

crYOLO

crYOLO is a deep-learning based particle picking software.

The crYOLO tutorial is available at: [For Relion 3.1 we have a \[Relion 3.1 crYOLO integration wrapper\]\(#\) for automated particle picking using the latest crYOLO generic models.](http://sphire.mpg.de/wiki/doku.php?id=pipeline>window:cryolo</p></div><div data-bbox=)

Loading the environment

CrYOLO is installed using the **CSSB anaconda**.

It can be used in 2 ways:

- **GPU-based:** train and predict
- **CPU-based:** mainly for predicting since the training takes too long on CPUs (version 1.3, can be installed on request)

To use it at CSSB/DESY call:

```
module load cryolo
```

and copy the CSSB templates as printed on the screen.

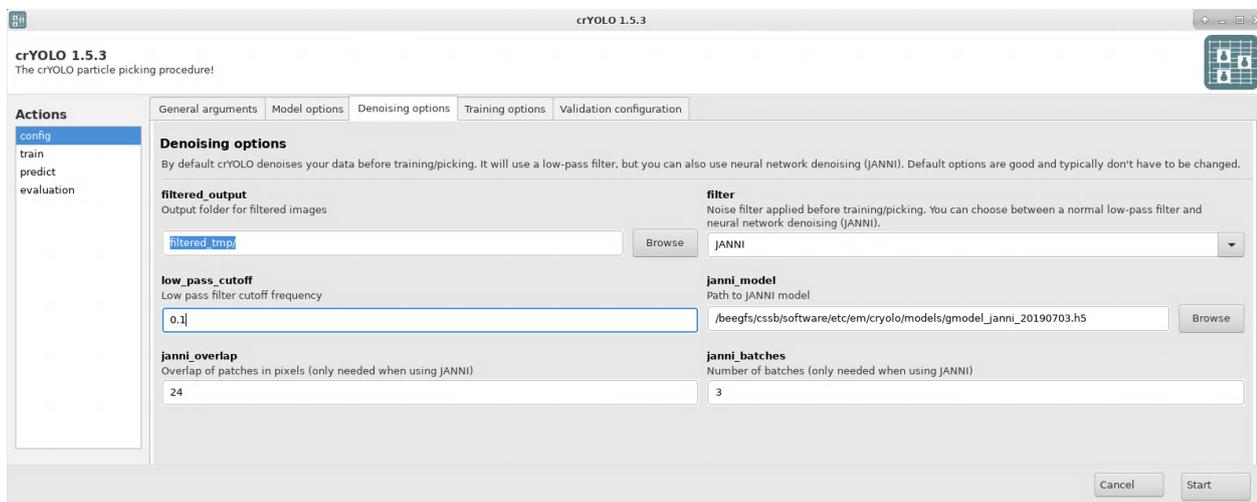
Replace *?:?* with the actual CrYOLO version printed by the *module show* command or the version you would like to use.

CSSB Beginner's Notes for crYOLO >= 1.5

Since crYOLO 1.5 the software comes with a graphical user interface for generating the configuration files but also for running the training and prediction on-the-fly on the local node. The official crYOLO tutorial helps you in using this. Since crYOLO 1.4 it is recommended to the PhosaurusNet instead of the YOLO algorithm in crYOLO. Start it now with:

```
cryolo_gui.py
```

Sometimes it is recommended to use the *JANNI* denoising software. The settings can be seen in the screenshot below.

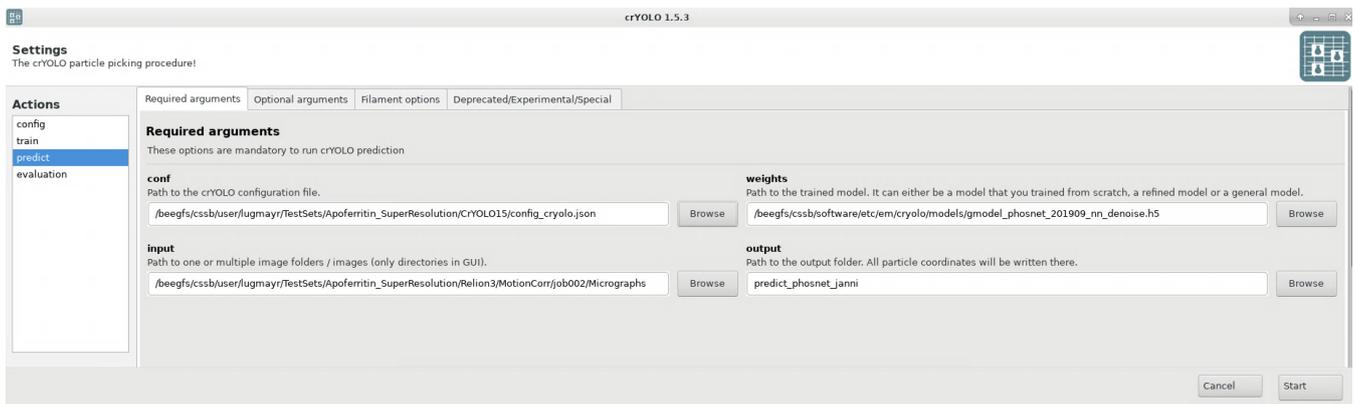


Use the pre-trained *JANNI* model from the following CSSB location and save the config file.

```
/beegfs/cssb/software/etc/em/cryolo/models/gmodel_janni_20190703.h5
```

Instead of training your own model you can try a generic pre-trained model.

So you just have to provide a motion corrected micrograph folder like below:



The pretrained models are at the following CSSB location:

```
/beegfs/cssb/software/etc/em/cryolo/models/gmodel_phosnet_202002_N63.json # config file
/beegfs/cssb/software/etc/em/cryolo/models/gmodel_phosnet_202002_N63.h5 # model file
```

For *JANNI* use the **_nn_*.h5* file.

```
/beegfs/cssb/software/etc/em/cryolo/models/gmodel_phosnet_202003_nn_N63.h5
```

If you want to submit the job to the cluster an updated submission template can be found as follows:

```
cp /beegfs/cssb/software/etc/em/cryolo/1.7/slurm_sbatch_cryolo.sh .
```

Just copy & paste the GUI command using the *Print cmd* button into it and submit it to the cluster using the *sbatch* command.

CSSB Beginner's Notes for crYOLO <= v1.4

This section describes shortly the CSSB specific setup and an example of the integration with *Relion*. The directory naming convention follows the *crYOLO* tutorial. Please read the tutorial first.

Import the micrographs

crYOLO needs motion-corrected single frame micrographs as input. The *Relion* 3.x naming conventions can be confusing (e.g. *Movies* directories instead of the previous *Micrographs*)!

Copy all micrographs as symbolic links from the original location to the *full_data* folder (*crYOLO v1.0, v1.1*).

crYOLO v1.2+ reads only the files from the *full_data* directory and pre-processes them to a temporary *tmp_filtered* directory. So just linking the original directory is enough.

```
cp -s MotionCorr/job002/Movies full_data
```

Prepare the training set

For training its recommended to use 10 micrographs or about 100 particles.

```
mkdir -p train_annotation train_image
cd train_image
for i in `ls ../full_data/*.mrc | head -10`; do cp -s $i .; done
cd ..
```

This statement copies the first 10 files in the *full_data* directory to the local *train_image* directory. You can also use a graphical file manager as alternative.

Pick particles in the training set

You can start picking directly (don't forget to load the environment):

```
cryolo_boxmanager.py
```

Open the files in the folder *train_image*, choose your boxsize and save the picked coordinates as box files in the folder *train_annotation*.

Here is how to pick particles (from the tutorial):

- LEFT MOUSE BUTTON: Place a box
- HOLD LEFT MOUSE BUTTON: Move a box
- CONTROL + LEFT MOUSE BUTTON: Remove a box

If you want to continue, open the files in the folder *train_image* and import the box files by choosing the folder *train_annotation*.

If you have already picked coordinates, copy them into the folder *train_annotation* and do as described in continue above.

Denosing your training data set

If you hardly see your particles for handpicking do the following in the activated crYOLO conda environment on a **GPU** node (*crYOLO v1.4+*):

```
mkdir -p janni_denoised
janni_denoise.py predict train_image janni_denoised /beegfs/cssb/software/etc/em/cryolo/models
/gmodel_janni_20190703.h5 -g 0
```

Then start the *cryolo_boxmanager.py* as described above and use as input micrograph folder *janni_denoised/train_image*.

Please note that the JANNI micrographs should only be used for handpicking, for all other steps use the original micrographs.

Train and build a model file

Prepare the *crYOLO* configuration file and the CSSB submission script as printed in:

```
module show cryolo
```

Edit the provided submission script and uncomment all lines with *#cryolo_train.py* by removing the *#* from the beginning of the line.

Comment all *cryolo_predict.py* by adding a *#* sign at the beginning of the line.

Save and submit the script to the cluster:

```
sbatch cryolo_yolo.sh
```

Once its done the trained model is in the *model_yolo.h5* file. This you could save with the corresponding *config_yolo.json* configuration file for future prediction without training again.

Pick the full dataset with the trained model

Edit the provided submission script and uncomment one line with *#cryolo_predict.py* by removing the *#* from the beginning of the line. The various examples are with different thresholds so if the result is not satisfying rerun the picking with a different one.

Comment all lines with *cryolo_train.py* by adding a *#* sign at the beginning of the line.

Save and submit the script to the cluster:

```
sbatch cryolo_yolo.sh
```

Once the job is finished you could check the amount of picked particles by:

```
wc -l predict_yolo/EMAN/*.box
```

View the picked coordinates

You can use the *crYOLO* boxmanager to view the results. Use **Open image folder:** *full_data* and **Import box files:** *predict_yolo/EMAN* as input.

Copy the picked coordinates to your project folder

To import the coordinates in *Relion* you need to copy the files to your original *Movies* folder (not the *MotionCorr/job002/Movies* folder!):

```
cp -v predict_yolo/STAR/*.star ./Movies
```

Now use the *Relion* import function pattern: ***Movies/*.star***, set your boxsize and extract the particles again.