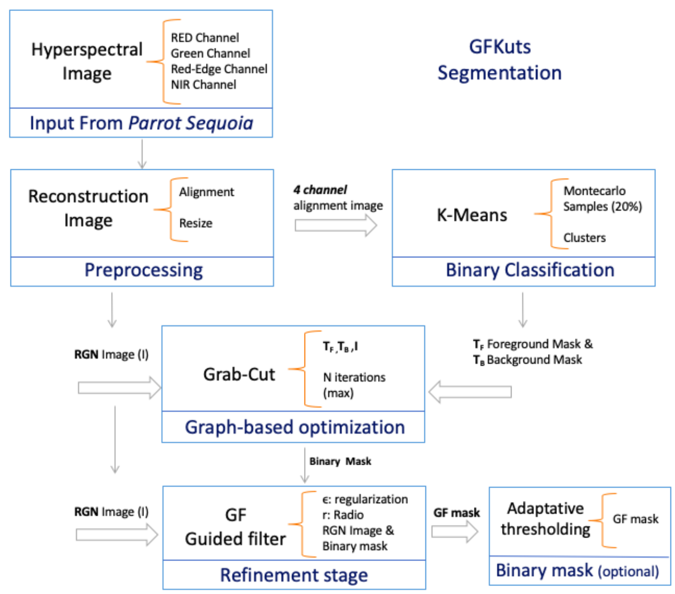

Fig. 5. Diagram of GFKuts methodological structure.

III. Results

Four segmentation strategies are implemented to evaluate and compare the performance of the presented segmentation strategy GFKuts. The evaluation of each technique is carried out using evaluation metrics, mainly F1-Score and Accuracy. Initially the techniques widely studied and validated in the literature, Threshold, k-means and Grab-Cut and finally the recently proposed technique GFKuts. To develop this evaluation, it is proposed to segment the image shown in figure 6 under the same conditions for each technique.

Algorithm 2 GFKuts,
I is the input image, N is the number of samples used in binary classification stage, in this case K-means, n is the number of iterations of GrabCut, r is the GF radius, ϵ is the regularization.

$\{T_B\ , T_F\} \leftarrow$ MontecarloSampledK-means (I, N)
while α converges or run n iterations **do**
  All pixels not set in T_B or T_F are set as a possible foreground pixels T_{UF}
  $\alpha \leftarrow$ GrabCut(I, T_B,T_F)
  Use the segmented image α as the new possible foreground pixels T_{UF}
end while
$\alpha_N \leftarrow$ GF $(Image = \alpha, Guidance = I)$
$\alpha_2 \leftarrow$ adaptiveBinaryThreshold(α_1)

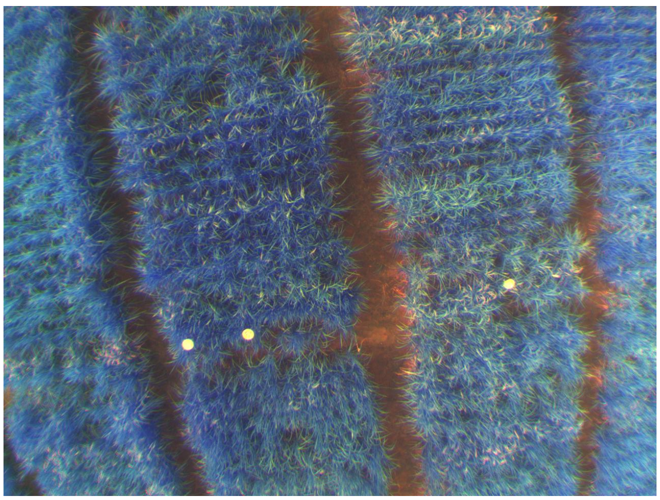

Fig. 6. Reconstructed and aligned RGN image from a parrot sequoia camera.

A. Threshold

The simplest strategy is to segment images through the threshold values of the histogram. In this case, the N(IR) channel threshold is used since most of the vegetation information is on this channel. The result is shown in figure 7.

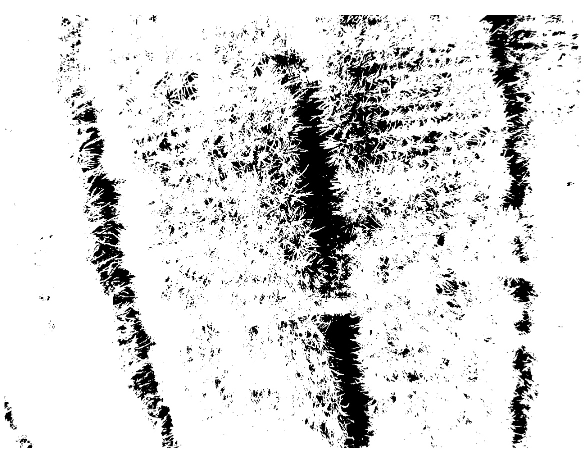

Fig. 7. Threshold segmentation strategy based mask.

B. GrabCut

In figure 1 the result of segmenting the RGN image with GrabCut is shown. It is important to highlight that the Tf Tb masks are developed manually, which does not guarantee repeatability due to the complexity of the image. Complexity related to the amount of detail of rice leaves, which are too many, long and thin.

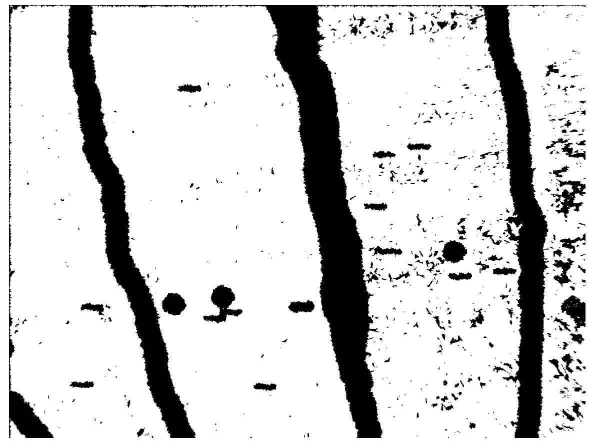

Fig. 8. GrabCut segmentation strategy based mask.

C. K-means

The k-mean segmentation is developed with a fraction of samples. The samples are selected with a uniform random distribution, the centroid and maximum distance of tow clusters are defined. With these two parameters, each pixel of the RGN image is evaluated, the result is a binary mask, a value ”1” represents the crop canopy and value ”0” represents the ground.The result is shown in figure 9.

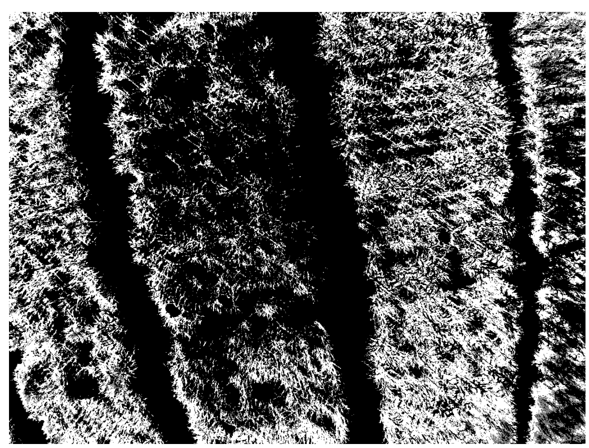

Fig. 9. Kmeans segmentation strategy based mask.

D. GFKuts

GFKuts presents an automatic approach as has been mentioned. The result is shown in figure 10, this segmentation is the result of points A, b, and C of section 2 whose diagram is presented in figure 5.

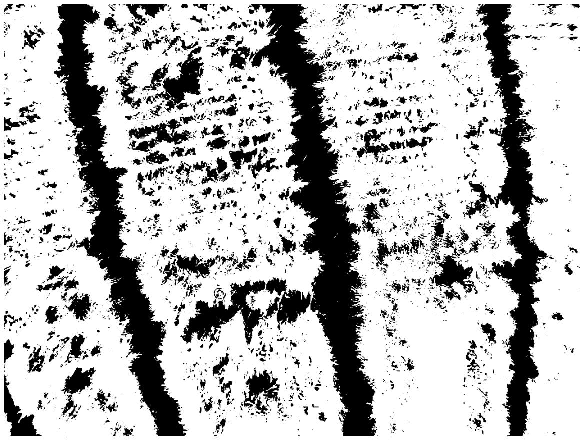

Fig. 10. GFKuts segmentation strategy based mask.

E. Metrics

Although the results can be subjectively intuited by observing the masks, it is important to present metrics that allow evaluating numerically the performance of each technique. in table I. The metrics presented are two representative (i) True positive rates (TPR) and (ii) a true negative rate (TNR). Additionally, three metrics widely used, (iii) F1-Score, (iv) accuracy (ACC), and (v) Matthews correlation coefficient (MCC) .

The first two related to vegetation values (TPR) and ground (TNR). Accuracy metric (ACC) assigns the same value to positives and negatives samples, while F1-Score privileges the positives, and the Matthews correlation coefficient (MCC) privileges the negatives.

Table I shows the five evaluation measures selected for the four segmentation techniques analyzed in this section.

TABLE I
IMAGE SEGMENTATION PERFORMANCE

	Threshold	**Kmeans**	**GrabCut**	**GFKuts**
True positive rate	0.91	0.34	0.88	0.81
True negative rate	0.55	0.92	0.66	**0.85**
Matthews coefficient	0.44	0.19	0.47	**0.52**
Accuracy (ACC)	0.86	0.42	0.85	0.82
F1-score	0.92	0.50	0.91	**0.90**
Standard deviation	0.155	0.207	0.239	0.1852

Most of the images used are characterized by having a higher content of pixels related to vegetation. Bearing in mind the imbalance in this equivalence allows evaluating the results, in table 1 the metrics that best represent the segmentation are highlighted in bold type. This is the key to understanding why the Threshold technique presents the best result in the True positive rate metric, but the overall result is the worst technique when evaluating with the Matthews coefficient metrics because it evens out the imbalance between samples by assigning greater weight to the samples that are in lower concentration in this case. It was also evidenced through the True negative rate metric.

IV. Conclusions

GFKuts can operate in (sRGB) the standard RGB color space, with single channel images, or images of a custom composite channel, this feature is importan since it is desirable to segment multispectral images.

It is usefull to use the samples selected by a binary classification algorithm like the kmeans shown in figure 2 to generate the initial convergence of the graph-based optimization. This binary classification presents better results when it has more information, in this case it uses the four channels of the parrot sequoia camera (Green, Red, Red-Edge, NIR)

Use of an independent four-channel multispectral camera requires preprocessing that involves reconstruction and alignment, otherwise the segmentation would have worse results.

Figure 11 shows a box plot with 400 samples of each technique. On each box, the central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually using the '+' symbol. The techniques present a large number of outliers due to the complex and varied nature of the images used. The best result is that of GFKuts with the characteristic of presenting an automatic focus in relation to GrabCut, although it also presents a good result, manual pixel selection of the image is required to obtain these results. K-means does not present favorable results due to light variations in the image. Finally threshold presents good results in the F1-Score metric associated with TPR, but not in the TNR metric, in contrast to K-means where TPR is low and TNR is high. GFKuts presents better results because, in addition to being automatic, it presents a good result in TPR and also in TNR.

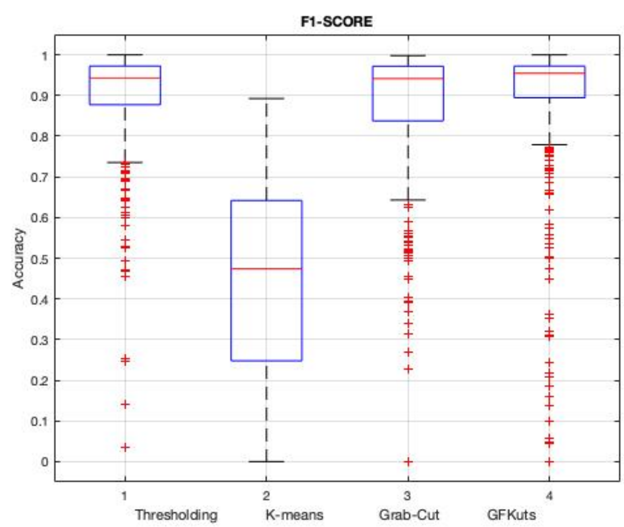

Fig. 11. F1-Score Metric in four segmentation techniques.

The standard deviation of the data plotted in figure 11 can be observed in the table II.

TABLE II
STANDARD DEVIATION

Threshold	**Kmeans**	**GrabCut**	**GFKuts**
0.155	0.207	0.239	0.1852

The motivation to develop image segmentation is based on generating useful information through Hyperspectral image processing. The characterization of the phenomic factors of different crop varieties, in an experimental way, allows training genomic selection models and evaluating the expression of traits of agronomic interest such as tolerance to variations in temperature and humidity, variations in radiation level, toxicity by aluminum in soils and biological attacks.

Quantifying the phenotypic characteristics through image segmentation is a tool that allows morphological modeling. This characterization is presented as valuable information in the task of validating the development of new agricultural varieties that allow greater productivity and food sustainability. For this purpose, segmentation allows estimating variables such as biomass and nitrogen, these variables have been presented as key variables to evaluate grain yield and the health status of the crop.

ACKNOWLEDGMENT

This work was funded by the OMICAS program: *Optimización Multiescala In-silico de Cultivos Agrícolas Sostenibles (Infraestructura y validación en Arroz y Caña de Azúcar)*, anchored at the Pontificia Universidad Javeriana in Cali and funded within the Colombian Scientific Ecosystem by The World Bank, the Colombian Ministry of Science, Technology and Innovation, the Colombian Ministry of Education and the Colombian Ministry of Industry and Turism, and ICETEX under GRANT ID: FP44842-217-2018 and OMICAS Award ID: 792-61187.

REFERENCES

1 N. R. Baker, "Chlorophyll fluorescence: A probe of photosynthesis in vivo," Annual Review of Plant Biology, vol. 59. Annual Reviews , pp. 89–113, Apr. 29, 2008, doi: 10.1146/annurev.arplant.59.032607.092759.

2 S. T. Ata-Ul-Karim, Q. Cao, Y. Zhu, L. Tang, M. I. A. Rehmani, and W. Cao, "Non-destructive assessment of plant nitrogen parameters using leaf chlorophyll measurements in rice," Front. Plant Sci., vol. 7, no. DECEMBER2016, Dec. 2016, doi: 10.3389/fpls.2016.01829.

3 Jimenez-Suancha, S. C., and BALAGUERA-LÓPEZ, H. E., "Fluorescence as an indicator of stress in Helianthus annuus L. A review", Revista Colombiana de Ciencias Hortícolas, 9(1), 149-160, 2015, doi: 10.17584/rcch.2015v9i1.3753.

4 A. V. Ruban et al., "Identification of a mechanism of photoprotective energy dissipation in higher plants," Nature, vol. 450, no. 7169, pp. 575–578, Nov. 2007, doi: 10.1038/nature06262.

5 Z. Chen and X. Wang, "Model for estimation of total nitrogen content in sandalwood leaves based on nonlinear mixed effects and dummy variables using multispectral images," Chemom. Intell. Lab. Syst., vol. 195, no. October, p. 103874, 2019, doi: 10.1016/j.chemolab.2019.103874.

6 Y. A. Djawad, H. Rehman, O. Jumadi, M. Tufail, S. Anwar, and N. Bourgougnon, "Discrimination of Nitrogen Concentration of Fertilized Corn with Extracted Algae and Polymer Based on Its Leaf Color Images," Ing. des Syst. d'Information, vol. 25, no. 3, pp. 303–309, 2020, doi: 10.18280/isi.250303.

7 D. F. Hoss, G. Lopes, C. R. Lajús, M. A. Moretto, and G. A. Tremea, "Multispectral aerial images for the evaluation of maize crops Imagens aéreas multiespectrais para avaliação da cultura do milho," 2020. doi: 10.1590/1413-7054202044004920.

8 Devia CA, Rojas JP, Petro E, Martinez C, Mondragon IF, Patino D, et al, High-Throughput Biomass Estimation in Rice Crops Using UAV Multispectral Imagery, Journal of Intelligent & Robotic Systems, 96(3-4):573 -589, 2019, doi: 10.1007/s10846-019-01001-5.

9 S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," Frontiers in Plant Science, vol. 7, no. September. 2016, doi: 10.3389/fpls.2016.01419.

10 Y. Xiong, L. Liang, L. Wang, J. She, and M. Wu, "Identification of cash crop diseases using automatic image segmentation algorithm and deep learning with expanded dataset," Comput. Electron. Agric., vol. 177, no. January, 2020, doi: 10.1016/j.compag.2020.105712.

11 S. Shirly and K. Ramesh, "Review on 2D and 3D MRI Image Segmentation Techniques," Curr. Med. Imaging Former. Curr. Med. Imaging Rev., vol. 15, no. 2, pp. 150–160, Jan. 2019, doi: 10.2174/1573405613666171123160609.

12 C. Rother, V. Kolmogorov, and A. Blake, "GrabCut - Interactive foreground extraction using iterated graph cuts," ACM SIGGRAPH 2004 Pap. SIGGRAPH 2004, vol. 1, no. 212, pp. 309–314, 2004, doi: 10.1145/1186562.1015720.

13 Caflisch RE., Monte Carlo and quasi-Monte Carlo methods, Acta Numerica,7:1-49, 1998, doi:10.1017/S0962492900002804.

14 He, K., Sun, J., and Tang, X., Guided image filtering, In European conference on computer vision (pp. 1-14). Springer, Berlin, Heidelberg,September 2010, doi:10.1007/978-3-642-15549-9_1

15 Siddiqui, F. U., and Isa, N. A. M., Enhanced moving K-means (EMKM) algorithm for image segmentation, IEEE Transactions on Consumer Electronics, 57(2), 833-841, 2011, doi:10.1109/TCE.2011.5955230

16 Gonzalez, R. C., and Woods, R. E., Digital Image Processing. 2nd edn Prentice Hall. New Jersey, 793, (2002).

17 Petschnigg, G., Szeliski, R., Agrawala, M., Cohen, M., Hoppe, H., and Toyama, K., Digital photography with flash and no-flash image pairs, ACM transactions on graphics (TOG), 23(3), 664-672, (2004).

18 He, K., Sun, J., Tang, X., Guided image filtering. In European conference on computer vision (pp. 1-14), 2010 September, Springer, Berlin Heidelberg.