



## Data Article

## Image dataset of urine test results on petri dishes for deep learning classification

Gabriel Rodrigues da Silva<sup>a</sup>, Igor Batista Rosmaninho<sup>a</sup>,  
 Eduardo Zancul<sup>a,\*</sup>, Vanessa Rita de Oliveira<sup>b</sup>,  
 Gabriela Rodrigues Francisco<sup>b</sup>, Nathamy Fernanda dos Santos<sup>b</sup>,  
 Karin de Mello Macêdo<sup>b</sup>, Amauri José da Silva<sup>b</sup>, Érika Knabben de  
 Lima<sup>b</sup>, Mara Elisa Borsato Lemo<sup>b</sup>, Alessandra Maldonado<sup>b</sup>,  
 Maria Emilia G. Moura<sup>b</sup>, Flávia Helena da Silva<sup>b</sup>,  
 Gustavo Stuani Guimarães<sup>b</sup>

<sup>a</sup> University of São Paulo, School of Engineering, Av. Prof. Luciano Gualberto, 1380, Cidade Universitária, São Paulo-SP, 05508-010, Brazil

<sup>b</sup> Grupo Fleury – Clinical Analysis, Av. Morumbi, 8860, São Paulo-SP, Brazil

## ARTICLE INFO

## Article history:

Received 16 June 2022

Revised 30 January 2023

Accepted 24 February 2023

Available online 2 March 2023

Dataset link: [Image Dataset of Clinical Urine Test Results on Petri Dishes \(Original data\)](#)

## Keywords:

Image Classification

Computational Vision

Urine Test Classification

Petri Dish

## ABSTRACT

Recent advancements in image analysis and interpretation technologies using computer vision techniques have shown potential for novel applications in clinical microbiology laboratories to support task automation aiming for faster and more reliable diagnostics. Deep learning models can be a valuable tool in the screening process, helping technicians spend less time classifying no-growth results and quickly separating the categories of tests that deserve further analysis. In this context, creating datasets with correctly classified images is fundamental for developing and improving such models. Therefore, a dataset of urine test Petri dishes images was collected following a standardized process, with controlled conditions of positioning and lighting. Image acquisition was conducted by applying a hardware chamber equipped with a led lightning source and a smartphone camera with 12 MP resolution. A software application was developed to support image classification and handling. Experi-

\* Corresponding author.

E-mail address: [ezancul@usp.br](mailto:ezancul@usp.br) (E. Zancul).

Social media: [@EZancul](#) (E. Zancul)

enced microbiologists classified the images according to the positive, negative, and uncertain test results. The resulting dataset contains a total of 1500 images and can support the development of deep learning algorithms to classify urine exams according to their microbial growth.

© 2023 The Author(s). Published by Elsevier Inc.  
This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

Specifications Table

Subject	Computer vision and pattern recognition.
Specific subject area	Image classification, urine test results classification.
Type of data	Image.
How the data were acquired	Images were acquired with a Samsung Galaxy Note 9 mobile phone with a 12 MP AF sensor (sensor rate of 4:3, sensor size of 1/3.4" and pixel size of 1.0µm). The cell phone was coupled in a hardware acrylic cubic box of 306 mm wide, height, and depth. The collected images were saved in a cloud database using a mobile application developed in React Native and Firebase.
Data format	The data are in jpeg format.
Description of data collection	The mobile phone was fitted into an acrylic chamber so that the Petri dishes were always placed centrally, at a fixed distance from the camera, and with standard artificial lighting using a 26cm diameter LED ring light (the chamber CAD file is provided in the Mendeley repository). A square cut has been made in the images before saving them to keep the Petri dish centered and eliminate the empty spaces of the rectangular sensor. A microbiologist classified the collected images. Additional information was included as metadata to justify the classification according to a standard.
Data source location	Institution: Fleury S.A. City/Town/Region: São Paulo, São Paulo state Country: Brazil Latitude -23.63765711436827 and longitude -46.64565897448785
Data accessibility	Repository name: Image Dataset of Clinical Urine Test Results on Petri Dishes Data identification number: <a href="https://doi.org/10.17632/xrbtd74pfj.3">10.17632/xrbtd74pfj.3</a> Direct URL to data: <a href="http://dx.doi.org/10.17632/xrbtd74pfj.3">http://dx.doi.org/10.17632/xrbtd74pfj.3</a> [1]

Value of the Data

- The dataset provides labeled images of urine exam results, addressing this field's lack of image datasets.
- The data can be applied by researchers and practitioners aiming to develop novel solutions to classify microbiology analysis results based on image recognition of bacterial colonies' growth on Petri dishes.
- This dataset can contribute to the creation of novel deep learning algorithms that enable the classification of clinical urine cultures in agar plates.
- The dataset can be applied to evaluate the performance of alternative agar plates' classification algorithms based on image recognition.
- The procedures, hardware, and software developed to collect the dataset can be adapted to similar contexts to construct Petri dishes images datasets.

1. Data Description

High-quality digital images of the bacterial cultures on solid agar plates are triggering a new digital revolution in the microbiology field [2]. Novel applications using deep learning algorithms can support task automation in clinical laboratories, not only helping technicians spend less time classifying no-growth results and focusing on positive or complex specimens that deserve further analysis [3,4], as well as leading to more efficient and precise diagnostics [2,5]. Examples of the use of computer vision combined with artificial intelligence techniques in laboratory practice include the detection of hemolyzed samples and parasites in blood tests [6,7].

However, medical image analysis usually requires sets of supervised data, which are often difficult or expensive to acquire [8]. To face the lack of classified images [8], a dataset is created by collecting and classifying photos of urinalysis tests in Petri dishes. The files are labeled according to the dataset structure, consisting of three subdirectories: Positive, Negative, and Uncertain. Each subdirectory contains image files in jpeg format and annotations in an XLSX file about the respective result of the exams. The distribution of the data in each category is represented in Table 1.

The tests' results were based on uroculture, as urines were seeded in plates of chromogenic medium (CPS elite - Biomerieux), in a quantitative way. The plates were initially incubated for 24h at 35°C for the first screening, and those that remained negative were incubated for another 24h, totaling 48h of incubation.

A specific description of each test result is provided below. For screening, cultures with growth of the same colonies greater than or equal to 2 Colony-Forming Units (CFU) were considered Positive (example in Fig. 1).

**Table 1**  
Exams results and respective quantity of images in the dataset.

Exam result	Number of images
Positive	498
Negative	500
Uncertain	502



**Fig. 1.** Positive sample image.



Fig. 2. Negative sample image.

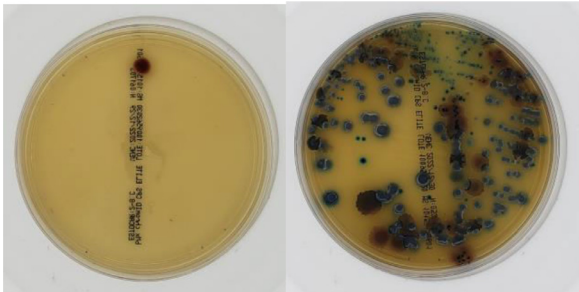


Fig. 3. Uncertain sample images with growths of 1 CFU (left) and mixed colonies (right).

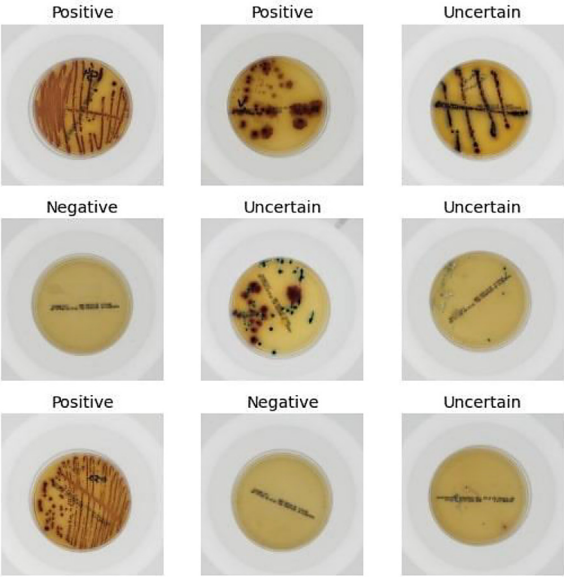
Cultures without growth were considered Negative (example in Fig. 2).

Cultures with growth of 1 CFU or growth of mixed colonies were considered Uncertain (example in Fig. 3).

In total, the dataset has 1500 images with  $3024 \times 3024$  pixel resolution each, which is the original camera output resolution considering a central square cut to convert the sensor's ratio from 4:3 to 1:1. Sample images are shown in Fig. 4.

File sizes range from 600 KB to 1 MB. Although larger files imply greater storage and processing capacity, it was decided to maintain the original image resolution since there are positive results with colorless microbial colonies (Fig. 5). In these situations, low image resolutions may lead to incorrect ratings of these tests as false negatives. This is a consequence of the high variability in bacterial growing patterns that can make a relatively trivial task for microbiologists a very complex task for a machine [2].

The annotations are in XLSX format and bring additional information about each image. The importance of this field is mainly related to bringing more knowledge about the exams, allowing further analyses, such as in the identification of polymicrobial growths. Table 2 provides a summary of the classification parameters and respective labels.



**Fig. 4.** Sample images.



**Fig. 5.** Positive test with colorless microbial colonies.

**Table 2**  
Classification parameters.

Label	Result	Annotation
A	Positive	2 a 99.000 CFU/mL
B	Positive	> 100.000 CFU/mL
C	Negative	No growth
D	Uncertain	1 CFU/mL
E	Uncertain	Polymicrobial growth

Every image in the dataset was labeled according to the classification parameters, as described in [Table 2](#).

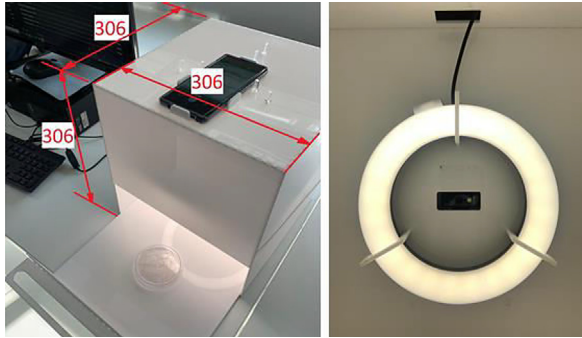


Fig. 6. Arrangement of the system with its dimensions and ring light position.

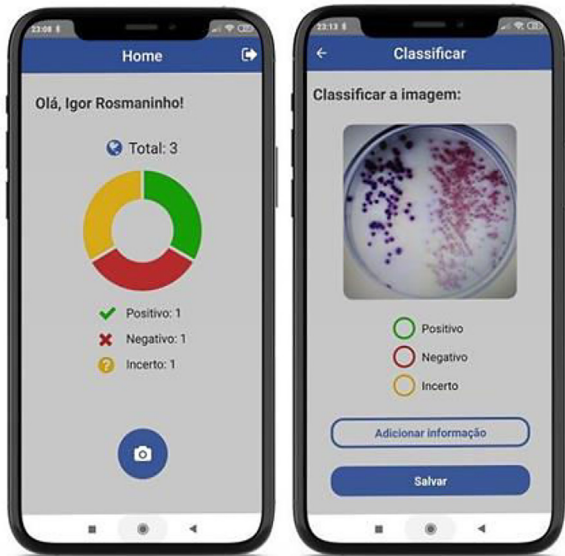


Fig. 7. Screenshots of the image acquisition and classification mobile application.

2. Experimental Design, Materials and Methods

A specific image acquisition system was developed to support data collection following a standardized process. The system is composed of a tailored hardware chamber integrated with a smartphone providing the camera and a software application.

The chamber consists of a white acrylic cubic box made to set a standard environment for the photos, with appropriate conditions of positioning and lighting. The chamber has a drawer with a positioning swell so that the Petri dishes can be easily positioned and removed. White acrylic 3mm thick sheets were used for chamber manufacturing. The light source is a 26 cm outer diameter ring light with 3800 lumens. The hardware was designed to reduce the interference of ambient light, which could generate different lighting conditions for each image (e.g., images with parts of the plate with the absence of light or the presence of shadows) [9,10]. The CAD file of the chamber is also available in the Mendeley data repository.

The camera used is embedded in a smartphone (Samsung Galaxy Note 9). The camera has 12 MP resolution (AF sensor, sensor rate of 4:3, sensor size of 1/3.4" and pixel size of 1.0µm). The

high-resolution cameras embedded in modern mobile phones can be used for image collection for computer vision solutions. A wide range of possibilities of its use for sensing systems, which includes the acquisition of Petri dishes' photos, is presented by Grossi [9]. In order to avoid non-uniform camera characteristics, the same smartphone model was used to capture all the images. Fig. 6 presents the system layout and main dimensions.

A mobile software application was developed to support image acquisition. The mobile application was developed using React Native as front-end technology and Firebase as back-end-as-a-service. The purpose of the application is to take pictures of the Petri dishes, provide a manual classification function for the user to define the state of the culture among the available labels (Positive, Negative, and Uncertain), and upload the images into the database. The app also allows the inclusion of complementary data for each image in a free text field. Image metadata, including date, time, user ID, is also stored. Moreover, the app provides the number of images that have already been acquired. Fig. 7 presents the classification screen of the mobile software application.

To obtain the final result of this dataset, no filtering or other type of pre-processing was performed.

Beyond the generated dataset, the image acquisition methods employed in this research may serve as a reference for other future works.

## Ethics Statements

The dataset is a result of secondary research using biospecimens not collected specifically for this study. The samples are anonymized and do not contain any identifiable private information. Ethics Committee approval 57858722.0.0000.5474 may/2022.

## CRediT Author Statement

**Gabriel Rodrigues da Silva:** Writing – Original Draft, Software; **Igor Batista Rosmaninho:** Writing – Original Draft, Software; **Eduardo Zancul:** Conceptualization, Supervision, Writing – review & editing, Project administration; **Vanessa Rita de Oliveira:** Investigation, Resources; **Gabriela Rodrigues Francisco:** Investigation, Resources; **Nathamy Fernanda dos Santos:** Investigation, Resources; **Karin de Mello Macêdo:** Investigation; **Amauri José da Silva:** Investigation; **Erika Knabben de Lima:** Investigation; **Mara Elisa Borsato Lemo:** Supervision; **Alessandra Maldonado:** Supervision; **Maria Emilia G Moura:** Supervision; **Flávia Helena da Silva:** Conceptualization, Supervision, Writing – review & editing; **Gustavo Stuani Guimarães:** Supervision.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data Availability

[Image Dataset of Clinical Urine Test Results on Petri Dishes \(Original data\)](#) (Mendeley Data).

## Acknowledgments

The authors would like to thank: Marcos Vinicius Ramos da Silva for supporting the hardware manufacturing; Maria Carolina Pintão, Helena Ogata and Ana Paula Leones for R&D support;

Clinical Analysis' team for the investigation support. Funding: This work was partially supported by Ocean R&D Program and Grupo Fleury R&D.

## References

- [1] G. da Silva, I. Rosmaninho, E. Zancul, V. Oliveira, G. Francisco, N. Santos, K. Macedo, A. Silva, E. Lima, M. Lemo, A. Maldonado, M.E. Moura, F. Silva, G. Guimarães, Image Dataset of Clinical Urine Test Results on Petri Dishes, Mendeley Data V3 (2023) doi:[10.17632/xrbtd74pfj.3](https://doi.org/10.17632/xrbtd74pfj.3).
- [2] A. Ferrari, S. Lombardi, A. Signoroni, Bacterial colony counting with Convolutional Neural Networks in Digital Microbiology Imaging, Pattern Recognit. 61 (2017) 629–640, doi:[10.1016/j.patcog.2016.07.016](https://doi.org/10.1016/j.patcog.2016.07.016).
- [3] M.L. Faron, B.W. Buchan, H. Samra, N.A. Ledebøer, Evaluation of WASPLab Software To Automatically Read chromID CPS Elite Agar for Reporting of Urine Cultures, J. Clin. Microbiol. 58 (1) (2019), doi:[10.1128/JCM.00540-19](https://doi.org/10.1128/JCM.00540-19).
- [4] A. Egli, J. Schrenzel, G. Greub, Digital microbiology, Clin. Microbiol. Infect. 26 (10) (2020) 1324–1331, doi:[10.1016/j.cmi.2020.06.023](https://doi.org/10.1016/j.cmi.2020.06.023).
- [5] Majchrowska, S., Pawłowski, J., Guła, G., Bonus, T., Hanas, A., Loch, A., Pawlak, A., Roszkowiak, J., Golan, T., Drulis-Kawa, Z. (2021). AGAR a microbial colony dataset for deep learning detection. [aRxiv2108.01234.pdf](https://arxiv.org/abs/2108.01234)
- [6] K.M. Lopes, F.H. da Silva, A.S. Gil Maldonado, S.A. Santiago, T.A. Pires, C.M. Ferrer, S.J. Mena, M.E.G. Moura, P.T. Domingues, L.M. Kawakami, E de Senzi Zancul, Portable Device for Measuring Blood Test Hemolyzed Samples Based on Computer Vision and Neural Network, J. Med. Devices 13 (2) (2019) 021004, doi:[10.1115/1.4043078](https://doi.org/10.1115/1.4043078).
- [7] M.C.C. Morais, D. Silva, M.M. Milagre, M.T.de Oliveira, T. Pereira, J.S. Silva, L.da F. Costa, P. Minoprio, R.M.C. Junior, R. Gazzinelli, M. de Lana, H.I Nakaya, Automatic detection of the parasite Trypanosoma cruzi in blood smears using a machine learning approach applied to mobile phone images, PeerJ 10 (2022) e13470, doi:[10.7717/peerj.13470](https://doi.org/10.7717/peerj.13470).
- [8] P. Andreini, S. Bonechi, M. Bianchini, A. Mecocci, F. Scarselli, Image generation by GAN and style transfer for agar plate image segmentation, Comput. Methods Programs Biomed. 184 (2020) 105268, doi:[10.1016/j.cmpb.2019.105268](https://doi.org/10.1016/j.cmpb.2019.105268).
- [9] M. Grossi, A sensor-centric survey on the development of smartphone measurement and sensing systems, Measurement 135 (2019) 572–592, doi:[10.1016/j.measurement.2018.12.014](https://doi.org/10.1016/j.measurement.2018.12.014).
- [10] C.S. Costa, E.C. Tetila, G. Astolfi, D.A. Sant'Ana, M.C. Brito Pache, A.B. Gonçalves, V.A. Garcia Zanoni, H.H. Picoli Nucci, O. Diemer, H. Pistori, A computer vision system for oocyte counting using images captured by smartphone, Aquacult. Eng. 87 (2019) 102017, doi:[10.1016/j.aquaeng.2019.102017](https://doi.org/10.1016/j.aquaeng.2019.102017).