It's a generative-adversarial-network-based model called PLocGAN, which could generate protein fluorescence images with quantitative fraction annotation to alleviate the insufficiency of the quantitative fraction of protein expression. The codes is composed of three parts, data preprocessing for single cell IF images, building PLocGAN model to generate quantitative labeled IF images, and applying PLocGAN to the unmixing model for validating the application value of the generated images.

part1 Data preprocess

There are two data sources. One is the real dataset, including subcellular location combination lysosomes & mitochondria, which can be accessed by 'https://murphylab.cbd.cmu.edu/software/2010_PNAS_Unmixing/', and the other, including four combinations, cytosol & nucleoplasm, cytosol & plasma membrane, mitochondria & nucleoplasm, and nucleoli & nucleoplasm, comes from the subcellular section in the Human Protein Atlas (HPA, https://proteinatlas.org). Run the codes step by step to get the single-cell images.

part2 PLocGAN

The 'model.py' and 'network.py' are the base model without contrastive learning module, and the 'model_cl.py' and 'network_cl.py' have the contrastive learning module. ‘options.py’ is the parameter setting code, such as the weight of loss function, batch size, and max epoch, etc. Run .\train.py to train a generative model. And the model will be applied to the unmixing model (part3).

part3 Unmixing model

The model is based on Bestfitting which the first place of 2019 global subcellular location classification competition, can be obtained by 'https://github.com/CellProfiling/HPA-competition-solutions/tree/master/bestfitting'. Run .\run\train_gan.py to get a quantitative prediction model that introduces PLocGAN. Run.\run\train.py to get a baseline if you want to compare to the unmixing model with PLocGAN.