

This readme file was modified on 2024-03-06 by Lucia Maddalena

GENERAL INFORMATION

Title of Dataset: ALFI

Author/Associate or Co-investigator Information

Name: Federica Polverino
ORCID: 0000-0002-0048-6122
Institution: National Research Council (CNR), Institute of Molecular Biology and Pathology
Address: c/o Sapienza University of Rome,
Via degli Apuli 4, 00185 Roma, Italy
Email: fede.polverino@gmail.com

Author/Alternate Contact Information

Name: Lucia Maddalena
ORCID: 0000-0002-0567-4624
Institution: National Research Council (CNR), Institute for High-Performance Computing and Networking (ICAR)
Address: Via P. Castellino 111, 80131 Napoli, Italy
Email: lucia.maddalena@cnr.it

Date of data collection: 2012 - 2022

Geographic location of data collection: Roma, Italy

Information about funding sources that supported the collection of the data:

This work has been partially funded by the BiBiNet project (H35F21000430002) within POR-Lazio FESR 2014-2020.

It was carried out also within the activities of the authors as members of the INdAM Research group GNCS and the ICAR-CNR INdAM Research Unit.

The work of M.R.G. was conducted within the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE).

The microscopy experiments were performed at the IBPM Nikon Reference Centre and have been partially supported by POR-Lazio FESR 2014-2020 (N. A0375-2020-36597 to F.D. and G.G.) and by AIRC (IG-2021 ID:25648 to G.G.).

We thank L. Lanzetti, Institute for Cancer Research at Candiolo, Italy, for the kind gift of U2OS H2B-GFP and RFP-alpha-tubulin cells. This work has been partially supported by the European Union PON "Ricerca e Innovazione 2014-2020" FESR (Grant no. ARS01_00734, project QUANCOM), POR-Lazio FESR 2014-2020 (project H35F21000430002 to M.R.G. and project A0375-2020-36597 to F.D. and G.G.), AIRC (IG-2021 ID:25648 to G.G.), and Consiglio Nazionale delle Ricerche - CNR (project DBA.AD005.225-NUTRAGE-FOE2021). It was carried out also within the activities of the authors as members of the INdAM Research group GNCS and the ICAR-CNR INdAM Research Unit. The work of M.R.G. was conducted within the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE). The implementation of the microscopy infrastructure was supported by MUR (Ministry of University and Research) funds to the PON project "IMPARA, Imaging from molecules to the preclinics". The microscopy experiments were performed at the IBPM Nikon Reference Centre. L.M. and L.A. would like to thank R. Mattiello and S. Sada for the technical support.

SHARING/ACCESS INFORMATION

Licenses/restrictions placed on the data: CC-BY

Links to publications that cite or use the data: <https://doi.org/10.1038/s41597-023-02540-1>

L. Antonelli, F. Polverino, A. Albu, A. Hada, I.A. Asteriti, F. Degrassi, G. Guarguaglini, L. Maddalena, M.R. Guarracino,
ALFI: Cell cycle phenotype annotations of label-free time-lapse imaging data from cultured human cells,
Scientific Data, accepted August 2023

Data was not derived from other sources.

Recommended citation for this dataset <https://doi.org/10.6084/m9.figshare.c.6436958.v1>: see <https://doi.org/10.1038/s41597-023-02540-1>.

DATA & FILE OVERVIEW

Data is divided into three folders:

1) Folder "Data&Annotations" includes images and annotations for all the sequences.

1.1) For sequences MI01 through MI08 (including annotations for both Task 1 and Task 2), the content is, e.g.,

```
|— Data&Annotations
|   |— MI01
```

```

| | | | Images
| | | | | I_MI01_0001.png
| | | | | ...
| | | | | I_MI01_0069.png
| | | | | MI01_DTLTruth.csv
| | | | | MI01_PhenoTruth.csv
| | | | | Masks
| | | | | | M_MI01_0001.png
| | | | | | ...
| | | | | | M_MI01_0069.png

```

For each j , $I_MI01_j.png$ is the j -th image of video MI01 and $M_MI01_j.png$ is the corresponding ground truth segmentation mask.

The MI01_DTLTruth.csv file includes annotation information consisting of

- a) Number of sequence image (ImNo),
- b) Cell ID (ID),
- c) Cell class (Class, either Interphase or Mitosis),
- d) Bounding box, specified by its upper left corner (xmin, ymin) and its dimensions (width, height), and
- e) Parent cell ID (Parent).

The MI01_PhenoTruth.csv file includes annotation information consisting of

- a) Number of sequence image (ImNo),
- b) Cell ID (ID),
- c) Cell class (Class, including EarlyMitosis, LateMitosis, CellDeath, and Multipolar), and
- d) Bounding box, specified by its upper left corner (xmin, ymin) and its dimensions (width, height).

1.2) For sequences CD01 through CD09 and TP01 through TP12 (including only annotations for Task 2), the content is similar, but there is no "Masks" subdirectory nor "*PhenoTruth.csv" file.

2) Folder "UseExamples" includes two usage examples and two directories including their output, in the form of Matlab/Python scripts.

```

| | | | UseExamples
| | | | | UseExample1.m
| | | | | UseExample1.py
| | | | | UseExample1.py.readme.txt
| | | | | UseExample1Out
| | | | | | TP06BBannots
| | | | | | | TP06_py_WithBB_0001.png
| | | | | | | ...
| | | | | | | TP06_py_WithBB_0016.png
| | | | | | | TP06_WithBB_0001.png
| | | | | | | ...
| | | | | | | TP06_WithBB_0016.png
| | | | | UseExample2.m
| | | | | UseExample2Out
| | | | | | MI01Lineage.png
| | | | | | MI01LineageCrops.png
| | | | | | MI02Lineage.png
| | | | | | MI02LineageCrops.png
| | | | | | MI03Lineage.png
| | | | | | MI03LineageCrops.png
| | | | | | MI04Lineage.png
| | | | | | MI04LineageCrops.png
| | | | | | MI05Lineage.png
| | | | | | MI05LineageCrops.png
| | | | | | MI06Lineage.png
| | | | | | MI06LineageCrops.png
| | | | | | MI07Lineage.png
| | | | | | MI07Lineage(Mitoses).png
| | | | | | MI07LineageCrops.png
| | | | | | MI08Lineage.png
| | | | | | MI08Lineage(Mitoses).png
| | | | | | MI08LineageCrops.png

```

2.1) UseExample1.m is a Matlab script for plotting the bounding box annotations over the original sequence images; the output is included in folder UseExample1Out/TP06BBannots, that contains images TP06_WithBB_0001.png, ..., TP06_WithBB_0016.png.

UseExample1.py is the analogous examples written in Python and UseExample1.py.readme.txt is the related readme file; the output is included in folder UseExample1Out/TP06BBannots, that contains images TP06_py_WithBB_0001.png, ..., TP06_py_WithBB_0016.png.

2.2) UseExample2.m is a Matlab script for plotting the cells lineage for one of the sequences MI01-MI08; the output of this example is the lineage representation and image crops of mitotic cells given in the PNG images included in folder UseExample2Out.

For the user convenience, we provide the output images for all MI* sequences, without need to run it. Please, observe that, for the very crowded sequences MI07 and MI08, we also provide the lineage representation reduced only to Mitoses (e.g., MI07Lineage(Mitoses).png), obtained setting the parameter OnlyMitoses = 1; (line 49 of the script).

3) Folder "Videos" includes the 29 videos stored in ND2 format.

```
└─ Videos
   └─ CD01.nd2
      ⋮
   └─ TP10.nd2
```

METHODOLOGICAL INFORMATION

Description of methods used for collection/generation of data: see <https://doi.org/10.1038/s41597-023-02540-1>.

Methods for processing the data: see <https://doi.org/10.1038/s41597-023-02540-1>.

Instrument- or software-specific information needed to interpret the data: The Use Example scripts were created using Matlab R2020a and Python.