19th IMEKO TC10 Conference
"MACRO meets NANO in Measurement for Diagnostics, Optimization and Control"
Delft, The Netherlands, September 21–22, 2023

Use of Artificial Intelligence in optical microscope imaging

M. Carratù, V. Gallo, V. Laino, C. Liguori

University of Salerno, Via Giovanni Paolo II, 132 Fisciano (SA)
{mcarratu,vgallo,vlaino,tliguori}@unisa.it

***Abstract* – Identification and morphological analysis of microorganisms are of high interest in scientific research, especially in the medical field and food industry. Identification allows rapid functional characterization based on similarities with known related species enabling to confirm the identity of an isolate used, for example, in a trademarked industrial process. Monitoring of microorganisms within a given ecosystem and analysis of the morphological characteristics of the observed species enable quality control of the process under analysis. Such procedures are carried out manually in specialized laboratories by trained personnel using the appropriate optical equipment; therefore, it may be of great interest to use automatic measurement approaches that enable rapid and effective process analysis. Artificial intelligence techniques in computer vision and especially deep learning are well suited for this purpose. This article describes the realization of an automatic measurement system based on deep learning for the identification and measurement of morphological parameters of Saccharomyces cerevisiae microorganisms present in brewer's yeast by returning for each of the objects identified within the image the confidence score, the coordinates, and the dimensions of the corresponding ellipsoid-shaped cell. The metrological characteristics of the system have been defined through a calibration process by comparing measurements with a reference system.**

Keywords – Saccharomyces cerevisiae, Deep Learning, YOLOv5, Identification and Morphological Analysis, Good Health and Well-being

I. INTRODUCTION

Yeasts are eukaryotic microorganisms reproducing by budding or fission that can be found in almost all environments, including soil, water, and plants. They play an important role in various biological processes as evidenced by their biotechnological potential in food industry, indeed they are responsible for the production of a wide range of fermented products such as alcoholic beverages, fermented milk products, cereal based leavened products and condiments by improving the taste, flavour, texture, nutritive values, functionality and reducing anti-nutritional factors [1]. Upon a morphological point of view, yeasts present a high morphological divergence, with ellipsoidal and oval shapes being the most common. In fact, in the identification processes, microscopic evaluation is the first resource followed by other more discriminatory tests such as microbiological and biochemical ones [2]. The most commonly known type of yeast is Saccharomyces cerevisiae (Fig. 1) which is a species of yeast classified as unicellular fungus and can be considered the best valuable tool for most aspects of basic research on eukaryotic organisms; unlike other model organisms in molecular and cell biology research, S. cerevisiae is easily amenable to genetic manipulation making it a powerful tool for studying gene function and regulation, easy to cultivate and concomitantly of great importance for various biotechnological applications due to its unique biological characteristics, i.e., fermentation capacity accompanied by the production of alcohol and CO2, resilience to adverse conditions of osmolarity and low pH [3].

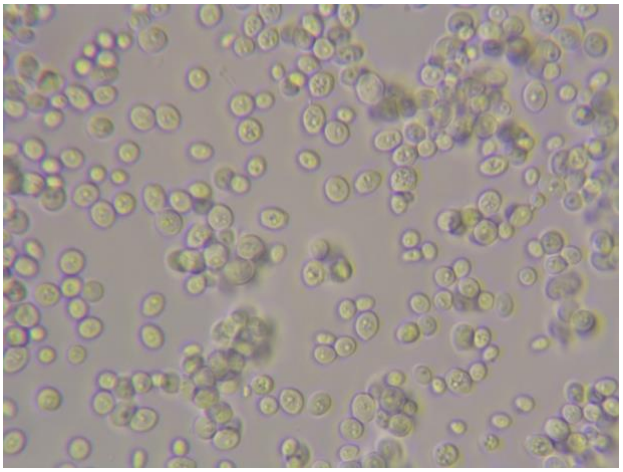

Figure 1. Saccharomyces cerevisiae under optical microscope

Moreover, it shares many cellular processes with more complex eukaryotes, including humans. As a result, researchers can use the yeast to model human diseases and study the underlying molecular mechanisms of disease. In this context the study of morphological properties is of great interest since morphology is one of the basic phenotypic characteristics of cells, and therefore conveys rich information about genetics, in fact abnormal yeast morphology can be a sign of cellular dysfunction and the study of Saccharomyces cerevisiae morphology can help researchers understand the molecular mechanisms underlying diseases such as cancer, neurodegenerative and metabolic disorders [4]. Saccharomyces cerevisiae has a distinctive oval to spherical shape with typical dimensions of approximately 3.5-9 μm in diameter appearing under a microscope as a round to oval-shaped cell with a smooth and a slightly refractive surface even if these dimensions can vary depending on the specific strain and the growth conditions in which it is cultured [5].
Due to its small size, it is difficult to observe, identify and analyse the cell morphology directly under a microscope indeed these procedures can often turn out to be inaccurate due to human error or inadequate setup. In this article an automatic measurement system based on deep learning is proposed to overcome the limitations of manual analysis and provide a fast and effective tool for the identification and measurement of morphological parameters of Saccharomyces cerevisiae microorganisms.

II. RELATED RESULTS IN THE LITERATURE

Several techniques can be used to compute quantitative and qualitative morphological analysis of yeast's cells enabling the possibility of on-line estimation: current mainstream methods are bright field or phase contrast microscopy (light microscopy) in which cells are stained with a suitable dye to observe the morphology of the cells, flow cytometry to analyze the size and shape of the cells in a population and electron microscopy to reveal cellular structure or surface details as in situ microscopy (ISM). An application of ISM based on an image analysis algorithm that uses detection of regional maxima can be used to determine the average size of single cells enabling on-line size monitoring [6]. With the progress in image processing technology, it turns out to be easier to extract information on several parameters, enabling the use of this technology for the measurement of microbial morphology, in fact image analysis and automated methods can be used to provide a better and detailed understanding of the morphology of Saccharomyces cerevisiae cells; in this context are well suited image-processing techniques to monitor yeasts cultivation directly using high-speed cameras [7], [8] and machine learning approaches [9] using classical segmentation algorithms [10], [11] and ones based on a set of relevant individual cell features based on first and second order histograms and wavelet-based texture measurement extracted from the microscope images of the yeast cells to represent the morphological characteristics in a more sophisticated way [12]. Counting procedures can be computed using the traditional microscopy or more effectively resorting to automatic systems as the image processing techniques and the segmentation algorithms cited before or bright-field microscopy and dye-exclusion methods [13].

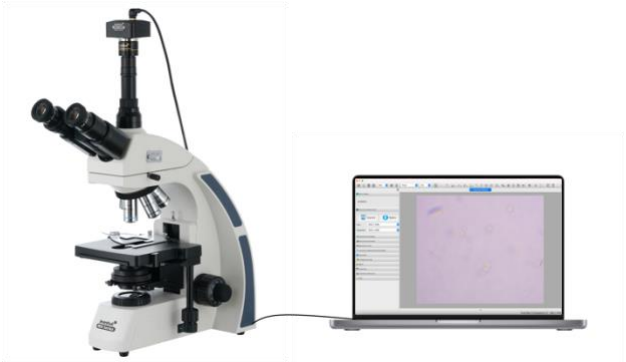

Figure 2. Measurement setup

III. PROPOSED METHOD

The methodology regarded firstly the investigation on morphological characteristics and aspects regarding the shape and the dimensions of the cells under analysis, subsequently the acquisition of the dataset needed to train the neural network and finally, the evaluation of network performances and the definition of its metrological characteristics. The preliminary dataset consisted of 500 images of a mixture of water and brewer's yeast that were acquired using the Levenhuk MED D40T microscope by resorting to 100x magnification and oil immersion of the sample. Levenhuk MED D40T is a digital microscope with a trinocular head and a 16 MP camera; the infinity-corrected planachromatic optical system produces a flat field, eliminates chromatic aberration, and enables observations with a high degree of detail enabling the transmission of clear and high-contrast images with an excellent level of flatness. The samples of the mixture were placed under the objective with the help of a mechanical shifter and moved in two directions while coarse and fine focus was used for sharpness adjustment so as to obtain images with as high a degree of detail as possible; the measurement setup is shown in the Figure 2. The images were initially labelled and then properly made through a pre-processing (images undergo self-orientation and 416x416 resizing) followed by a data augmentation process as a way to obtain new images created by horizontally rotating labelled images in order to expand the size and the variability of the dataset. Downstream of the data augmentation process, a total number of 748 images is obtained; 80% of this dataset was used for training, 14% for validation and 6% for test.

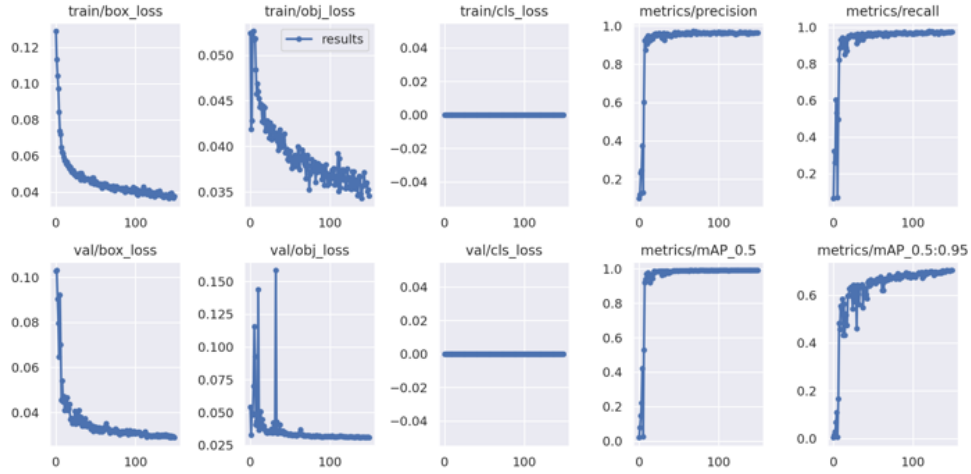

Figure 3. Training results

For the purpose of this work a convolutional neural network (CNN) [14], [15] based on the YOLO (You Only Look Once) algorithms family was employed for its fast processing and high accuracy [16]. More specifically the used algorithm is YOLOv5 [17] whose architecture consists of three main components: backbone, neck, and head. The backbone is a convolutional neural network that reduces the amount of calculation and increases the speed of inference using a Cross Stage Partial network (CSP) and aggregates ad forms image features at different granularities (feature extraction) using a Spatial Pyramid Pooling (SPP). The neck neural network uses the Feature Pyramid Network (FPN) and PANet to mix and combine features and pass them forward to the prediction. The head is used to generate final vectors with class probabilities, confidence scores and bounding boxes for each of the object detected in the image [18]. The algorithm splits the image of interest into cells using a S x S grid, if the centre of an object falls into a grid cell this one is responsible for the detection of that object. Each grid cell is responsible for the prediction of B bounding boxes and confidence scores obtaining 5 predictions: x_c, y_c, w, h, and confidence where the x_c, y_c coordinates represent the centre of the box and w, h are respectively width and height thereof while confidence score reflects how confident the model is that the box contains an object and also, how accurate the predicted box is; moreover each grid cell predicts C conditional class probabilities that are probabilities conditioned on the grid cell containing an object. Multiplying the conditional class probabilities by the confidence predictions of the individual boxes gives class confidence scores (1) giving a class-specific confidence score for each box [19].

$$Pr(class_i) * IoU_{pred}^{truth} \qquad (1)$$

These scores encode both the probability of that class appearing in the box and how well the predicted box fits the object.

IV. RESULTS AND DISCUSSIONS

The training of the model made on 150 epochs required a proper hardware to manage the computational load in order to decrease the training process duration; the results of the process are shown in Figure 3.

In order to define the metrological characteristics of the realized system, a calibration process was used by comparing measurements with a reference system. The first step regarded the calibration of the camera through the reference software used to acquire the dataset by resorting to the use of the appropriate calibration target (Fig. 4); the result of the calibration process was the conversion factor (equal to 38.480 $\frac{pixel}{\mu m}$) with which was possible to make measurements on the objects of interest for a given magnification value (in the case of interest equal to 100x) according to the selected unit of measurement (µm). In this way the outputs of the trained network that were expressed in normalized coordinates were able to be converted in µm.

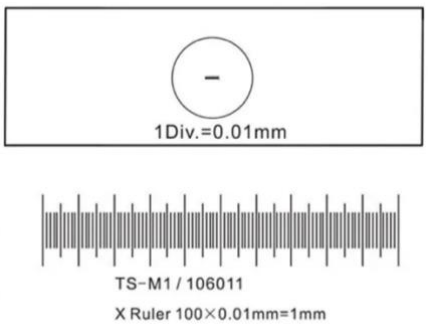

Figure 4. Calibration target

Twenty images for a total of 152 cells were analysed; the measurements made on each cell (position and diameters of ellipses) were made first with LevenhukLite software

(Fig. 5) and then with the automatic measurement system realized (Fig. 6) in the reference coordinate system (Fig. 7). Reference measurements were made by drawing ellipses on the objects of interest exploiting LevenhukLite software. For each of the plotted ellipses, it was possible to derive information regarding the location of the centre of the ellipse in the image and to the major and minor diameters expressed in µm using the related measurements sheet (Table 1); correspondingly the realized system returned the matching parameters of interest for each of the identified cells in the image (Table 2).

Figure 5. Reference measurements

	x_c [µm]	y_c [µm]	w [µm]	h [µm]
1	83.80	45.18	5.90	5.90
2	84.12	36.41	7.38	6.96
3	94.76	49.32	6.42	6.19
4	98.61	46.04	6.42	6.00
5	99.56	50.81	6.96	6.91

Table 1. Measurements sheet

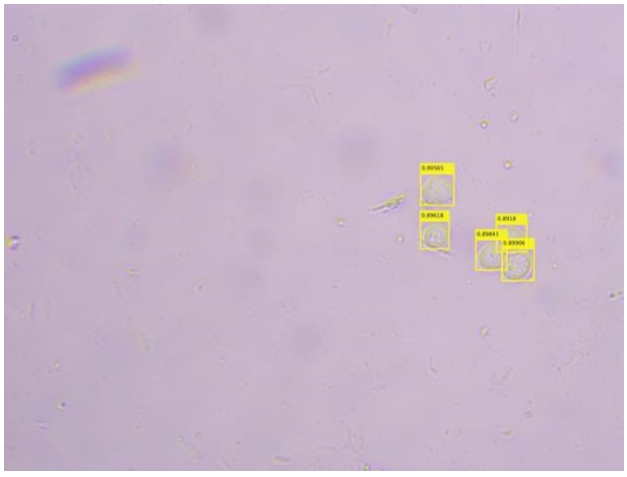

Figure 6. Automatic system detections

	x_c [µm]	y_c [µm]	w [µm]	h [µm]
1	83.65	45.33	5.78	5.72
2	84.15	36.51	6.70	6.50
3	94.64	49.20	6.25	6.16
4	98.57	45.93	6.12	5.66
5	99.85	51.12	6.44	6.58

Table 2. Automatic system measurements

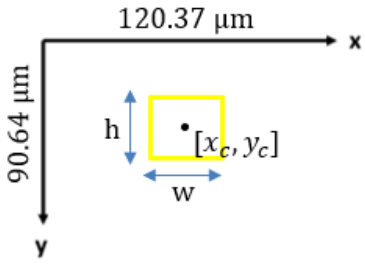

Figure 7. Reference

The calibration curves obtained downstream the correction of systematic effects are shown below (Fig. 8-11); the x-axis reports the values measured by the realized model while the y-axis reports the reference values. The regression used is of the linear type.
The parameter R^2 (Table 3), also called the coefficient of determination, is a proportion between the variability of the data and the correctness of the statistical model used, it expresses the "goodness" of the curve obtained and in particular it indicates the extent to which the dependent variable is predictable; if R^2 is close to 1 it means that the regressors predict well the value of the dependent variable, while if it is 0 it means that they do not.

	x_c	y_c	w	h
$\mathbf{R^2}$	0.99998	0.99997	0.91239	0.93039

Table 3. R^2

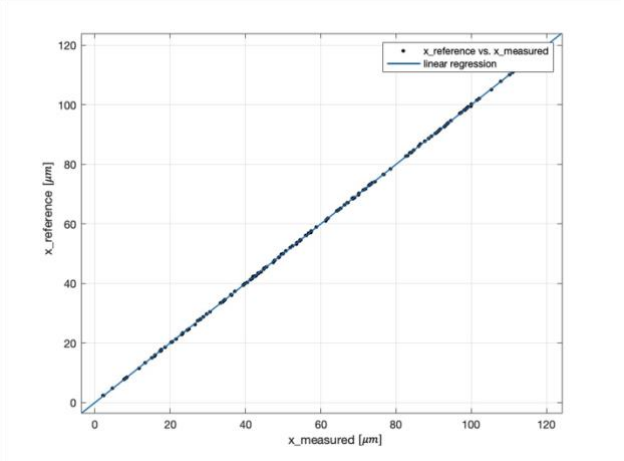

Figure 8. x_c coordinate calibration curve

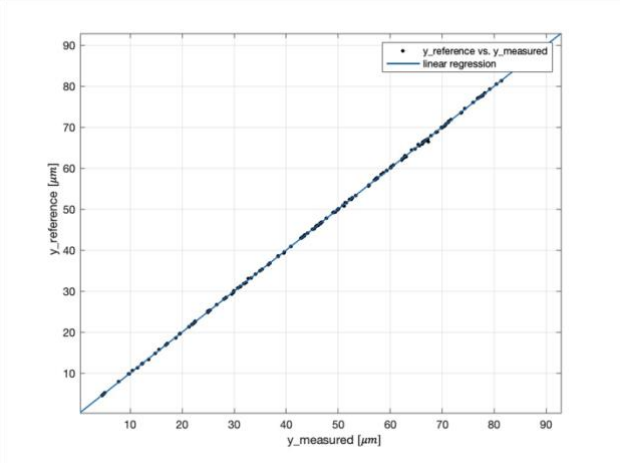

Figure 9. y_c coordinate calibration curve

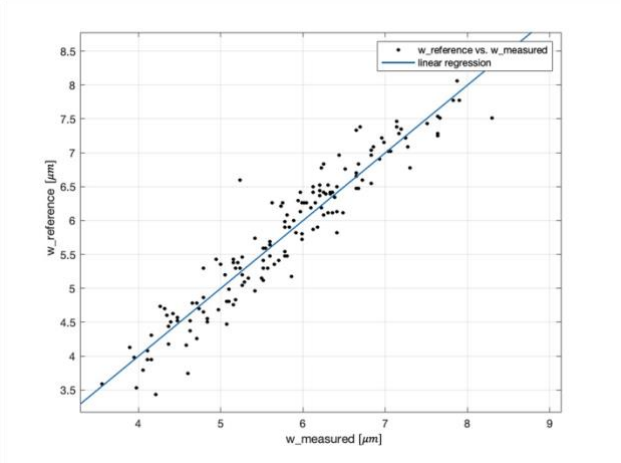

Figure 10. w calibration curve

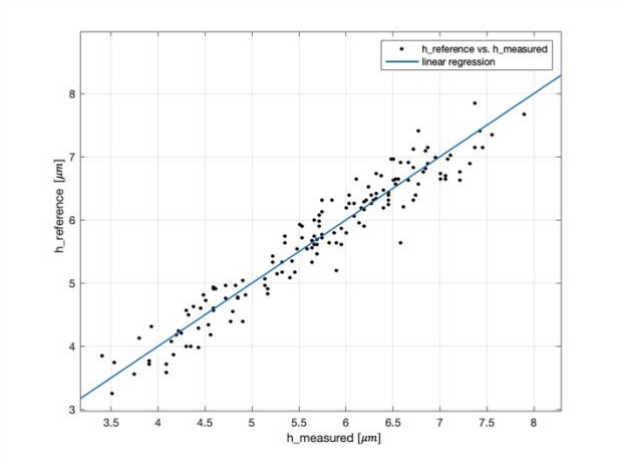

Figure 11. h calibration curve

For each of the measured parameters (x_c, y_c, w, h), mean (2) and maximum (3) relative errors RE were calculated after the correction of systematic effects.

$$mean\ RE = \frac{1}{N}\sum_{i=1}^{N}\frac{|reference_i - measured_i|}{reference_i} \quad (2)$$

$$max\ RE = max\left(\frac{|reference_i - measured_i|}{reference_i}\right) \quad (3)$$

Relative errors were found to be larger for w and h parameters than for x_c and y_c coordinates because each refers to a pair of points while the couple (x_c, y_c) represents a single point in the image.

mean RE **x_c**	***mean RE*** **y_c**	***mean RE*** **w**	***mean RE*** **h**
0.33%	0.24%	4.52%	4.14%

Table 4. mean relative errors

max RE **x_c**	***max RE*** **y_c**	***max RE*** **w**	***max RE*** **h**
6.90%	2.30%	23.09%	16.73%

Table 5. maximum relative errors

V. CONCLUSIONS AND OUTLOOK

The purpose of this work was to implement an automatic measurement system based on deep learning that could detect and measure morphological parameters of Saccharomyces cerevisiae microorganisms automatically. The developed system allows a rapid functional characterization of the observed species and to realize quality control of the process under analysis. The network could be trained to recognize the presence of microorganisms unrelated to those of interest such as bacteria and mites. The results of this work have revealed that the automatic measurement system realized can be considered as an effective method to compute cell's measurements quickly and with a good degree of reliability.

Morphological analysis of microorganisms is of great interest in the field of medical research; when a microorganism is identified, the clinicians can perform analyses to determine which drugs are most effective against it, so that effective treatment can be arranged as soon as possible. The considerations made in this paper can therefore be extended to this area as an aid to the doctor in the identification and morphological analysis of the microorganism responsible for a given disease.

REFERENCES

[1] **Rai, Amit Kumar and Kumaraswamy Jeyaram**: "Role of Yeasts in Food Fermentation." (2017).

[2] **Maicas, Sergi**: 2020. "The Role of Yeasts in Fermentation Processes" Microorganisms 8, no. 8: 1142. https://doi.org/10.3390/microorganisms8081142.

[3] **Parapouli M, Vasileiadis A, Afendra AS, Hatziloukas E**: Saccharomyces cerevisiae and its industrial applications. AIMS Microbiol. 2020 Feb 11;6(1):1-31. doi: 10.3934/microbiol.2020001. PMID: 32226912; PMCID: PMC7099199.

[4] **Kaminska, J.; Soczewka, P.; Rzepnikowska, W. Zoladek, T.**: Yeast as a Model to Find New Drugs and Drug Targets for VPS13-Dependent Neurodegenerative Diseases. Int. J. Mol. Sci. 2022, 23, 5106. https://doi.org/10.3390/ijms23095106.

[5] **Aon, J.C., Tecson, R.C. & Loladze, V.**: Saccharomyces cerevisiae morphological changes and cytokinesis arrest elicited by hypoxia during scale-up for production of therapeutic recombinant proteins. Microb Cell Fact 17, 195 (2018). https://doi.org/10.1186/s12934-018-1044-2.

[6] **Belini VL, Suhr H, Wiedemann P**: Online monitoring of the morphology of an industrial sugarcane biofuel yeast strain via in situ microscopy. J Microbiol Methods. 2020 Aug;175:105973. doi: 10.1016/j.mimet.2020.105973. Epub 2020 Jun 6. PMID: 32522492.

[7] **Saito TL, Ohtani M, Sawai H, Sano F, Saka A, Watanabe D, Yukawa M, Ohya Y, Morishita S.**: SCMD: Saccharomyces cerevisiae Morphological Database. Nucleic Acids Res. 2004 Jan 1;32(Database issue):D319-22. doi: 10.1093/nar/gkh113. PMID: 14681423; PMCID: PMC308847.

[8] **Kitahara Y, Itani A, Oda Y, Okamura M, Mizoshiri M, Shida Y, Nakamura T, Kasahara K, Ogasawara W**: A real-time monitoring system for automatic morphology analysis of yeast cultivation in a jar fermenter. Appl Microbiol Biotechnol. 2022 Jun;106(12):4683-4693. doi: 10.1007/s00253-022-12002-0. Epub 2022 Jun 10. PMID: 35687157.

[9] **D. Capriglione, M. Carratu, A. Pietrosanto, and P. Sommella**, "Soft Sensors for Instrument Fault Accommodation in Semiactive Motorcycle Suspension Systems," IEEE Transactions on Instrumentation and Measurement, vol. 69, no. 5, pp. 2367–2376, 2020, doi: 10.1109/TIM.2019.2963552.

[10] **Tleis, Mohamed & Verbeek, Fons**: (2015). Machine learning approach to segment Saccharomyces cerevisiae yeast cells. 278-281. 10.1109/ICABME.2015.7323306.

[11] **Dietler N, Minder M, Gligorovski V, Economou AM, Joly DAHL, Sadeghi A, Chan CHM, Koziński M, Weigert M, Bitbol AF, Rahi SJ**: A convolutional neural network segments yeast microscopy images with high accuracy. Nat Commun. 2020 Nov 12;11(1):5723. doi: 10.1038/s41467-020-19557-4. PMID: 33184262; PMCID: PMC7665014.

[12] **Tleis MS, Verbeek FJ**: Machine Learning approach to discriminate Saccharomyces cerevisiae yeast cells using sophisticated image features. J Integr Bioinform. 2015 Oct 6;12(3):276. doi: 10.2390/biecoll-jib-2015-276. PMID: 26673792.

[13] **Thomson, K., Bhat, A., and Carvell, J.**: (2015), Comparison of a new digital imaging technique for yeast cell counting and viability assessments with traditional methods. J. Inst. Brew., 121, 231– 237. doi: 10.1002/jib.224.

[14] **M. Carratu, C. Liguori, A. Pietrosanto, M. O'Nils, and J. Lundgren,** "Data fusion for timber bundle volume measurement," in I2MTC 2019 - 2019 IEEE International Instrumentation and Measurement Technology Conference, Proceedings, 2019, vol. 2019-May. doi: 10.1109/I2MTC.2019.8826961.

[15] **I. Shallari, V. Gallo, M. Carratu, M. O'Nils, C. Liguori and M. Hussain**: "Image Scaling Effects on Deep Learning Based Applications," 2022 IEEE International Symposium on Measurements & Networking (M&N), Padua, Italy, 2022, pp. 1-6, doi: 10.1109/MN55117.2022.9887705.

[16] **Buonocore, Daniele & Carratù, Marco & Lamberti, Maria**: (2023). Classification of coffee bean varieties based on a deep learning approach. 14-19. 10.21014/tc10-2022.002.

[17] **Ultralytics,** YOLOv5**,** https://github.com/ultralytics/yolov5

[18] **Horvat, Marko & Jelečević, Ljudevit & Gledec, Gordan**: (2022). A comparative study of YOLOv5 models performance for image localization and classification.

[19] **J. Redmon, S. Divvala, R. Girshick and A. Farhadi**: "You Only Look Once: Unified, Real-Time Object Detection," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 779-788, doi: 10.1109/CVPR.2016.91.