Supplementary Information (SI) for Journal of Materials Chemistry B.

Quantification of *Candida* spp. using fluorescence and SERS spectroscopy for bloodstream infections diagnosis

SUPPLEMENTARY INFORMATION

*Manuel García-Algar, Anna Mur, Sergi Casellas-Díaz, Marina B. Velasco, Paulo Alexandre Gomes, Laura Pol, Juan Sagales, and Sara Gómez-de Pedro**

Medcom Advance, Carrer de Marcel·lí Domingo 2-4, Edifici N5, 43007 Tarragona, Spain

*E-mail: saragomez.medcomadvance@gmail.com

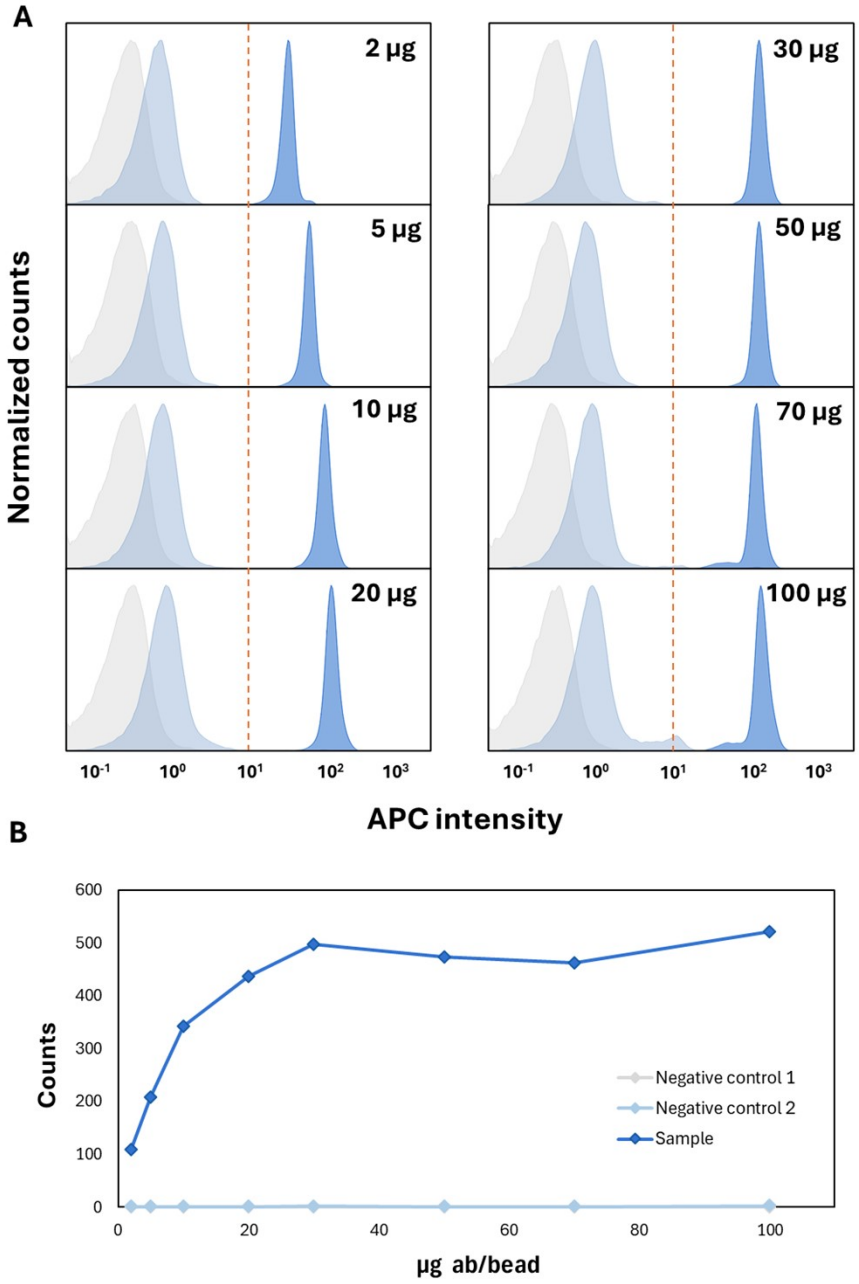


Figure S1. A) Flow cytometer plots for the optimization of the antibody immobilization to the mSERStag_Ca; left column, from top to bottom: 2 µg ab/bead, 5 µg ab/bead, 10 µg ab/bead and 20 µg ab/bead; right column, from top to bottom: 30 µg ab/bead, 50 µg ab/bead, 70 µg ab/bead and 100 µg ab/bead. B) Fluorescence dependent antibody concentration for the mSERStag_Ca (blue). Negative controls are also shown in the plot (light grey for particles without secondary antibody (Negative control 1) and light blue for particles with anti-mouse secondary antibody Negative control 2)).

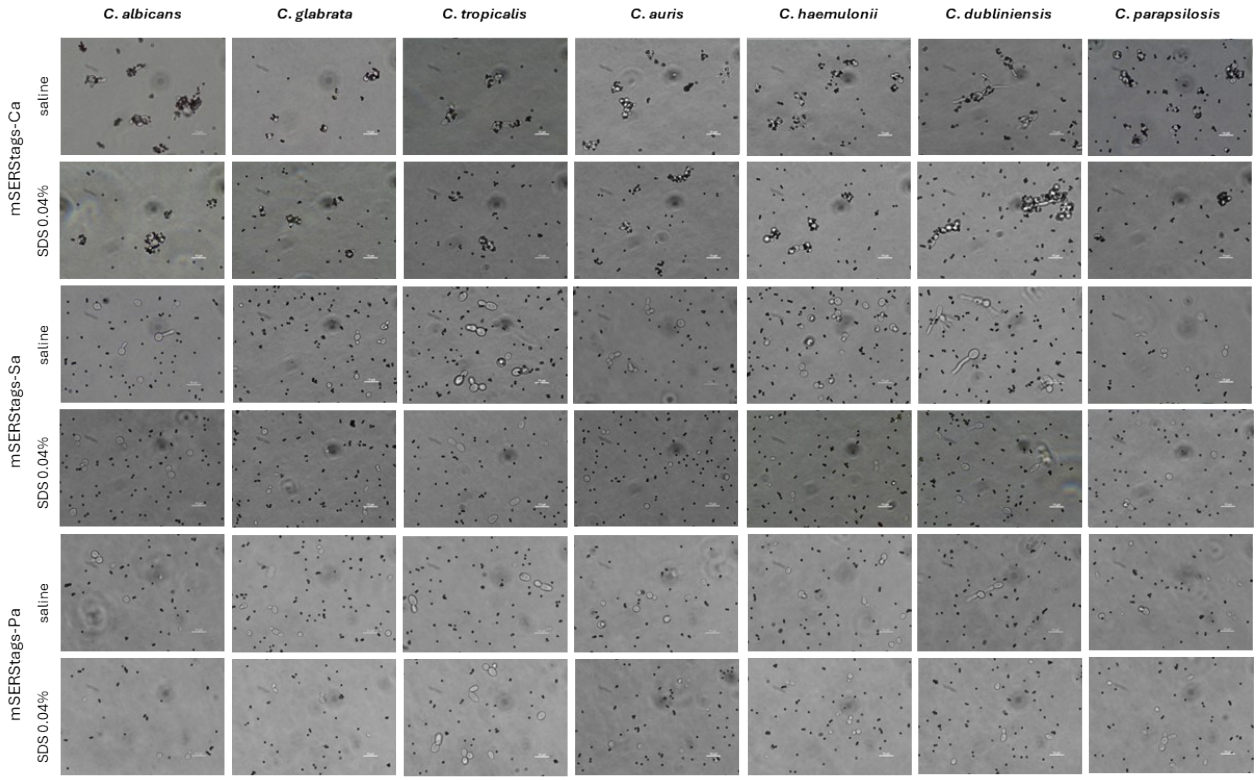


Figure S2. Microscope images of magnetic SERS Tags reactivity with diverse *Candida spp.*, in saline solution and 12.5-fold diluted SDS (previously incubated 30 minutes with 0.5% SDS). Results from recognition of mSERStags-Ca and cross-reactivity of mSERStags-Sa and mSERStags-Pa are shown.

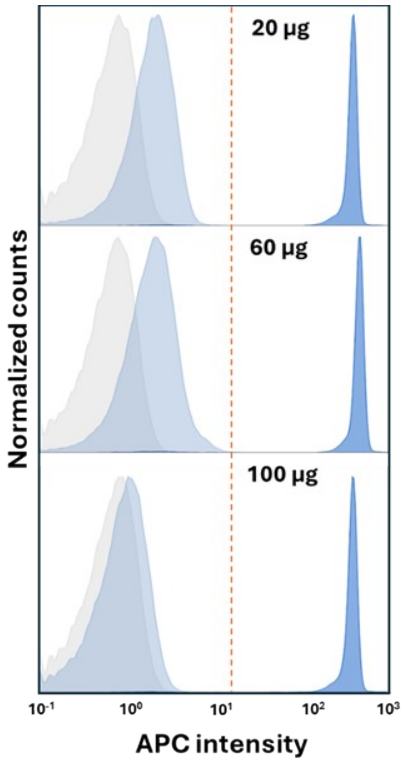


Figure S3. Characterization of conjugated magnetic beads for retention of *Candida spp.* with different ratios of antibody: 20µg ab, 60µg ab and 100 µg ab per mg bead. Plots include negative controls with only beads (gray), negative controls adding anti-mouse secondary antibody (light blue) and a positive sample of beads with anti-rabbit secondary antibody (dark blue). As it can be seen, 60µg ab/mg bead was found as the optimal condition due to the higher intensity attained.

Data analysis:

Images were analyzed with a developed real-time analysis code in Python. A color filter based on the Hue, Saturation and Value (HSV) was threshold for the detection of fluorescent spots. A MobileNetV3Small Neural Network was used, since code can be run at same time as sample acquisition. CNN training was based on binary classification, which included “candida” and “not candida” types. A dataset of 260 “candida” and 1725 “not candida” fluorescent images were used, which were collected by samples from experiments with a high and low concentration of different *Candida* spp. Data augmentation was also applied to the images by performing random rotations and inversions. A total of 10k images were generated per class, and an 80/20 dataset split was performed for training and validation, respectively.

Training and validation were performed based on 50 epochs, achieving near 100% for both with near zero loss (Figure S4A). Training and validation scores were perfectly aligned, showing no signs of overfitting or underfitting.

A second CNN was trained based on the white light image obtained for the final fluorescent CNN. Same classification, CNN architecture and model training procedure was used, “candida” and “not candida”, MobileNetV3Small and decreasing learning rate, respectively. Concluded the training, the model also achieved a near 100% validation and training accuracy, with perfectly aligned scores [Figure S4B].

Finally, last component of the sample analysis involved the development of a SOM model capable of classifying the different magnetic SERS tags. A model was trained using 200 SERS spectra of each label class: POT (yellow), MCN (orange) and MBA (green), which were manually selected. Before training, the Raman spectra were preprocessed by applying cosmic-ray removal, baseline subtraction and normalization. With these considerations, most peaks were very well-defined resulting in a high, 96%, accuracy score. To improve the inherent SERS signal spread when measuring Raman maps, a preprocessing algorithm was created to confine SERS signal to the location of the particles, overlapping its correspondent white light image.

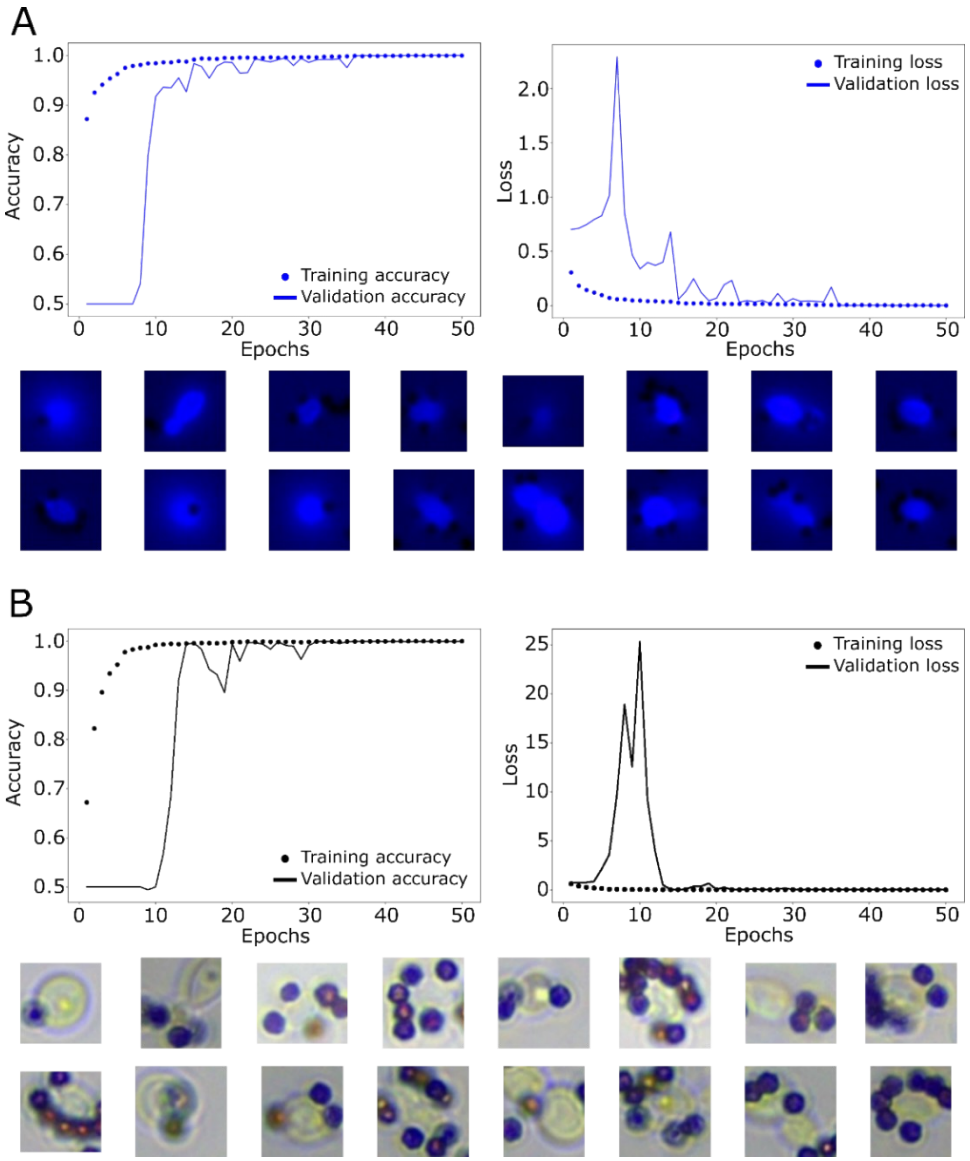
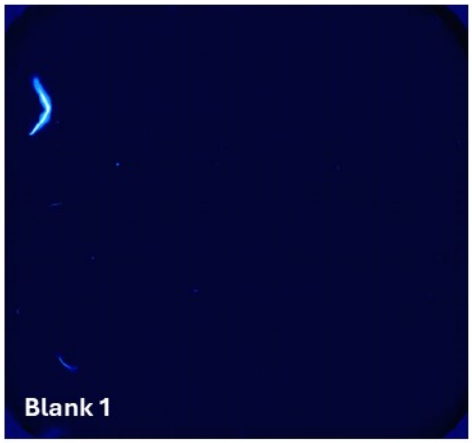
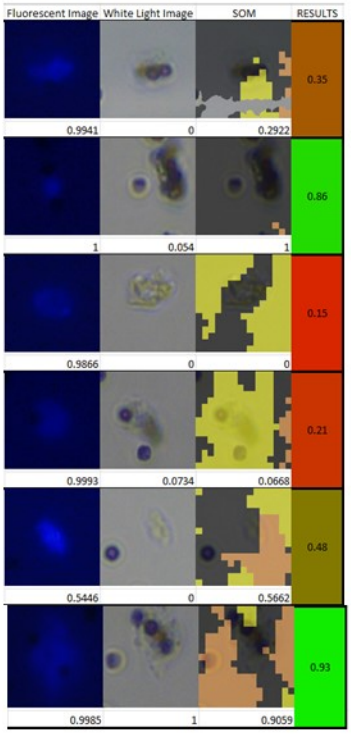


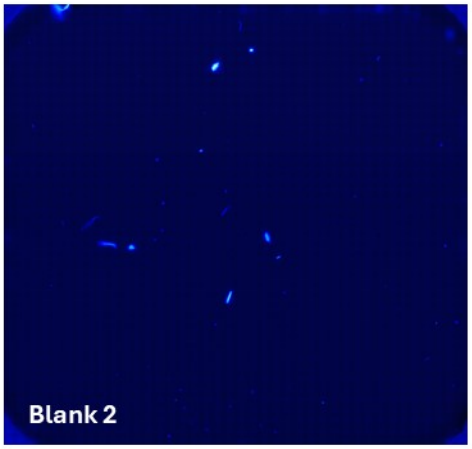
Figure S4. CNN models. A) Fluorescence CNN training, showing the evolution of the training and validation accuracy and loss though the 50 epochs run. Showing consists of convergence and high accuracy scores with low overfit and underfit. Below are shown some examples of the raw input data from the “candida” class. B) White light CNN training, showing the evolution of the training and validation accuracy and loss though the 50 epochs run. Showing similar consistency of convergence and high accuracy scores with low overfit and underfit. Below are shown some examples of the raw input data from the “candida” class.

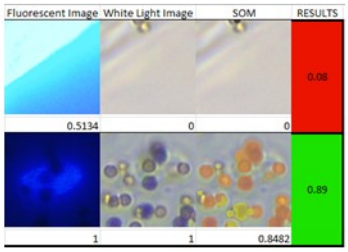
a)


Blank 1

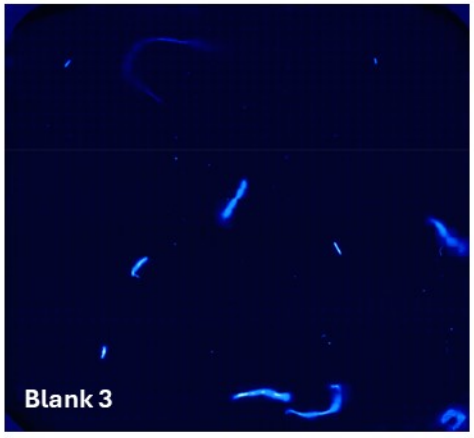

Fluorescent Image
White Light Image
SOM
RESULTS
0.35
0.9941
0
0.2922
0.86
1
0.054
1
0.15
0.9866
0
0
0.21
0.9993
0.0734
0.0668
0.48
0.5446
0
0.5662
0.93
0.9985
1
0.9059

b)


Blank 2


Fluorescent Image
White Light Image
SOM
RESULTS
0.08
0.5134
0
0
0.89
1
1
0.8482

c)


Blank 3

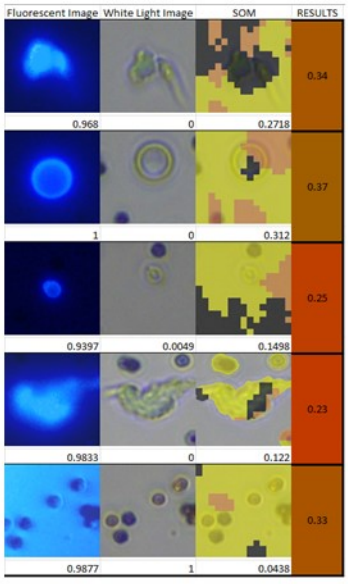

Fluorescent Image
White Light Image
SOM
RESULTS
0.34
0.968
0
0.2718
0.37
1
0
0.312
0.25
0.9397
0.0049
0.1498
0.23
0.9833
0
0.122
0.33
0.9877
1
0.0438

Figure S5. A-C) Blank samples. On the left, Fluorescence microscope image from the whole montage. On the right, fluorescent and white light images, and SOM analysis obtained from the final report generated, including the scoring from the three different detections. Positive or negative result, referred as “candida” or “not candida”, was identified in green or red, respectively.

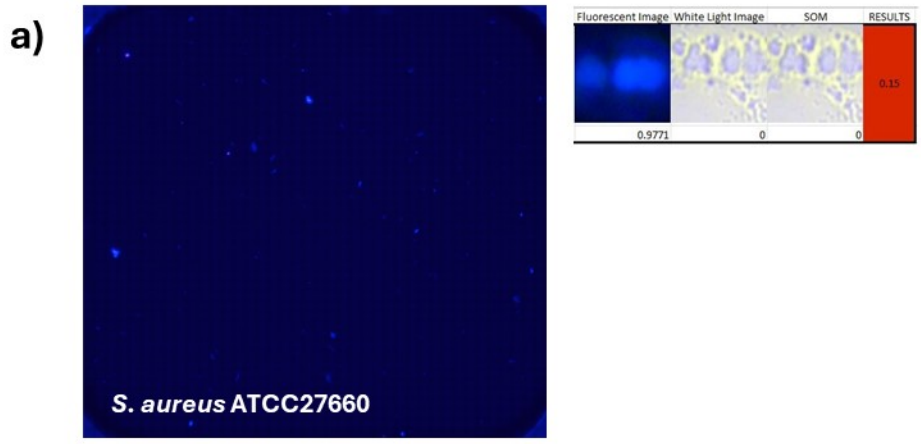
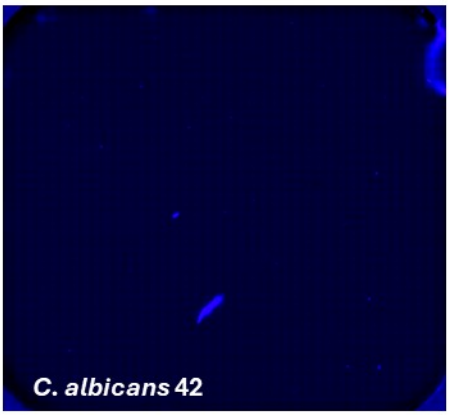
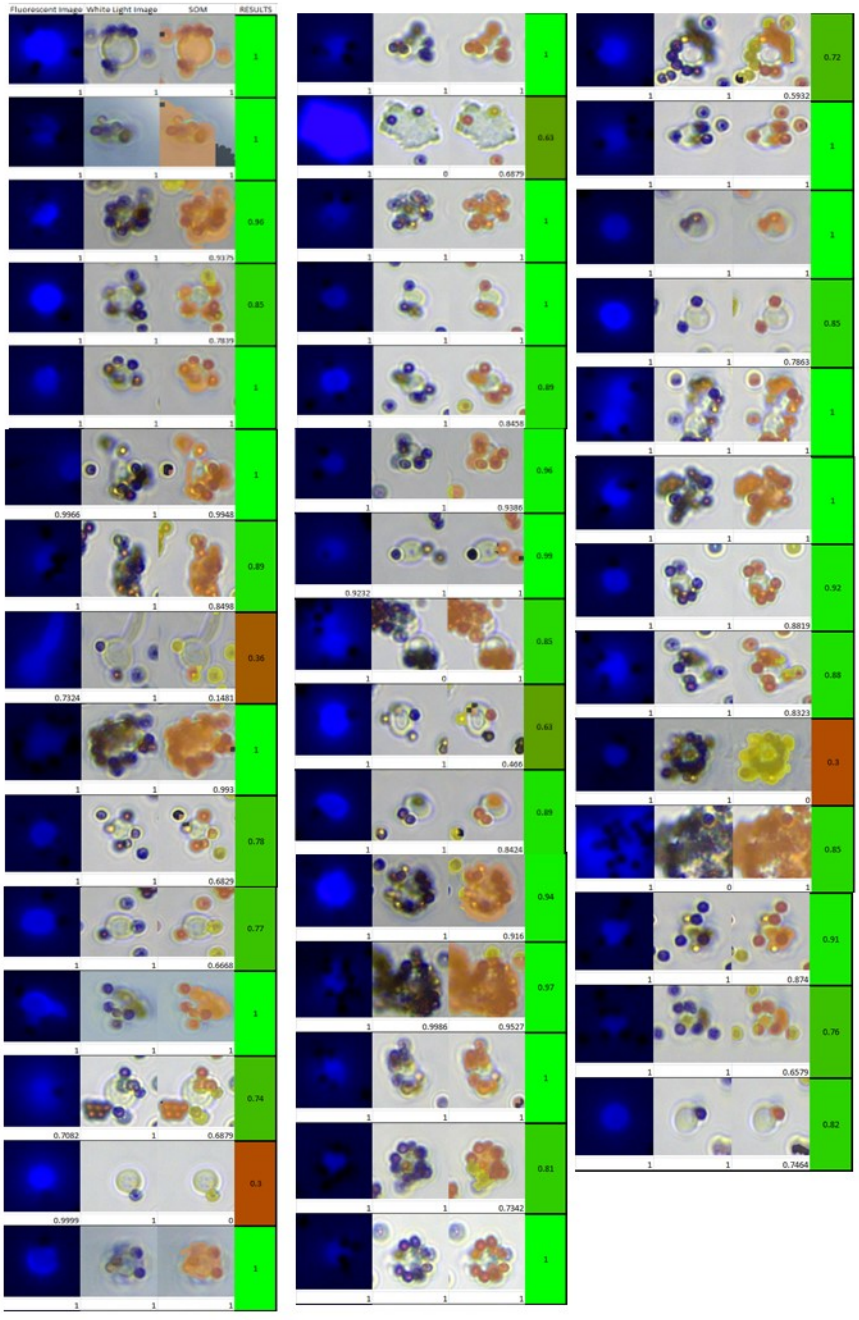


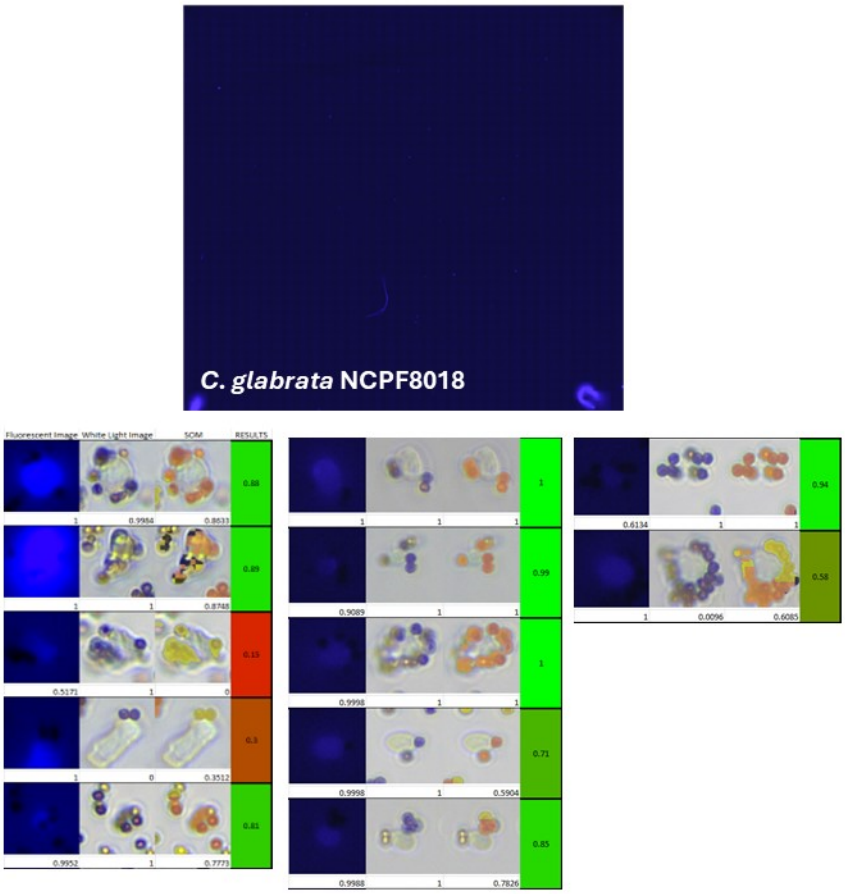
Figure S6. Negative control samples. A) *S. aureus* sample. B) *P. aeruginosa* sample. On the left, fluorescence microscope image from the whole montage. On the right, fluorescent and white light images, and SOM analysis obtained from the final report generated, including the scoring from the three different detections. Positive or negative result, referred as “candida” or “not candida”, was identified in green or red, respectively.

a)

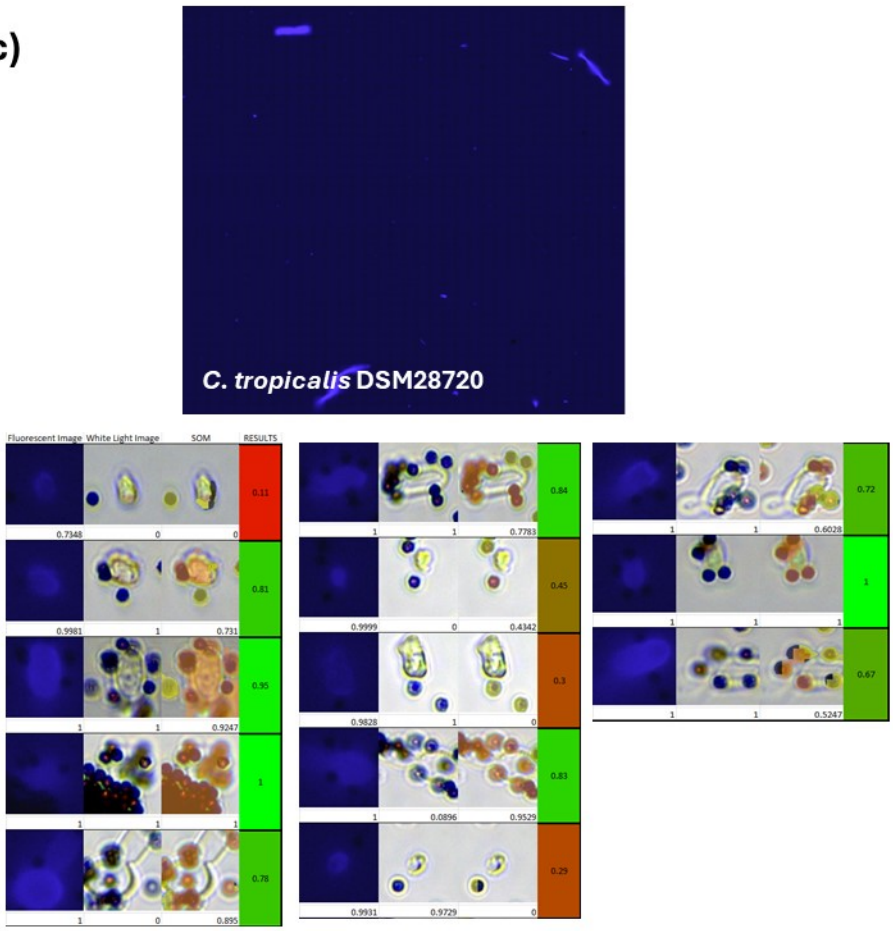

C. albicans 42


Fluorescent Image
White Light Image
SOM
RESULTS

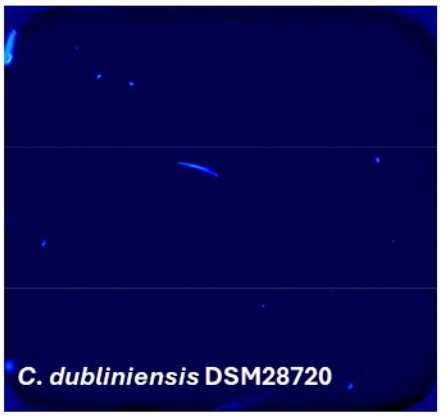
b)

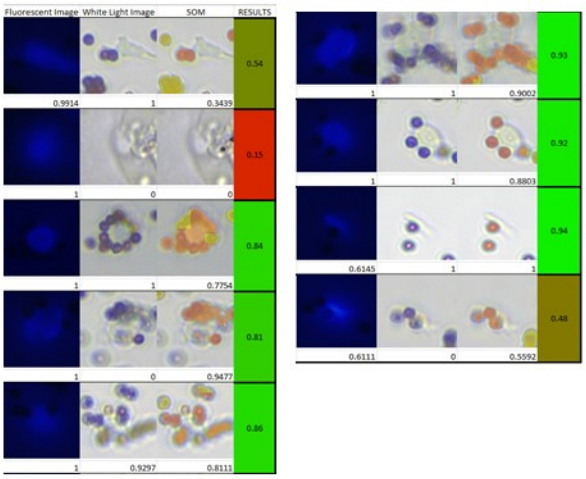

C. glabrata NCPF8018
Fluorescent Image
White Light Image
SOM
RESULTS

c)

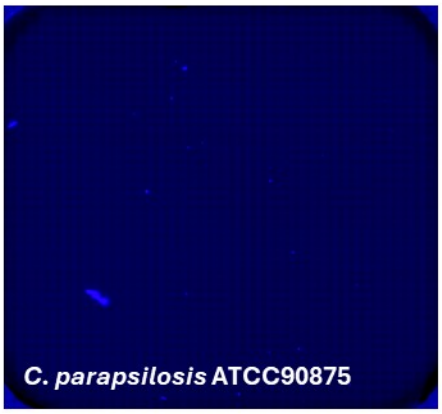

C. tropicalis DSM28720
Fluorescent Image
White Light Image
SOM
RESULTS

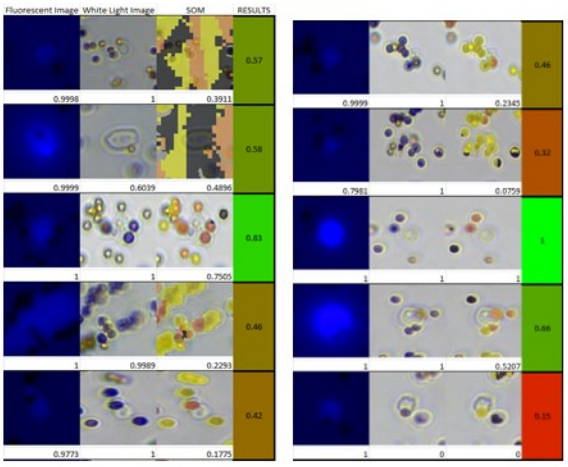
d)


C. dubliniensis DSM28720

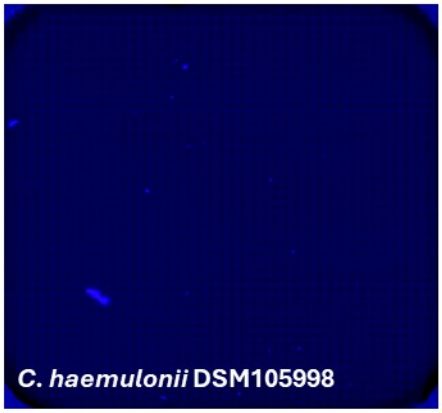

Fluorescent Image
White Light Image
SOM
RESULTS

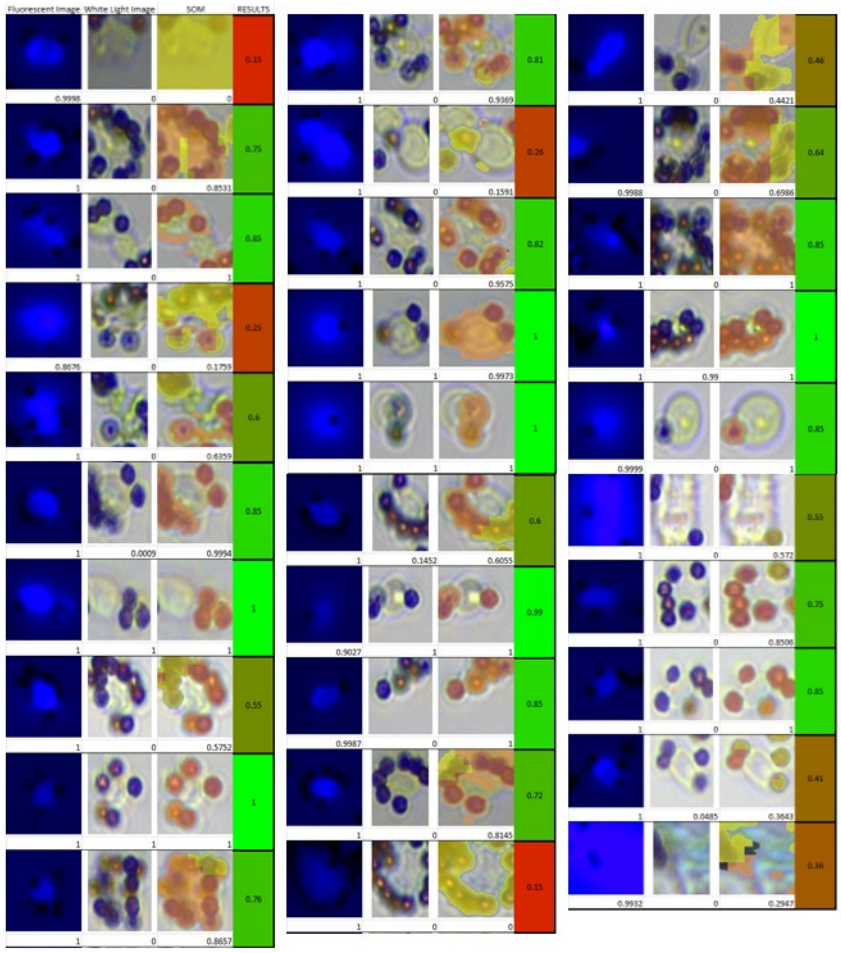
e)


C. parapsilosis ATCC90875


Fluorescent Image
White Light Image
SOM
RESULTS

f)


C. haemulonii DSM105998


Fluorescent Image
White Light Image
SOM
RESULTS

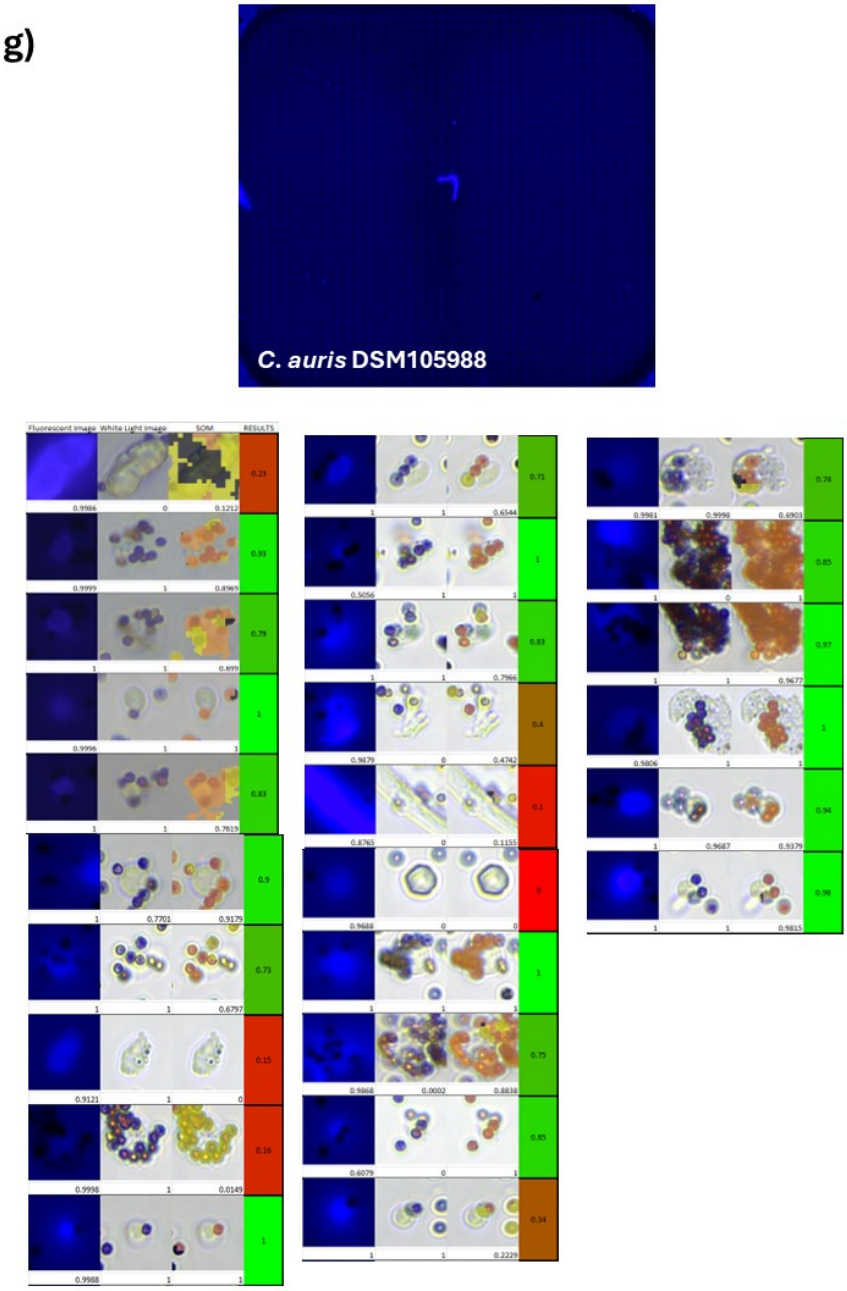


Figure S7. *Candida* spp. samples. A) *C. albicans* 42*, B) C. glabrata* NCPF8018*, C) C. tropicalis* DSM28720, D) *C. dubliniensis* DSM13268 *E) C. parapsilosis* ATCC90875 F) *C. haemulonii* DSM105988, and *G) C. auris* DSM105988. On the top, fluorescence microscope image from the whole montage. On the bottom, fluorescent and white light images, and SOM analysis obtained from the final report generated, including the scoring from the three different detections. Positive or negative result, referred as “candida” or “not candida”, was identified in green or red, respectively.