

# ReadMe

---

This is the official data set for the [NeurIPS 2022 competition](#): cell segmentation in multi-modality microscopy images.

## Background

This is an instance segmentation task where each cell has an individual label under the same category (cells). The training set contains both labeled images and unlabeled images. You can only use the labeled images to develop your model but we encourage participants to try to explore the unlabeled images through weakly supervised learning, semi-supervised learning, and self-supervised learning.

The images are provided with original formats, including tiff, tif, png, jpg, bmp... The original formats contain the most amount of information for competitors and you have free choice over different normalization methods. For the ground truth, we standardize them as tiff formats.

Here are some commonly used image readers that you might be interested in:

skimage: <https://scikit-image.org/docs/dev/api/skimage.io.html>

tifffile: <https://github.com/cgohlke/tifffile>

cv2: [https://docs.opencv.org/4.x/db/deb/tutorial\\_display\\_image.html](https://docs.opencv.org/4.x/db/deb/tutorial_display_image.html)

simpleitk: <https://simpleitk.readthedocs.io/en/master/IO.html>

**We aim to maintain this challenge as a sustainable benchmark platform. If you find the [top-performing solutions](#) don't perform well on your dataset, welcome to send us ([neurips.cellseg@gmail.com](mailto:neurips.cellseg@gmail.com))! We will include them in the new testing set!**

## Main difficulties for cell segmentation

There are at least three challenges to develop universal cell segmentation methods

1. Cell images are extremely diverse. There are many factors that can lead to different image appearances, such as the microscopy imaging technologies (e.g., brightfield, fluorescent, phase-contrast, and differential interference contrast), cell types (various tissues and cultured cells), and staining types (e.g., immunostaining, Jenner-Giemsa). Thus, the segmentation model should have enough capacity to handle these diversities. Moreover, to test the model's generalization ability, the testing images will include images from unseen domains during evaluation.
2. Touching objects and various shapes. In many images, the cells are touched together and have various morphologies. Simply using binary mask to represent the cells may not be optimal. Thus, you may need to design new cell representation methods that can separate the touched cells and be robust to different morphologies.
3. Trade-off between segmentation accuracy and efficiency. Please keep in mind that most of the end users of your segmentation model do not have powerful server/desktop to run the model. Thus, the evaluation metrics will focus on both segmentation accuracy and efficiency. Moreover, there will be

some whole-slide images (~10,000x10,000) in the testing set. The evaluation platform has a RAM of 28G and a GPU with of 10GM of memory. Please keep the resource consumption of your model within this constraint.

## Tips to handle the data

During evaluation, the testing images will be predicted by your docker container one-by-one. Thus, multiprocessing cannot obtain gains in the running efficiency metric.

```
docker container run --gpus "device=0" -m 28G --name teamname --rm -v
$PWD/CellSeg_Test:/workspace/inputs/ -v $PWD/teamname_seg:/workspace/outputs/
teamname:latest /bin/bash -c "sh predict.sh"
```

To save the segmentation results, please use `tifffile` and compress the label

```
import os
import tifffile
from skimage import segmentation
input_path = ''
output_path = ''
names = sorted(os.listdir(input_path))

for name in names:
    # 1. read image: img_data = imread(os.path.join(input_path, name))
    # 2. infer the image with your model (the output should be instance mask)
    inst_mask = model(img_data)
    # relabel the mask
    inst_mask, _, _ = segmentaiton.relabel_sequential(inst_mask)
    # save results; Please compress the mask by setting the `compression`
    save_name = name.split('.')[0] + '_label.tiff' # the suffix of the
    segmentation results should be '_label.tiff'
    tifffile.imwrite(join(output_path, save_name), np.int16(inst_mask),
compression='zlib')
```

Some related papers that may help you further understand this task and develop state-of-the-art models.

- U-Net: <https://www.nature.com/articles/s41592-018-0261-2>
- Cellpose: <https://www.nature.com/articles/s41592-020-01018-x>
- Mesmer: <https://www.nature.com/articles/s41587-021-01094-0>

We also developed an out-of-the-box baseline: <https://github.com/JunMa11/NeurIPS-CellSeg>

Taken together, our goal is to work towards a cell segmentation foundation model that can handle all the needs of biologists for microscopy image-based cell segmentation. Looking forward to your solutions!

If you have any questions, feel free to reach out to [NeurIPS.CellSeg@gmail.com](mailto:NeurIPS.CellSeg@gmail.com) or post the questions on the forum at <https://grand-challenge.org/forums/forum/weakly-supervised-cell-segmentation-in-multi->

modality-microscopy-673/.

Hope you enjoy the competition!