

Circular Hough Transform to improve viable *Saccharomyces cerevisiae* identification

Felipe Silveira Brito Borges ¹, Diogo Soares da Silva ², Lia Nara Balta Quinta ³, Arnaldo Ibanhe Mongelo ³, Marney Pascoli Cereda ⁴, Hemerson Pistori ^{2,3}

¹University of Debrecen (UNIDEB), Debrecen, Hungary

²Universidade Católica Dom Bosco (UCDB), Campo Grande, MS, Brazil

³Universidade Federal de Mato Grosso do Sul (UFMS), Campo Grande, MS, Brazil

⁴Universidade Estadual Paulista (UNESP), Botocatu, SP, Brazil

Abstract—Yeast counting is an important step in monitoring the fermentation process in sugarcane mills to optimize ethanol production. There is a need for a faster method to count viable cells in place of the fastidious and operator-dependent traditional method. In this paper, the application of a slightly modified version of the standard Circular Hough Transforms to automate the inoculated fermentation process of *Saccharomyces cerevisiae* is reported. The results of several experiments with different preprocessing algorithms and parameter adjustments are presented. The resulting system will be part of a microbiological control procedure that is being developed to respond to Brazilian ethanol sugarcane mill's demands.

Index Terms—Yeast detection. Hough transform. Machine Vision. Object Counting.

I. INTRODUCTION

The scarcity of fossil fuels and climate changes caused by the burning of these fuels is on the agenda of discussions worldwide and renewable energy sources, like bio-energy, are of increasing importance on the global stage [1]. Brazil has one of the cleanest energy matrices, sugar cane culture to produce alcohol representing 14% of the primary energy generated in the country [2].

The alcoholic fermentation is due to the action of the *Saccharomyces cerevisiae* yeast, which ferments sugars of sugarcane juice diluted with water for their growth and ethanol production. As industrial fermentation is done without asepsis, secondary contamination occurs by bacteria that consume sugar without producing ethanol. This contamination can hurt the yeast's performance. The number of viable yeasts is also essential to increase the yield and make the fermentation time shorter. Thereby controlling the number of viable yeasts is essential. In sugarcane mills, this control can be made by sampling and manual counting yeasts using a Neubauer chamber, which consists of a microscope slide divided into squares with an established volume, used to facilitate cell counting [3]. In this paper, some results on the automation of this time-consuming and tedious method, using computer vision techniques, are presented.

Corresponding author: Felipe Silveira Brito Borges (email: felipesilveira@ieee.org).

The method presented is part of a software that performs tests of the viability of yeast samples [3], performing the screening of microorganisms and calculating its area, which is found through the value of its radius, for the subsequent classification.

The main contribution is an extensive experimental investigation on the use of several preprocessing steps coupled with a Hough transform-based template matching strategy to automatically count the number of yeasts in microscopic images. In the preprocessing step, different smoothing and edge detection algorithms have been tested. A slightly modified version of the standard Circular Hough Transform has been used to detect yeast cells. It has been also found that, for this problem, the enhancement of the color saturation (in HSB space), before grayscale conversion, improves the efficiency of the Hough Transform. The results of the proposed method are compared, using precision, recall, and f-score metrics, to ground-truth images where yeast cells have been manually counted by a specialist.

In the next section, related works are presented and the proposed approach is discussed in Section III. The experimental setup, results and analysis are exposed in Section IV, V and VI. The last section presents conclusions and future works.

II. RELATED WORK

Milka [4] used a newly synthesized DNA fluorescent along with an Easycounter YC instrument to determine the viability of *Saccharomyces carlsbergensis* yeast cells in beer production. The counting is done after the samples incubation, which is loaded into the apparatus chip, then imaged and counted using a cytometer. The procedure achieved theoretical viability R^2 of 0.9988, which is a much better result in comparison with the manual method using methylene blue, however, it is still a manual method that requires a specific machine, the Easycounter.

In contrast to manual yeast counting procedures, Nicholas [5] applied an automated image analysis using the ImageJ software. This work did not obtain a statistically higher counting result if compared with the manual counts at 95% of confidence. On the other hand, it requires much less time

since as high as 400 Colony Forming Unit (CFUs) quickly counting was done, abruptly reducing the analysis time.

One new example of the Hough transform is the recognition of angle aided circles and the welding spot detection on the automobile body. The method proposed in [6] utilized a novel algorithm to reduce the computational complexity of the traditional Randomized Hough transform, and reached more than 98% accuracy on a dataset of 37 welding spot test images.

Hough transforms are also being applied to the problem of fruit detection. Guichao et. al. [7] proposed a novel technique to detect fruit using partial shape matching with the Hough transform. The authors used 7 different datasets, each one respective to a different fruit, containing no balanced images. The method reported achieving, for citrus, tomato, pumpkin, bitter gourd, towel gourd, and mango, the precision of 0.783, 0.848, 0.745, 0.762, 0.807, and 0.919, respectively.

The literature also presents a work about the Hough transform voting process to detect multiple circles to avoid false positives even with a large number of circles [8]. A set of 17 coins of varying sizes and colors were used, and the standard approach was used to be compared. The number of false positives is impressive 99% lower on the presented work.

Another interesting example was developed by Halil et. al. [9], where their studies were directed to the Optic Disk Localization detection. The team used the Hough Transform to detect optic disk (OD) in retinal images for diagnosis diseases, such as diabetic retinopathy, papilledema, and glaucoma. According to the authors, the methods were able to detect with 100%, 96.92%, and 98.88% accuracy for the DRIVE, DIARETDB0, and DIARETDB1 public datasets, respectively.

Automation in the radiography field was also the subject of the work of Lena et. al. [10], who developed approaches on 13 clinical radiography images. The manual results achieved 96.46%, which was almost surpassed by the 96.24% HT with edge detection.

None of the interesting techniques above mentioned can be used to find the viability of yeast by color since the first two examples are not targeted for detection and classification of the yeast *Saccharomyces cerevisiae*, and the third example above, can only be used in the counting of yeast grown in a Petri dish [11].

III. PROPOSED METHODS

The Hough Transform is a very well-known template matching strategy [12], [13] highly dependent on some preprocessing stages, like smoothing and edge detection. To find the best combination of smoothing and edge detection algorithms and also the best parameters values for the specific problem handled in this paper, several experiments have been conducted and led to the choice of a Gaussian smoothing filter combined with a Canny Edge detector.

Using context information, the usual sizes, and agglomeration properties of yeast cells, the standard Circular Hough Transform has been slightly adapted to work with ranges of circle radii and to discard circles whose Euclidean centers fall below a parameter's value. The yeast cells

present a bi-modal radius distribution and working with two ranges of radii during the calculation of the 3D Hough space helped to lower the processing time without compromising the detection performance. Another preprocessing step proposed in this work was to enhance image color saturation in HSB space before grayscale conversion. Hence, the steps of the proposed method are:

- Enhance Color Saturation
- Smooth using a Gaussian Filter
- Convert to grayscale
- Detect edges using the Canny Filter
- Identify yeast cells using the modified Circular Hough Transform

IV. EXPERIMENTAL SETUP

The experiments were performed on a set of 144 images captured in laboratory through the optical microscope Eclipse, from Nikon, and digitized using the CB25 CCD camera.

The yeast cell counting is made with the help of a Neubauer chamber, which presents a grid as seen in Figure 1 (a, b and c) that is used to guide the technician during the counting process. Also, there is a methylene pigment process that causes a blue color on nonviable cells, which results in its differentiation from viable cells.

The chamber used in this experiment, presented in Figure 1(a), is organized in outer and inner squares. There are 20 outer squares. Figure 1(b) presents a region of the chamber containing 4 complete inner squares. Each outer square is formed by 16 inner small squares, organized in a 4 x 4 grid. A detailed view of 4 inner squares (2x2), which were used as the input images for the algorithms tested in this paper is presented.

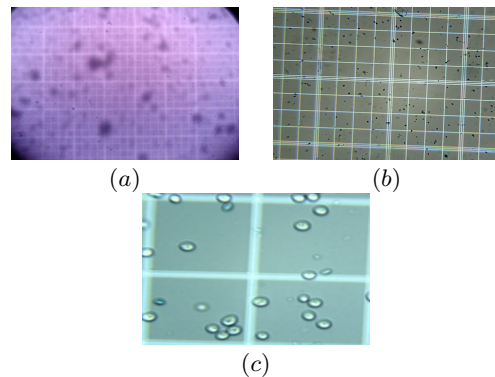


Fig. 1. Neubauer chamber and grids used for cell counting: (a) Neubauer Chamber. (b) Four complete outer squares. (c) Example of the 2x2 inner squares images used in the experiments. In addition, a blue cell in the 1st square, which represents a nonviable yeast.

From each 144 2x2 inner squares images, the number and position of yeast cells, identified by a technician, were available. These 144 images are grouped into 6 datasets, representing images captured in different fermentation moments. The names of the datasets refer to the Brix (Bn), which is related to the number of soluble solids in a fruit juice [14].

Table I describes each of the six datasets used in this work. Note that not all outer and inner squares are used in the manual counting process. The exact number and position of the squares that are counted are part of a standard laboratory procedure used to statistically infer some quality parameters of the fermentation process from the number of viable and non-viable yeast cells. Therefore, the number of viable cells per liter is counted throughout a formula, which subsequently allows determining the fermentation process period.

TABLE I
DESCRIPTION OF THE 6 DATASETS USED IN THE EXPERIMENTS

Dataset	Description
R1-B3	first repetition with Brix 3
R1-B6	first repetition with Brix 6
R2-B3	second repetition with Brix 3
R2-B6	second repetition with Brix 6
R3-B3	third repetition with Brix 3
R3-B6	third repetition with Brix 6

Four experiments are reported in this paper. The first one intended to find the best values for the Canny edge detector. The second experiment determined the threshold that should be used by the Hough transform algorithm to find the peaks in the Hough space (accumulator) that correspond to actual circles in the original image. The third evaluates a smoothing parameter. Finally, the last experiment was conducted to find the best set of parameters for the preprocessing, segmentation, and smoothing, along with some changes in the saturation of images to highlight the edges.

To evaluate the performance of the proposed method, three metrics were used: Recall, Precision, and the F-score [15]. The metrics are presented in equations 1, 2 and 3, where hy is the number of yeast cells that have been correctly detected as a circle by the Hough transform, ty is the total number of yeast cells present in the ground truth image and hc is the number of circles detected by the Hough transform.

$$recall = \frac{hy}{ty} \quad (1)$$

$$precision = \frac{hy}{hc} \quad (2)$$

$$F = 2 * \frac{precision * recall}{precision + recall} \quad (3)$$

All metrics vary from 0 to 1, with 0 associated with the worst and 1 with the best performance. Recall and precision give complementary information regarding the classification performance of the algorithm and the F-score combines recall and precision in just one metric.

V. RESULTS

The first experiment tested the α parameter of the Canny filter. This parameter controls the strength of edge suppression: the higher the value of α the more the strength (gradient magnitude) of an edge surrounded by other edges is reduced [16]. The range of α reported here varies from 0.8 to 1.3

because preliminary exploratory experiments showed that the performance outside this range is very poor. For this experiment, the parameters σ and the Hough space threshold of the Hough transform, determined experimentally in previous experiments, were set to 1.6 and 16, respectively. The results are presented in Table II.

In other words, the Petri dish counting method was used to find the real number of viable cells on blue and colorless images. Further, it was compared with an evaluation method used by sugar cane mills. The resulting values for the used metrics correspond to the mean value from the application of the method to all images of the dataset TF-R2-B3 and the greater values are shown in bold case.

TABLE II
RESULTS RELATED TO THE VARIATION OF THE α PARAMETER

Dataset	α	Recall	Precision	F
R2-B3	0.8	0.9804	0.7235	0.8221
R2-B3	0.9	0.9830	0.7388	0.8335
R2-B3	1.0	0.9778	0.8104	0.8795
R2-B3	1.1	0.9727	0.8546	0.9058
R2-B3	1.2	0.9552	0.8899	0.9174
R2-B3	1.3	0.9247	0.9170	0.9170

The second test was conducted to find the threshold used to determine the peaks in the Hough space that correspond to circles. Threshold values from 15 to 18, determined in previous exploratory experimentation, were tested. The smooth parameter σ was fixed at 1.6 and the best three α values, regarding the F-score, found in the first experiment, were used. The results are presented in Table III.

TABLE III
RESULTS RELATED TO THE HOUGH SPACE THRESHOLD EXPERIMENT

Dataset	α	Threshold	Recall	Precision	F
R2-B3	1.3	15	0.9440	0.8545	0.8933
R2-B3	1.3	16	0.9248	0.9170	0.9170
R2-B3	1.3	17	0.8737	0.9399	0.9020
R2-B3	1.3	18	0.8272	0.9553	0.8834
R2-B3	1.2	15	0.9718	0.8158	0.8817
R2-B3	1.2	16	0.9552	0.8899	0.9174
R2-B3	1.2	17	0.9132	0.9323	0.9190
R2-B3	1.2	18	0.8624	0.9485	0.9006
R2-B3	1.1	15	0.9800	0.7546	0.8457
R2-B3	1.1	16	0.9727	0.8546	0.9058
R2-B3	1.1	17	0.9303	0.8974	0.9098
R2-B3	1.1	18	0.8946	0.9280	0.9094

In the experiment to determine the best σ value for the smoothing parameter, a variation from 1.4 to 1.8 was used. The value adopted for α was 1.2 (best F-score), and the Hough space threshold used was 16 and 17 (best F-score in experiment 1). The results are presented in Table IV.

In all experiments, the radii ranges used by the Hough transform search were determined visually, based on the usual size of the yeast cells. Table V summarizes the results of the application of the proposed method on each of the six datasets.

Figure 2 shows examples of the resulting images obtained during the experiments. Figure 2.(a) is the original image and

TABLE IV
RESULTS RELATED TO THE σ PARAMETER VALUE DETERMINATION

Dataset	σ	Threshold	Recall	Precision	F
R2-B3	1.4	16	0.9237	0.9056	0.9105
R2-B3	1.6	16	0.9552	0.8899	0.9174
R2-B3	1.7	16	0.9507	0.9004	0.9217
R2-B3	1.8	16	0.9504	0.9017	0.9222
R2-B3	1.4	17	0.9677	0.8661	0.9098
R2-B3	1.6	17	0.9132	0.9323	0.9190
R2-B3	1.7	17	0.9045	0.9317	0.9148
R2-B3	1.8	17	0.8919	0.9264	0.9056

TABLE V
RESULTS FROM THE APPLICATION OF THE PROPOSED METHOD

Dataset	Recall	Precision	F
R1-B3	0.908	0.970	0.936
R1-B6	0.815	0.950	0.874
R2-B3	0.952	0.900	0.922
R2-B6	0.888	0.859	0.868
R3-B3	0.888	0.957	0.917
R3-B6	0.902	0.838	0.861

from Figure 2.(b) to Figure 2, examples of the following procedures are shown: (b) saturation reinforcement, (c) smoothing, (d) gray-scale conversion, (e) edge detection and (f) circles detected by the Hough transform.

VI. DISCUSSION

Table V shows that the performance of the proposed method presents important variation between the datasets. It can be seen that the best F-Score and Precision were achieved in the dataset, TF-R1-B3, which has not been used for parameter tuning. However, the Recall is significantly higher in the TF-R2-B3 dataset, used for tuning. The minimum F-score was 86%, in the TF-R3-B6, and the maximum was 93%, which seems to be an encouraging result.

The proposed method has minor failures to detect yeast cells (false negatives) or to mistakenly identify as a yeast the circles that do not correspond to yeasts (false positives). As an example, a false positive should be a blue cell counted as a white cell, since the blue coloring corresponds to the nonviable yeast and the viable ones are presented in white color. To identify what might be causing this problem, all images where the failures occurred were visually analyzed with the help of an expert in the field.

Two main categories of mistakes were identified and are presented here. The most common failure is due to a lack of focus in some of the yeasts because of its rounded surface. The microscope focus is adjusted according to the trained human vision at the time that the images are captured. Sometimes, a satisfactory image for a specialist may have a poor yeast border demarcation, resulting in poor edge detection. The edge pixels of these out-of-focus yeasts have a color value that is very similar to the background and are not correctly detected by the Canny algorithm. This problem results in some false-negative errors, as illustrated in Figure 3.

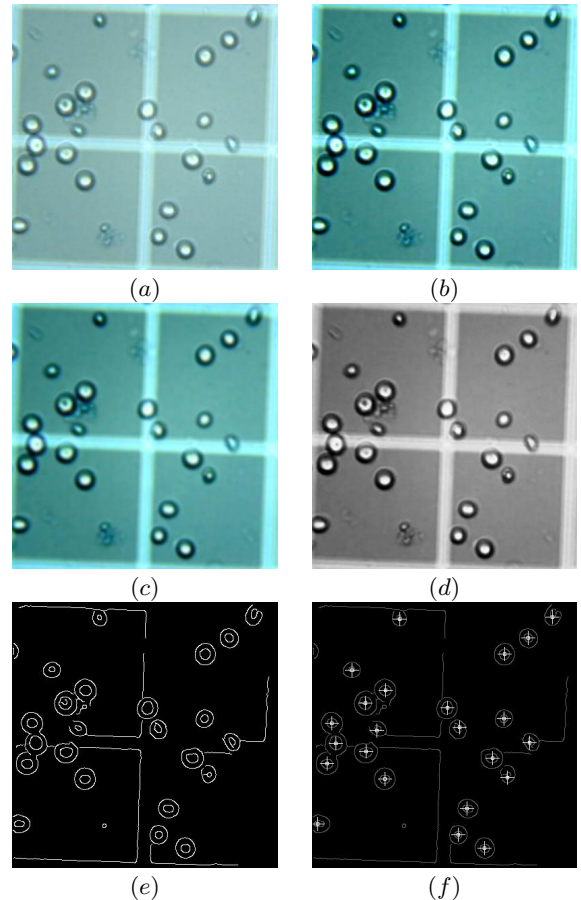


Fig. 2. Resulting images obtained during the experiments: (a) Original image. (b) Saturation enhancement. (c) Gaussian smoothing. (d) Grayscale conversion. (e) Canny filter. (f) Circles detected using the Hough transform.

The second cause of errors can be attributed to the presence of impurities in the liquid samples, which are unexceptional on sugarcane. Some of these impurities have well-marked borders that result in many undesirable edges found by the Canny algorithm that further disrupt the accumulation of evidence in the Hough Circles transformed space. Additionally, the possibility of chamber traces interference may occur. Two examples of images that exhibit this problem are shown in Figure 4, along with images of the CHT results.

It is worth mentioning the quick method used by sugar cane mills only evaluates the quality of ethanol cells. In addition, the Petri dish method is a unique counting procedure that results in the real number of viable cells. However, visual perceptiveness mistakes are present through the technician analysis, and the proposed method addresses a solution to this issue.

VII. CONCLUSION

The experiments indicate that the proposed method can achieve good precision (above 95%) and recall rates (above 98%) in some of the datasets tested and serve as baseline performance for further tests using this kind of image. Also, it is important to notice the precision of the sugarcane mill's measures is unspecified.

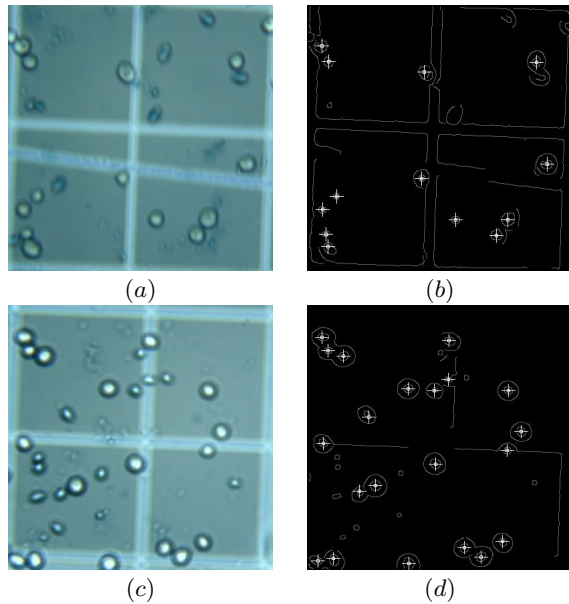


Fig. 3. Examples of images containing out of focus yeast cells: (a) Original image. (b) Circles detected by the proposed method. (c) Original Image. (d) Circles detected by the proposed method.

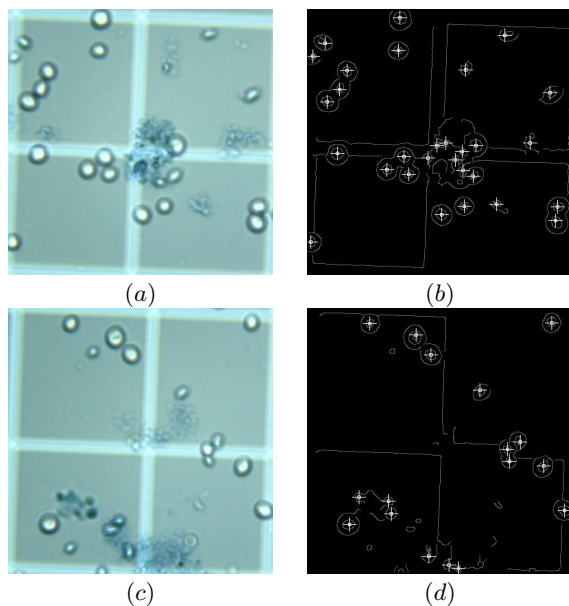


Fig. 4. Examples of highly compacted clusters of dust that impair the proposed method performance: (a) Original image. (b) Circles detected by the proposed method. (c) Original Image. (d) Circles detected by the proposed method.

A few other works are dealing with *Saccharomyces cerevisiae* cells counting however not use the same type of dye and control procedure studied in this paper. This is, the used method requires 24 hours to considerable analysis and the fermentation ends after 10 hours as well. Thus, the presentation of a new solution is also a contribution of this paper. The mean running time of the proposed approach has also been calculated, for further comparisons, and was 270.1806ms with a standard deviation of 30.032ms over 144 repetitions.

In the final system, the Hough transform yeast cell identification module is planned to serve as a preprocessing step to another, distance-based template matching module, that will refine this search and also separate viable from nonviable cells. So, high recall rates are more important, in this project, than precision, as this other module will have the opportunity to discard the regions detected by the Hough transform that do not correspond to yeast cells.

For future works, comparisons with other approaches are being planned. The running time may also be further reduced using the Random or the Segment Hough transforms, which are reported to significantly improve the processing time in other application domains.

ACKNOWLEDGMENT

This work has received financial support from the Dom Bosco Catholic University, UCDB and the Foundation for the Support and Development of Education, Science and Technology from the State of Mato Grosso do Sul, FUNDECT. Some of the authors have been awarded with Scholarships from the the Brazilian National Council of Technological and Scientific Development, CNPq and the Coordination for the Improvement of Higher Education Personnel, CAPES.

REFERENCES

- [1] E. Campelo and I. L. Michels, "A expansão da agroindústria canavieira em Mato Grosso do Sul," *Revista de Política Agrícola*, no. 3, 2009.
- [2] S. et al, *Balanço nacional de cana-de-açúcar e agroenergia*. Ministério da Agricultura, Pecuária e Abastecimento, Brasília, DF, 2007.
- [3] M. Sullivan and B. Bradford, "Viable cell yield from active dry yeast products and effects of storage temperature and diluent on yeast cell viability1," *Journal of Dairy Science*, vol. 94, no. 1, pp. 526 – 531, 2011.
- [4] M. Atanasova, G. Yordanova, R. Nenkova, Y. Ivanov, T. Godjevargova, and D. Dinev, "Brewing yeast viability measured using a novel fluorescent dye and image cytometer," *Biotechnology & Biotechnological Equipment*, vol. 33, no. 1, pp. 548–558, 2019.
- [5] N. Stolze, C. Bader, C. Henning, J. Mastin, A. E. Holmes, and A. L. Sutlief, "Automated image analysis with imagej of yeast colony forming units from cannabis flowers," *Journal of microbiological methods*, vol. 164, p. 105681, 2019.
- [6] Q. Liang, J. Long, Y. Nan, G. Coppola, K. Zou, D. Zhang, and W. Sun, "Angle aided circle detection based on randomized hough transform and its application in welding spots detection," *Mathematical biosciences and engineering: MBE*, vol. 16, no. 3, p. 1244, 2019.
- [7] G. Lin, Y. Tang, X. Zou, J. Cheng, and J. Xiong, "Fruit detection in natural environment using partial shape matching and probabilistic hough transform," *Precision Agriculture*, pp. 1–18, 2019.
- [8] W. O. Barbosa and A. W. Vieira, "On the improvement of multiple circles detection from images using hough transform," *TEMA (São Carlos)*, vol. 20, no. 2, pp. 331–342, 2019.
- [9] H. M. Ünver, Y. Kökver, E. Duman, and O. A. Erdem, "Statistical edge detection and circular hough transform for optic disk localization," *Applied Sciences*, vol. 9, no. 2, p. 350, 2019.

- [10] L. D. Novović, V. S. Ostojić, D. S. Starvević, and V. S. Petrović, "Radiography calibration marker detection using hough transformation," *Telfor Journal*, vol. 11, no. 1, pp. 30–34, 2019.
- [11] N. Stolze, C. Bader, C. Henning, J. Mastin, A. E. Holmes, and A. L. Sutlief, "Automated image analysis with imagej of yeast colony forming units from cannabis flowers," *Journal of Microbiological Methods*, v.164, 2019.
- [12] R. O. Duda and P. E. Hart, "Use os the hough transformation to detect lines and curves in pictures," *Graphics and Image Processing*, vol. 15, no. 01, 1972.
- [13] P. V. C. Hough, "Method and means for recognizing complex patterns," *USENIX Security Symposium*, 1962.
- [14] P. Hull, *Glucose Syrups: Technology and Applications*. Wiley-Blackwell, March 2010.
- [15] A. Gosiewska, M. Bakala, K. Woznica, M. Zwolinski, and P. Biecek, "Epp: interpretable score of model predictive power," *arXiv preprint arXiv:1908.09213*, 2019.
- [16] L. Gun, Z. Ran, and C. Honglei, "A contour detector with improved corner detection," *Multimedia Tools and Applications*, vol. 76, no. 4, pp. 5965–5984, 2017.