Application for pixel-level segmentation and quantification of lignified fibers (sclerenchyma) and parenchyma in microscopic images of sugarcane culms

Gabriela Ferreira Ferrari Rubio[1], João Marcos Ferrari Rubio[1], Mariana Francato Logli[1], Lilian Elgalise Techio Pereira[1], João Paulo Rodrigues Marques[1], Adriano Rogério Bruno Tech[1]

[1]Faculty of Animal Science and Food Engineering (FZEA), University of São Paulo (USP)

Correspondence: gabriela.fr@usp.br

A paper from the Proceedings of the
17th International Conference on Precision Agriculture and 11th Brazilian Congress on Precision and Digital Agriculture
13-15 July 2026
Porto Alegre, Brazil

Abstract.

Quantifying lignified tissues in sugarcane culms (Saccharum spp.) is essential for anatomical characterizations and for inferences related to biomass quality and the potential uses of plant material. Conventional methods may require specific laboratory procedures and manual steps in digital analysis, increasing processing time and reliance on skilled operators. In this context, computer vision techniques applied to microscopic images constitute an accessible and reproducible alternative, with potential integration into digital agriculture workflows. The objective of this work was to develop and validate an application for the automatic detection and quantification of the area fraction occupied by lignified fibers (biomass with calorific value) and parenchyma (sucrose-containing cells), based on pixel counting in microscopic images of the sugarcane culm. Samples were obtained by freehand transverse sectioning of the mid-culm region of Cultivar CTC 4, followed by phloroglucinol staining within a three-minute reaction time. Sections were mounted on histological slides, and images were captured using a Belphotonics stereomicroscope at 2.5× magnification. In the images, lignified areas appeared as red pixels, necessitating conversion to a binary image (black and white) for improved visualization. Digital processing consisted of converting images from RGB to HSV color space, followed by binarization via segmentation based on the S (saturation) channel. Lignified regions, represented as white pixels in the binary image, were identified by automatic thresholding, with optional interactive user adjustment. For segmentation refinement, a morphological closing operation was applied, followed by filling regions enclosed by external contours. The application was developed in Python, using the Streamlit library to implement the interactive graphical interface. The system automatically computes the total number of pixels analyzed and the pixels classified as lignified fibers (white pixels) and parenchyma (black pixels), providing absolute and percentage values,

along with visualization of intermediate processing steps. Validation was performed by comparison with ImageJ software (n = 10 images), yielding a coefficient of determination (R^2 = 0.97). In addition to the high agreement between the methods, the developed application featured a fully automated pipeline, reducing the need of manual segmentation and measurement steps required in ImageJ, thereby simplifying analysis and broadening applicability for non-specialist users. We have concluded that S-channel-based segmentation in the HSV space enabled consistent identification of lignified regions, with an equivalent performance to the reference method, while offering greater operational ease. The application has shown to be a promise as a low-cost digital tool for the anatomical quantification of lignified and sucrose-accumulating regions in sugarcane culms, with prospects for integration into digital agriculture systems. Future studies should correlate the derived area metrics with laboratory indices of lignification and sucrose to expand quantitative validation and the system's applicability.

Keywords.
computer vision, image segmentation, lignification, Saccharum spp.

Introduction

Sugarcane (*Saccharum* spp.) occupies a strategic position in Brazilian agribusiness with production reaching a record 713.2 million ton in 2023/2024 (Companhia Nacional de Abastecimento 2024), consolidating Brazil as the world's largest producer of the crop. The strategic importance to national agribusiness relies on its economic relevance for industrial sectors, such as for sugar–ethanol production, and the extensive applications for animal production.

For livestock systems, sugarcane is an essential forage source, especially during dry periods when pasture availability declines. Despite the high biomass potential and energy density, sugarcane express nutritional limitations, particularly related to the high fibrous cell walls from stalks (stem). Then, the structural composition of the stalk becomes particularly relevant, influencing both industrial processing and biomass digestibility in animal feeding systems (Empresa Brasileira de Pesquisa Agropecuária 2022).

While the parenchymatous tissues are associated with sucrose accumulation and primary cell wall, the lignified tissues contribute to the stalk's fibrous structure with secondary cell wall (Marques et al. 2021). Lignin is a major structural component of the plant cell wall. Its distribution and abundance affect fiber quality and constitute relevant parameters in selecting plant materials for different purposes. From a nutritional perspective, higher lignin contents reduce fiber digestibility by hindering the action of ruminal microorganisms and limiting enzymatic access to cellulose and hemicellulose (Berchielli et al. 2011). For industrial applications, the ratio between lignified regions and regions associated with sucrose accumulation allow assessments of stalk quality.

Anatomical quantification of these tissues is traditionally performed using histological techniques and digital image analysis procedures that often require manual delineation of regions of interest (ROI) using ImageJ (Marques et al., 2018). While useful, such methods can increase processing time, reduce reproducibility, and demand greater operator intervention. In this context, computer vision emerges as a promising alternative to automate microscopic images processing and to generate quantitative metrics with greater speed and standardization.

An essential step in computer-vision pipelines that enables the identification and separation of regions of interest for subsequent quantitative analysis is called image segmentation. Several techniques have been applied to identify pixels associated with the ROI, although the accuracy depends on the image and object features. In histological images stained with phloroglucinol, lignified regions acquire a reddish coloration. This facilitates their distinction from other tissues and enables color-channel thresholding approaches.

However, the main challenge for color-channel thresholding lies in identifying the specific color components to isolate which, ideally, should maximize the contrast between the ROI and background pixels, producing a binary (black and white) image, ready for automated pixel

counting and quantification.

In these binarized images, pixel counting can be used to estimate the proportion of pixels occupied by regions of interest, making it a practical strategy for anatomical quantification (Barelli 2019; Santos 2014).

Accordingly, the objective of this study was to develop and validate an application for the automatic detection and pixel-count-based quantification of the area fraction occupied by lignified fibers and parenchyma in microscopic images of sugarcane stalks. The proposed approach aims to enhance analytical reproducibility, reduce manual steps, and provide an accessible digital tool to support anatomical studies and applications in digital agriculture. To validate the developed application, the segmentation and quantification results were compared with those obtained using ImageJ, a widely validated and extensively used software for digital image analysis, employing the same segmentation metrics.

Material and Methods

The study was conducted at the Laboratory of Technology and Information Systems (LTSI) using images provided by the Study Group on Applied Botany (GEBAP); both are affiliated with the Faculty of Animal Science and Food Engineering (FZEA) at the University of São Paulo (USP).

Figure 1 illustrates the complete computational pipeline developed for the anatomical quantification of lignified fibers and parenchymatous tissues from microscopic sugarcane images. The workflow was organized into four sequential stages, including image acquisition, preprocessing, segmentation, post processing and feature extraction.

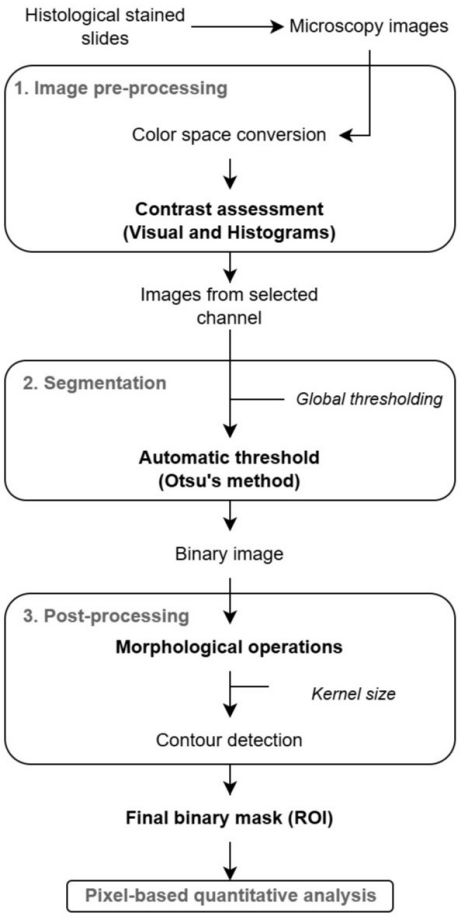

Figure 1- Workflow of the methodological pipeline for anatomical quantification of lignified fibers and parenchyma tissues in sugarcane images

Image Acquisition

Sugarcane stalks of the *Saccharum* spp. cultivar CTC 4 were collected from Sítio dos Ipês, located in Mogi Mirim, São Paulo, Brazil. Initially, the materials were selected and cleaned to remove surface impurities, followed by manual transverse sectioning with a stainless-steel blade to obtain freehand sections from the mid-stalk region.

The sections collected from the middle third of the sugarcane stalk were subjected to histochemical staining with a phloroglucinol solution, with a reaction time of approximately three minutes (Marques and Soares, 2021). This procedure imparts pink-red coloration to lignified tissues, enabling their differentiation from other cellular structures. After staining, the samples were mounted on microscope slides. A drop of distilled water was placed on each sample, and a coverslip was carefully applied to maintain structural integrity and avoid the formation of air bubbles.

The slides were examined under a Belphotonics stereomicroscope at 2.5× magnification. Illumination conditions were standardized throughout acquisition to minimize variations in intensity and hue among images. Image capture was performed using a digital system coupled to the microscope, ensuring adequate resolution and focus for subsequent computational processing.

Image preprocessing

Initially, the acquired microscopic images were imported into the computational environment for digital processing. During the preprocessing stage, the images were converted from the RGB color space to the HSV color space to allow independent evaluation of the H (hue), S (saturation), and V (value) channels. Subsequently, the RGB and HSV channels were visually compared regarding their ability to distinguish lignified tissues from parenchymatous regions.

Segmentation

After channel selection, image segmentation was performed through automatic binarization using Otsu's thresholding method. In the resulting binary images, lignified tissues were represented by white pixels, whereas parenchymatous tissues were represented by black pixels. To ensure methodological consistency during validation, the same thresholding parameters and morphological kernel sizes adopted in ImageJ were applied in the developed software.

Post processing

To improve segmentation quality and reduce small discontinuities, a morphological closing operation (MORPH_CLOSE) was applied. This procedure, consisting of dilation followed by erosion, promoted the connection of adjacent lignified regions and minimized noise in the segmented images. Subsequently, external contours of the identified regions were detected using the RETR_EXTERNAL function, which considers only the outer boundaries of the segmented structures. These contours were then filled to generate a continuous and consistent final binary mask suitable for quantitative analysis.

Feature extraction

Based on the final binary mask, feature extraction was performed through pixel-based quantitative analysis. The extracted metrics included: (i) the total number of pixels corresponding to the image area, (ii) the number of pixels classified as lignified fibers, and (iii) the number of pixels associated with parenchymatous tissues. From these values, the percentage area occupied by lignified fibers and parenchyma relative to the total image area was calculated, enabling objective, reproducible, and automated anatomical quantification.

Interface development

Finally, the graphical user interface (GUI) of the application was developed in Python using the Streamlit library within the Visual Studio Code (VS Code) environment. The computational pipeline integrated OpenCV for image-processing operations, NumPy for matrix manipulation, and Matplotlib for image visualization. The interface was designed to be interactive and user-friendly, allowing users to upload microscopic images and monitor all stages of processing in real time. In addition, adjustable controls were implemented for segmentation parameters, including manual threshold definition and morphological kernel size selection, providing greater flexibility under different experimental and image-acquisition conditions.

Application Validation

Validation was performed by comparing the quantitative results produced by the developed application with the values obtained in the ImageJ software, adopted as the reference method. Ten images were used for this comparison.

The degree of agreement between the methods was assessed using the coefficient of determination, computed in Excel, taking the ImageJ values as the reference and the application's values as the prediction. The metric was obtained as shown in Equation 1:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_{true,i} - \hat{y}_{true,i})^2}{\sum_{i=1}^{n}(y_{true,i} - \hat{y}_{true})^2} \quad (1)$$

Where:
$y_{true,i}$ = denotes the fiber pixel count measured in ImageJ;
$\hat{y}_{pred,I}$ = the count provided by the application
n = the number of images analyzed

Moreover, the agreement between the methods was also examined using the Mean Absolute Percentage Error (MAPE), which was likewise computed in Python with the support of the scikit-learn and NumPy libraries. This metric expresses the average percentage difference between the values provided by the application and the reference values obtained from ImageJ, allowing the relative accuracy of the developed method to be assessed. The MAPE was calculated as presented in Equation 2.

$$MAPE = \frac{100}{n}\sum \left| \frac{y_{true,i} - \hat{y}_{true,i}}{y_{true,i}} \right| \quad (2)$$

Where:
$y_{true,i}$ = denotes the fiber pixel count measured in ImageJ;
$\hat{y}_{pred,I}$ = the count provided by the application
n = the number of images analyzed

Results and Discussion

Digital color images are commonly represented in different color spaces, which are mathematical representations that describe each image pixel through numerical values. Among the most widely used color spaces are RGB (Red, Green, Blue), in which colors are represented by the combination of red, green, and blue intensities, and HSV (Hue, Saturation, Value), which separates color information into hue (dominant wavelength), saturation (color purity), and value (brightness). The individual evaluation of the channels within each color space is essential for achieving more efficient segmentation, as different color representations may enhance specific characteristics of the target structures (Barelli 2019).

In color thresholding, segmentation is performed by defining intensity ranges within a selected channel, allowing pixels with similar color properties to be isolated from the background or other tissue components. Since biological tissues and histological stains often exhibit distinct responses in different channels, the effectiveness of thresholding can vary substantially depending on the

chosen color space, thereby directly influencing the identification of edges and structural features within the regions of interest (ROI) (Bueno et al. 2008; Guan et al. 2014).

Color thresholding

In the present research, the stained slides show the typical spatial distribution and intensity of lignification within sugarcane internodes (Figure 2). The sclerenchyma fibers surrounding the vascular bundles retain the highest proportion of lignified walls (pink-red intense regions), acting as the primary sites of lignin deposition. The peripheral regions are characterized by an increased number of cell layers, which progressively decreases toward the central vascular bundles (Brienzo et al. 2016). This gradient reinforces the mechanical strength of the outer internode tissues, a pattern consistent with the structural demands of the stalk.

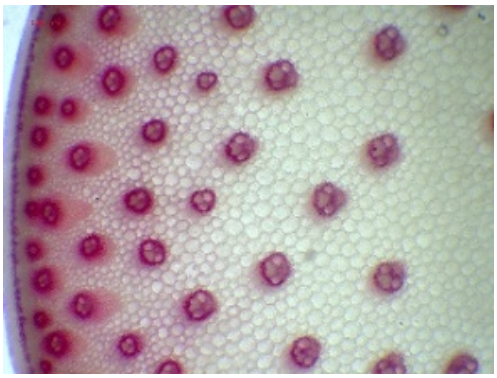

Figure 2. Original image of sugarcane (*Saccharum spp.* 'CTC 4') stalk cross-section (2.5× magnification) stained with a phloroglucinol solution.

Thus, for computer vision–based systems aimed at identifying biological structures in microscopic images, segmentation should be based on the most discriminative color information available in the image, since computers do not directly recognize anatomical structures such as fibers, vessels, or parenchyma. Therefore, representing the image in different color spaces enables the enhancement of specific contrasts, facilitating the separation and quantification of different tissues.

The RGB (Red, Green, and Blue) model is one of the most widely used color spaces in digital image processing, as it corresponds to how most acquisition systems (digital cameras, microscopes coupled to sensors, and monitors) record and represent image information. In this model, each pixel is represented by three channels: red, green, and blue. In 8-bit images, each channel typically takes values from 0 to 255, where values near 0 indicate low light intensity for that component, while values near 255 indicate high intensity. Thus, the final color of each pixel results from the combination of the intensities of the three channels (Barelli 2019).

RGB is classified as an additive model because colors are formed by the sum of red, green, and blue light in different proportions. When the three channels have low values, the pixel tends toward black; when they have high values, the pixel tends toward white. In histological images, this representation allows one to decompose the original image into three independent intensity matrices, making it possible to observe which channel shows the greatest contrast between the region of interest and the background.

The initial adoption of RGB in this study is justified by its simplicity, widespread use, and direct relationship with the original format of digital images. Guan et al. (2014) note that color images are generally stored in the RGB format and that many classical color spaces are obtained by linear or nonlinear transformations from this model. Therefore, beginning the analysis with RGB makes it possible to assess the chromatic information closest to the captured image before applying transformations to other color spaces.

In addition, RGB has low computational cost and straightforward interpretation. Each channel can

be analyzed separately, allowing verification of whether any of them enhances the structure of interest. Falcioni et al. (2022) describe that the Wiesner reagent, based on phloroglucinol–HCl, enables the identification of lignin through staining in different shades of red. Accordingly, one would expect the red channel to contribute more in regions of lignified fibers.

Moreover, Liu et al. (2007) used the RGB model to segment microscopic images of pathological cells. The authors found that the blue channel (B) presented greater contrast between cell nuclei and the image background, enabling the automatic segmentation of these structures. From the segmented regions, morphological and colorimetric features were extracted and used to classify normal and cancerous cells. This example shows that, despite being a simple model, RGB can be effective when one of its channels provides sufficient contrast between the object of interest and the background.

In the present study, because the lignified fibers were stained with phloroglucinol, these regions were expected to exhibit greater contrast in the R channel. However, although lignified fibers display a reddish hue after the Wiesner reaction (Falcioni et al., 2022), the optical density of the lignin–phloroglucinol complex results in light absorption that reduces the reflected intensity compared to the background. This apparent contradiction arises from the principles of image formation in brightfield microscopy (Murphy and Davidson, 2013) and from the additive RGB model, in which the parenchyma tends toward maximum saturation R,G,B≈255. This indicates that the R channel does not exclusively represent the “amount of red,” but rather a combination of chromaticity and luminance (Figure 3).

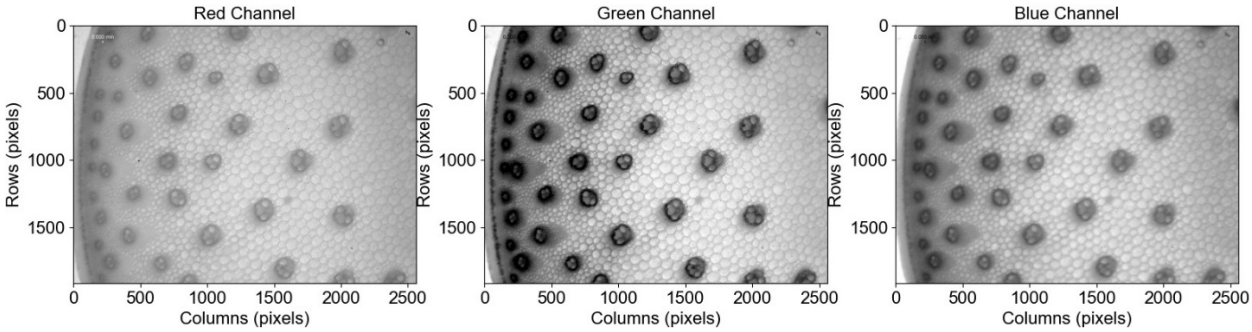

Figure 3. Decomposition of image of sugarcane (*Saccharum spp.* ‘CTC 4’) stalk cross-section (2.5× magnification) stained with a phloroglucinol solution in the individual red (R), green (G) and blue (B) channels.

Although the three channels preserved part of the anatomical organization of the stem cross-section, it is acknowledged that the RGB model has limitations for segmentation tasks, since color and luminance information are not independent (Gonzalez and Woods, 2018; Bueno et al., 2008). In this regard, Barelli (2019) cautions that selecting a channel for segmentation should not rely solely on subjective visual inspection, as human perception is optimized to detect edges and local contrasts, whereas thresholding algorithms operate on the global intensity distribution.

Indeed, the analysis of the histograms (Figure 4) demonstrates the shortcomings of the RGB model for this purpose. A substantial overlap is observed among the distributions of the three bands, with a high pixel density concentrated between 140 and 220. This indicates that structural information is tied to luminance, making the system sensitive to illumination variations and noise (Bueno et al., 2008). Moreover, saturation peaks at 255 (B channel) and low intensities near 0 (G channel) suggest overexposed areas and shadows, respectively, which can lead to classification errors (false positives and false negatives).

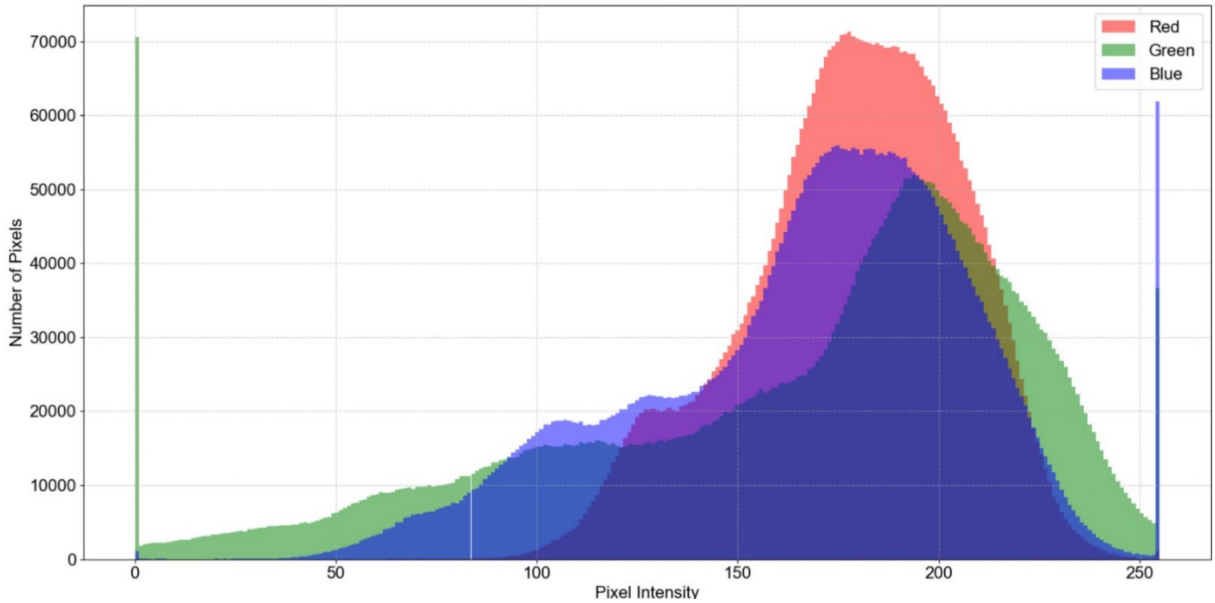

Figure 4. Histograms of red (R), green (G) and blue (B) channels of sugarcane (*Saccharum spp.* 'CTC 4') stalk cross-section (2.5× magnification) image slides stained with a phloroglucinol solution.

This numerical instability reinforces that, although RGB is the native acquisition format, this color space is insufficient to handle the anatomical complexity and the heterogeneity of staining in sugarcane stalks. Therefore, it becomes necessary to transition to color models that enable data decorrelation, such as the HSV model, with the aim of isolating the chromatic information from phloroglucinol staining (Hue and Saturation) from brightness (Value) fluctuations inherent to slide preparation.

Thus, the HSV model emerges as a robust alternative to mitigate the illumination-sensitivity limitations observed in the RGB model, especially in biological analyses where acquisition conditions may vary (Pedrini and Schwartz, 2008). The HSV space decomposes chromatic information into Hue (H), Saturation (S), and Value (V), allowing color to be analyzed independently of illumination (Gonzalez and Woods, 2018). This separation is essential for identifying the product of the reaction of phloroglucinol with lignin, which results in a reddish coloration with high chromatic purity.

When examining the decomposition of the image into HSV channels (Figure 5), each component responds differently to the anatomical elements of the stem. The H (Hue) channel exhibits a noisy response, hindering a clear delineation of fiber boundaries. The V (Value) channel, in turn, resembles a grayscale conversion of RGB, retaining illumination gradients that can interfere with binarization. By contrast, the S channel stands out visually by isolating the color purity present in the sample: lignified fibers appear at high intensity (bright pixels), whereas the parenchyma remains at low intensity (dark pixels), creating a natural binary contrast superior to the other channels.

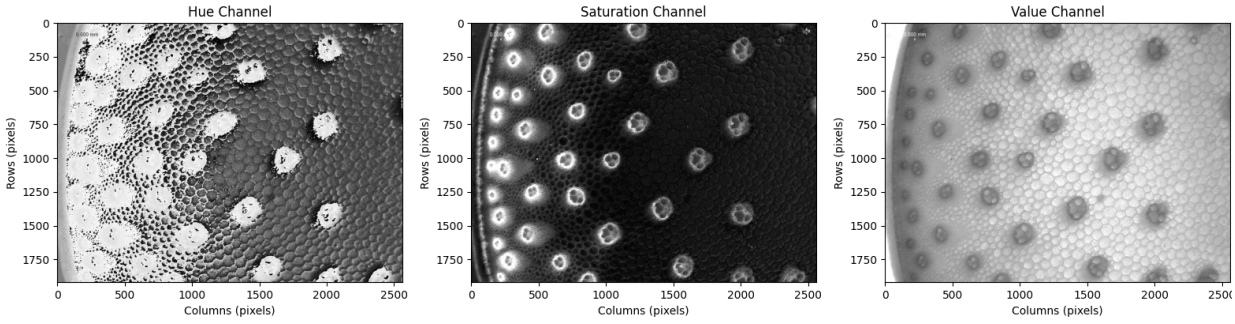

Figure 5. Decomposition of image of sugarcane (*Saccharum spp.* 'CTC 4') stalk cross-section (2.5× magnification) stained with a phloroglucinol solution in the individual hue (H), saturation (S) and value (V) channels.

The results presented in the HSV histogram (Figure 6) corroborate this visual observation. While the H and V channels exhibit more spread-out and overlapping distributions, the S channel shows a prominent peak at low intensities (between 0 and 50), corresponding to the vast area of low-saturation parenchyma. A second group of pixels, with intensity values above 120 and extending to near 200, accurately identifies the regions of lignified fibers. This clear separation, with a well-defined "valley" between peaks, is what allows automatic thresholding algorithms to operate with higher precision and a lower error rate.

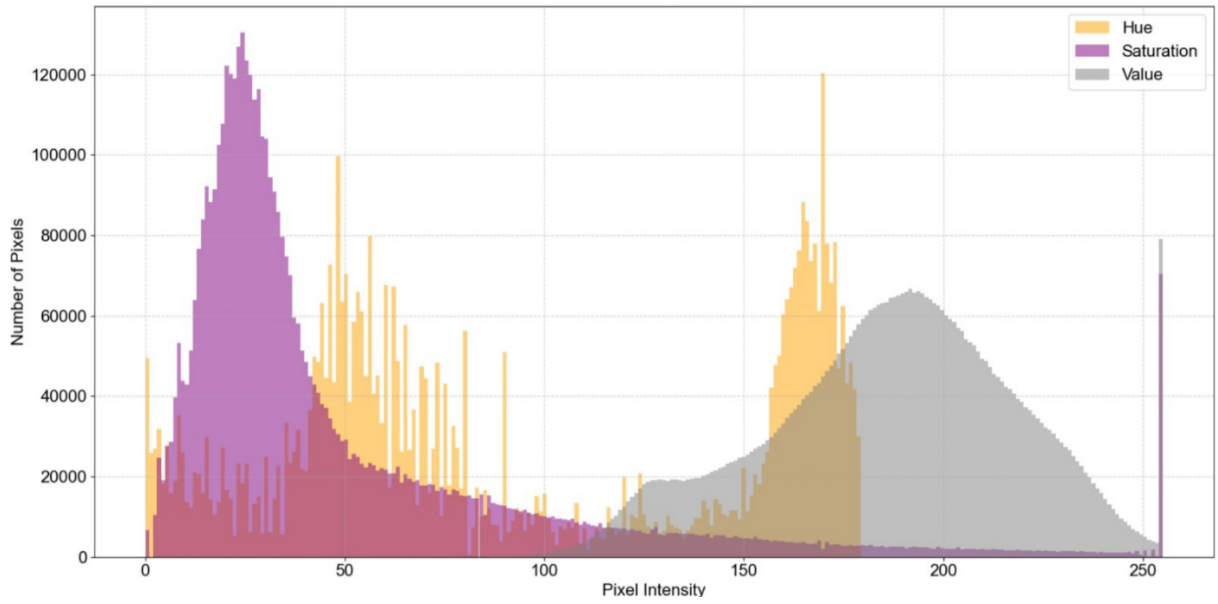

Figure 6. Histograms of hue (H), saturation (S) and value (V) channels of sugarcane (*Saccharum spp.* 'CTC 4') stalk cross-section (2.5× magnification) image slides stained with a phloroglucinol solution.

This trend of the superiority of saturation-based channels for segmenting-colored objects against neutral or light backgrounds is consistent with the findings of Bai et al. (2013), who identify saturation as the most stable parameter for isolating regions of interest in plant images under different lighting conditions. Unlike the V channel or the RGB bands, which are directly influenced by sensor exposure and the microscope's light source, the saturation component captures the vividness of the pigment present in lignified fibers.

Accordingly, it is concluded that the S channel is the most stable parameter for the automatic segmentation of fibers and parenchyma. Scientifically, the effectiveness of the S channel is explained by the fact that the phloroglucinol reaction not only shifts the tissue hue but also increases the chromatic purity (saturation) of the fibers relative to the parenchymal tissue, which remains low in saturation. Because the S channel does not vary with the intensity of white light (brightness), it provides a stable binarization threshold that isolates the chemical staining response, minimizing false positives in high-reflectance parenchyma regions and ensuring faithful quantification of the lignified area.

Performance assessment

Once defined as the most suitable channel for color thresholding, the performance of the proposed pipeline was compared with a standard method (ImageJ). For this stage of the research (calibration), images were individually analyzed using the standard method and then replicated using the same settings in the proposed application. This approach allowed for a precise evaluation of the possible impacts of different threshold and kernel sizes required to achieve high-fidelity segmentation before applying them to a larger dataset. Moreover, since the same spatial filters and binarization thresholds were replicated, a side-by-side comparison provides a baseline to assess the accuracy of the proposed pipeline against an established benchmark.

Table 1 presents all the data obtained and the parameters used for each sample. Among the total

pixels in the images, those identified as lignified fibers represented from 18.5 to 29.6% when analyzed with ImageJ and from 17.5 to 30.2% when using the proposed method. The relative differences between the refence method and segmented images using the S channel varied from -0.86 to -6.03%, with the proposed method resulting in slightly lower values in eight from the ten samples analyzed. For nine of the ten evaluated samples, the difference in lignified fiber proportion was lower than 1.5 percentage points, indicating that the proposed method can accurately quantify lignified tissues in sugarcane internode cross-sections.

Sample 8 represented an outlier, exhibiting the largest discrepancy observed in the dataset. Two main factors may have contributed to this result: image-specific characteristics and analysis settings that influenced segmentation performance. Visual inspection of the histological slide revealed the presence of air bubbles, which generated false edges and disrupted color homogeneity, contributing to the formation of highly noisy regions. Under these conditions, the ImageJ parameter settings, particularly the combination of a threshold value close to the decision boundary and a large kernel size, may have compromised the accuracy of the area fraction estimates.

Despite that, the developed application achieved a coefficient of determination (R^2) of 0.97 (Figure 7) with a mean absolute percentage error (MAPE) of 5.35% relative to ImageJ. Considering the inherent variability associated with boundary definition of lignified tissues, the magnitude of error can be regarded as low. Therefore, the proposed method demonstrated good agreement with the reference software while providing a more automated and user-friendly workflow.

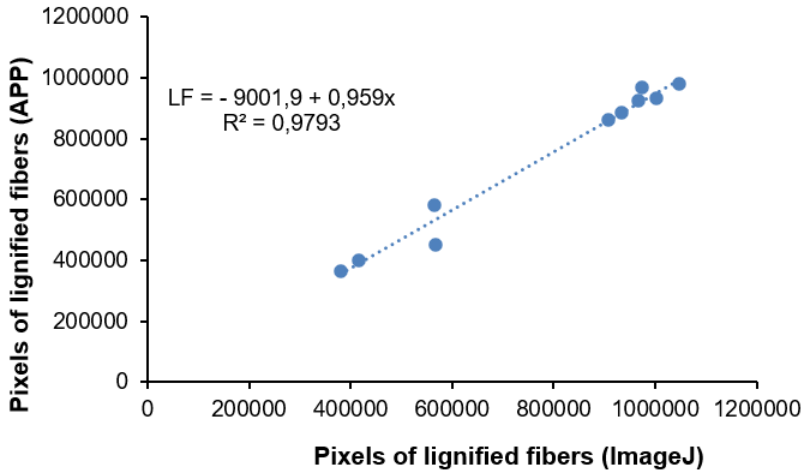

Figure 7. Scatter plot comparing lignified fiber pixel counts of sugarcane (*Saccharum spp.* 'CTC 4') stalk cross-section (2.5× magnification) image slides stained with a phloroglucinol solution obtained by the developed application (APP) and ImageJ. The blue line represents the ideal agreement (y = x).

Another important aspect of the proposed application is its usability. The Streamlit-based interface enables straightforward user interaction (including multiple image uploads, parameter adjustment, and immediate visualization of results). For the analysis of small number of samples, the application enables visualization of the intermediate processing steps, including the original image, histogram, binarization, segmentation, morphological cleanup, and final quantitative results. This organization aids users in interpreting the pipeline and facilitates adjustments when needed, particularly for images with variations in contrast or staining intensity.

This user-friendly design broadens its applicability to researchers and non-specialists in digital image processing, offering greater practicality and potential integration into digital agriculture workflows for automated anatomical analysis.

Table 1 - Pixel-count–based quantification from the developed application (APP) and ImageJ (reference method): thresholds, kernel, and absolute counts of parenchyma and lignified fibers.

Sample	Settings		Developed application (APP)		Reference method (ImageJ)		
n	Threshold	Kernel (pixels)	Parenchyma (pixels)	Lignified fibers (pixels)	Parenchyma (pixels)	Lignified fibers (pixels)	Total pixels
1	93	9	3989973	925227	3948338	966862	4915200
2	86	13	3980303	934897	3911856	1003344	4915200
3	105	19	4029806	885394	3980913	934287	4915200
4	96	9	4055051	860149	4006430	908770	4915200
5	94	15	3947347	967853	3941174	974026	4915200
6	92	15	3934792	980408	3867911	1047289	4915200
7	94	5	1554043	365957	1540300	379700	1920000
8	80	15	1467495	452505	1351629	568371	1920000
9	85	5	1518737	401263	1503276	416724	1920000
10	91	5	1339703	580297	1356275	563725	1920000

Conclusion

The developed application successfully enabled the automatic segmentation and quantification of lignified fibers and parenchymatous tissues in microscopic images of sugarcane stalks, using computer vision techniques. Among the color channels evaluated, the saturation channel (S) of the HSV color space provided the most effective identification between lignified and non-lignified tissues, thus allowing consistent threshold-based segmentation and minimizing the influence of lighting variations.

The proposed workflow, which integrates HSV-based thresholding, morphological post-processing, and quantification by pixel counting, demonstrated high agreement with the reference method used by ImageJ, reaching a coefficient of determination (R^2) of 0.97 and a mean absolute percentage error (MAPE) of 5.35%. These results indicate that the application can accurately estimate the fraction of the area occupied by lignified fibers, substantially reducing the manual intervention required in conventional image analysis procedures.

In addition to its quantitative performance, the Streamlit-based interface provides a practical and intuitive environment for image processing, facilitating the adoption of the method by researchers with little experience in digital image analysis. Therefore, the developed application represents a low-cost and accessible tool for the anatomical characterization of sugarcane stalks and has the potential for integration into digital agriculture workflows.

Future studies should expand the validation dataset to include different sugarcane cultivars, nutrition management, staining conditions, and image acquisition settings, as well as investigate correlations between image-derived metrics and laboratory measurements of lignin and sucrose content. Such efforts can further strengthen the applicability of the proposed approach for biomass quality assessment and decision support in agricultural and industrial systems.

Acknowledgments

We gratefully acknowledge the Laboratory of Technology and Information Systems (LTSI) and the Study Group on Applied Botany (GEBAP), both at the School of Animal Science and Food Engineering, University of São Paulo, for their support in developing this study. We also express our sincere gratitude to the Fundação de Estudos Agrários Luiz de Queiroz (FEALQ) for the financial support that made this research possible.

References

Bai, X., Cao, Z., Wang, Y., Li, Z., & Zhang, L. (2013). *Crop segmentation from images by morphology-modeling-based adaptive thresholding*. Computers and Electronics in Agriculture, 99, 21-31.

Barelli, F. C. (2019). *Introdução à visão computacional: uma abordagem prática com Python e OpenCV*. São Paulo, Brazil: Casa do Código.

Berchielli, T. T., Pires, A. V., & Oliveira, S. G. (2011). *Nutrição de ruminantes*. Jaboticabal, Brazil: Funep.

Brienzo, M., Abud, Y., Ferreira, S., Corrales, R. C. N. R., Ferreira-Leitão, V. S., De Souza, W., & Sant'Anna, C. (2016). Characterization of anatomy, lignin distribution, and response to pretreatments of sugarcane culm node and internode. *Industrial Crops and Products, 84*, 305–313. https://doi.org/10.1016/j.indcrop.2016.01.039

Bueno, G., González, R., Déniz, O., González, J., & García-Rojo, M. (2008). Colour model analysis for microscopic image processing. *Diagnostic Pathology, 3*(Suppl. 1), S18. https://doi.org/10.1186/1746-1596-3-S1-S18

Companhia Nacional de Abastecimento. (2024). *Acompanhamento da safra brasileira de cana-de-açúcar 2023/2024: 4º levantamento*. Brasília, Brazil: Conab.

Empresa Brasileira de Pesquisa Agropecuária. (2022). *Cana como forragem*. https://www.embrapa.br/agencia-de-informacao-tecnologica/cultivos/cana-de-acucar/producao/cana-como-forragem. Accessed 26 April 2026.

Falcioni, R., Moriwaki, T., Furlanetto, R. H., Nanni, M. R., & Antunes, W. C. (2022). Simple, fast and efficient methods for analysing the structural, ultrastructural and cellular components of the cell wall. *Plants, 11*(7), 995. https://doi.org/10.3390/plants11070995

Gonzalez, R. C., & Woods, R. E. (2018). *Digital image processing* (4th ed.). New York, NY, USA: Pearson.

Guan, T., Zhou, D., Xu, C., & Liu, Y. (2014). A novel RGB Fourier transform-based color space for optical microscopic image processing. Robotics and Biomimetics, 1, 16. https://doi.org/10.1186/s40638-014-0016-1

Liu, B., Yin, C., Liu, Z., Zhang, Z., Gao, J., Zhu, M., Gu, J., & Xu, K. (2007). Microscopic image analysis and recognition on pathological cells. In *Proceedings of the IEEE International Conference* (pp. 1022-1025).

Marques, J. P. R., and Soares, M. K. M. (2021). *Manual de técnicas aplicadas à histopatologia vegetal*. Piracicaba, Brazil: FEALQ.

Marques, J. P. R., Caetano, V. S., Pereira, F. J., Castro, E. M., & Pereira, M. P. (2021). *Sugarcane as a forage plant: structural and chemical traits that affect fiber quality.* BioResources, 16(4), 7623-7634.

Marques, J. P. R, Hoy J. W, Appezzato-da-Glória, B., Viveros, A. F. G, Vieira, M. L. C and Baisakh N (2018) Sugarcane Cell Wall-Associated Defense Responses to Infection by Sporisorium scitamineum. Front. Plant Sci. 9:698. doi: 10.3389/fpls.2018.00698

Murphy, D. B., & Davidson, M. W. (2013). *Fundamentals of light microscopy and electronic imaging (2nd ed.).* Hoboken, NJ, USA: John Wiley & Sons.

Pedrini, H., & Schwartz, W. R. (2008). *Análise de imagens digitais: princípios, algoritmos e aplicações.* São Paulo, Brazil: Thomson.

Santos, A. B. (2014). *Estudo da deposição de lignina em genótipos de cana-de-açúcar sob condição de déficit hídrico e baixa temperatura* (Doctoral dissertation, Universidade de São Paulo).