TOPAZ – PICK MORE OF ANYTHING

TOPAZ: REVOLUTIONIZING PARTICLE PICKING IN CRYO-EM

- **Background:** Cryo-electron microscopy (cryo-EM) is crucial for determining protein structures, but manual particle picking is time-consuming and prone to bias.
- **Topaz:** A particle-picking pipeline using convolutional neural networks (CNNs) trained with a positive-unlabeled (PU) learning method.
- **Advantages:**
  - Efficient and accurate, reducing the need for manual effort.
  - Capable of detecting challenging particles, including small, non-globular, and asymmetric ones.
  - Low false-positive rate and minimal need for post-picking curation.
- **Key Features:** Modular, standalone, free, open-source, and compatible with various cryo-EM software suites.

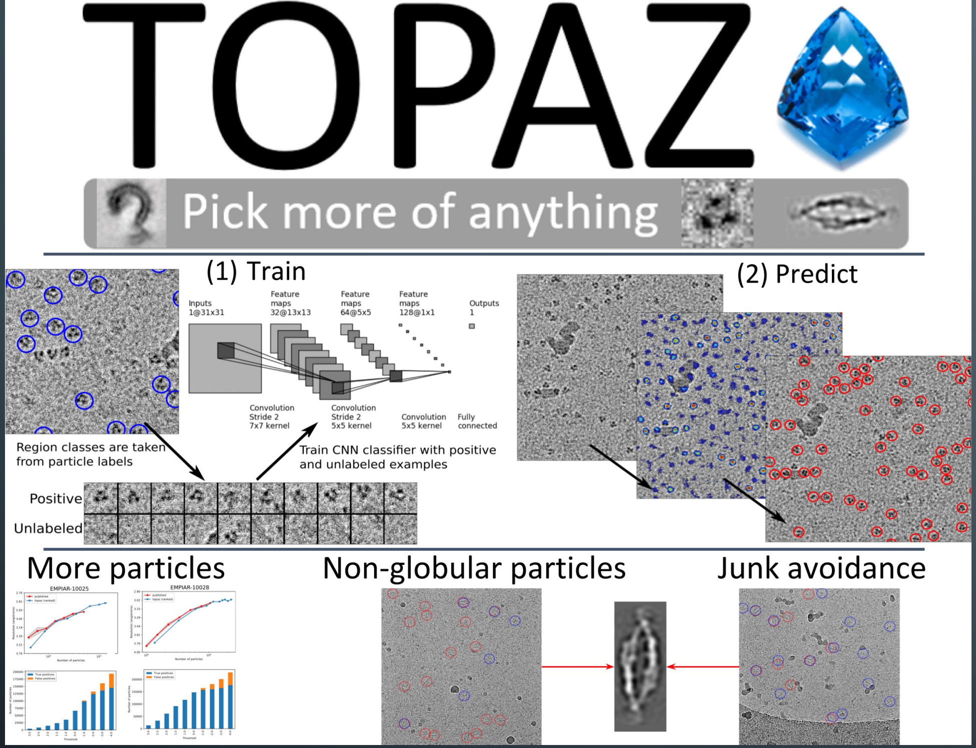
Bepler, T., Morin, A., Rapp, M. *et al.* Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat Methods* **16**, 1153–1160 (2019). https://doi.org/10.1038/s41592-019-0575-8

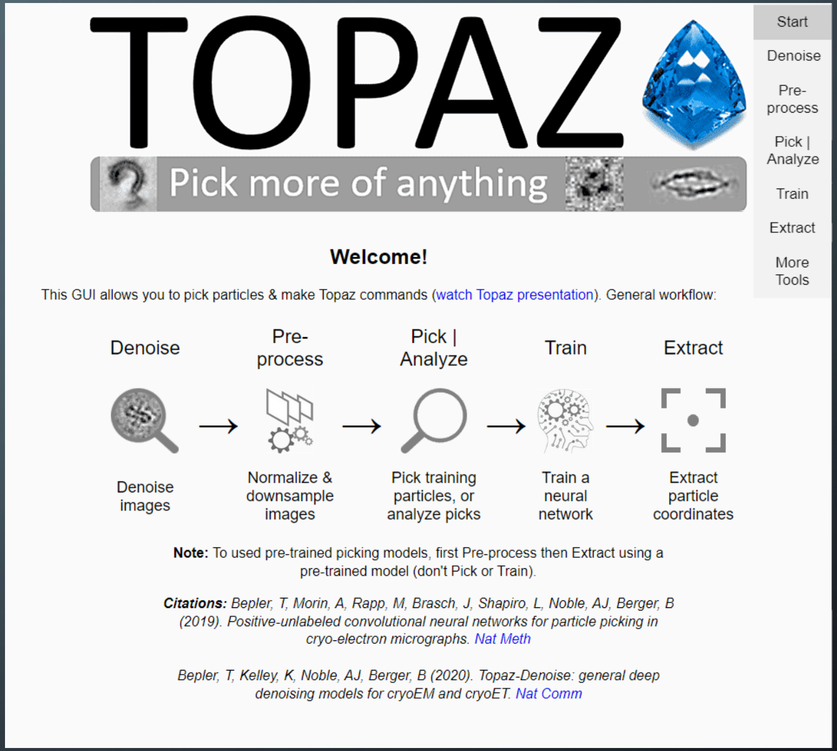
PERFORMANCE AND IMPACT OF TOPAZ IN CRYO-EM

- **High-Resolution Reconstruction:** Topaz enables high-resolution reconstructions with fewer false positives and better representation of particle views.
- **Case Studies:** Demonstrated effectiveness on datasets with difficult particles, such as the Toll receptor and clustered protocadherin.
- **Integration and Efficiency:** Appion Relion and CryoSparc. Runs efficiently on a single GPU computer.

Bepler, T., Morin, A., Rapp, M. *et al.* Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat Methods* **16**, 1153–1160 (2019). https://doi.org/10.1038/s41592-019-0575-8

OVERVIEW AND INTERFACE(S)





Welcome!

This GUI allows you to pick particles & make Topaz commands (watch Topaz presentation). General workflow:

Denoise images → Normalize & downsample images → Pick training particles, or analyze picks → Train a neural network → Extract particle coordinates

Note: To used pre-trained picking models, first Pre-process then Extract using a pre-trained model (don't Pick or Train).

Citations: *Bepler, T, Morin, A, Rapp, M, Brasch, J, Shapiro, L, Noble, AJ, Berger, B (2019). Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. Nat Meth*

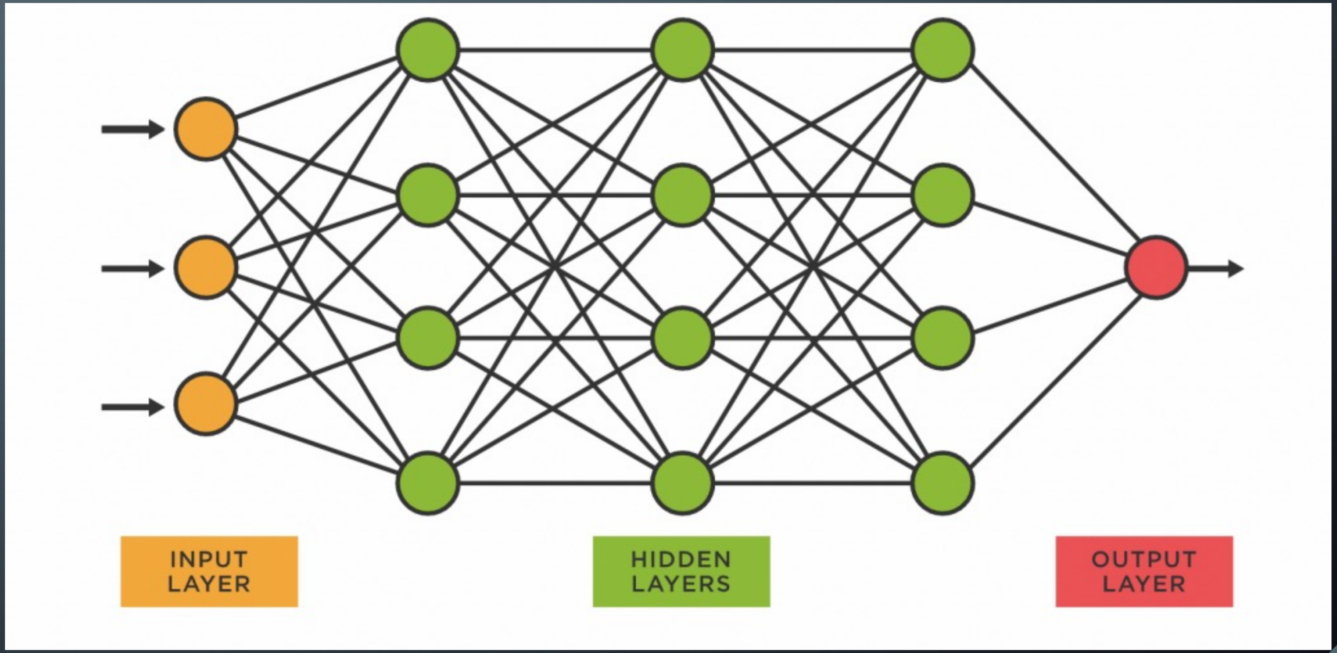
Bepler, T, Kelley, K, Noble, AJ, Berger, B (2020). Topaz-Denoise: general deep denoising models for cryoEM and cryoET. Nat Comm

Bepler, T., Morin, A., Rapp, M. *et al*. Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat Methods* **16**, 1153–1160 (2019). https://doi.org/10.1038/s41592-019-0575-8

CNN ARCITECTURE

Convolutional neuronal networks are used to analyse multi-dimensional data and present it in human readable form. The operate on reducing multidimensionality to a perceivable dimension.

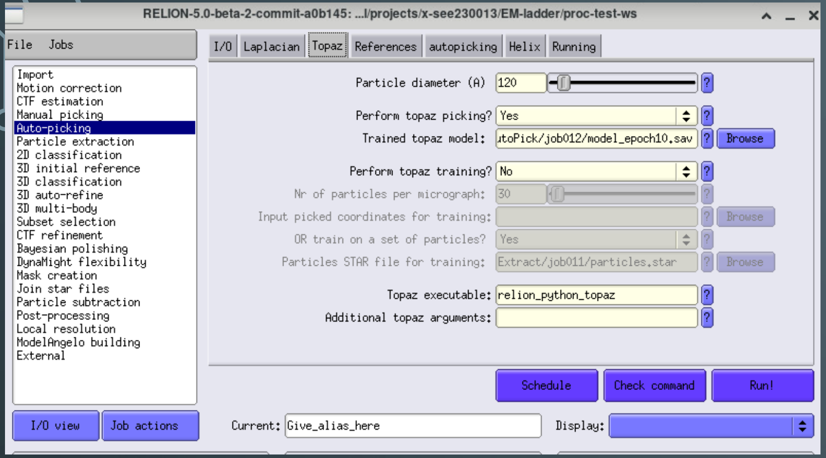
- **Input Layer:** Represents the features of the dataset, with each neuron corresponding to one feature.
- **Hidden Layers:** Multiple layers of neurons that process input features, learning complex patterns and relationships.
- **Output Layer:** Final layer that provides the prediction or classification result, with each neuron representing a possible output category.



DEMO

https://www.tibco.com/reference-center/what-is-a-neural-network

TOPAZ EXAMPLES



RELION-5.0-beta-2-commit-a0b145: .../projects/x-see230013/EM-ladder/proc-test-ws

Particle diameter (A) 120

Perform topaz picking? Yes

Trained topaz model: utoPick/job012/model_epoch10.sav

Perform topaz training? No

Nr of particles per micrograph: 30

Input picked coordinates for training:

OR train on a set of particles? Yes

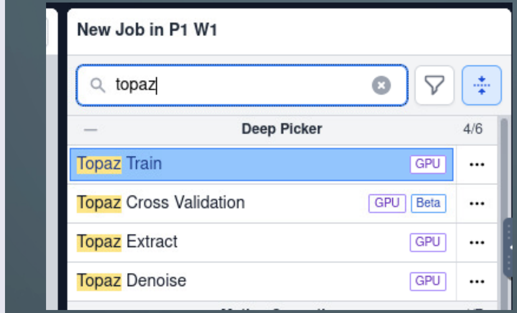
Particles STAR file for training: Extract/job011/particles.star

Topaz executable: relion_python_topaz

Additional topaz arguments:

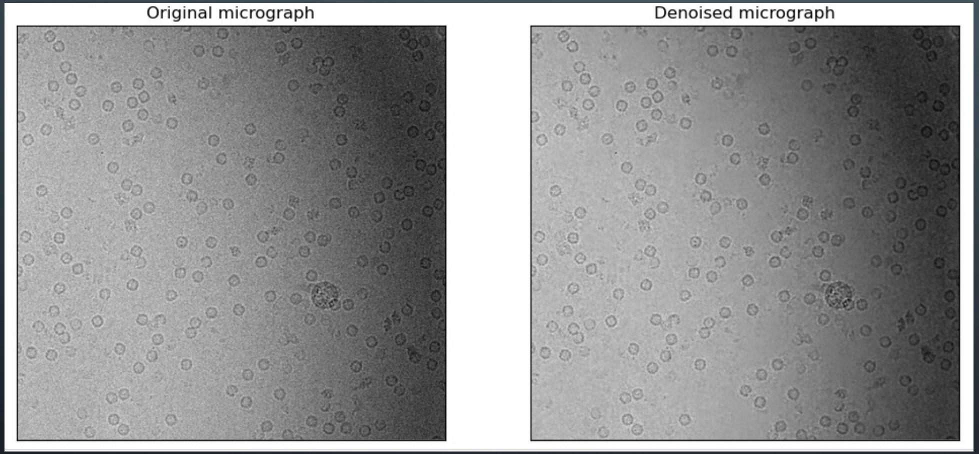
Current: Give_alias_here

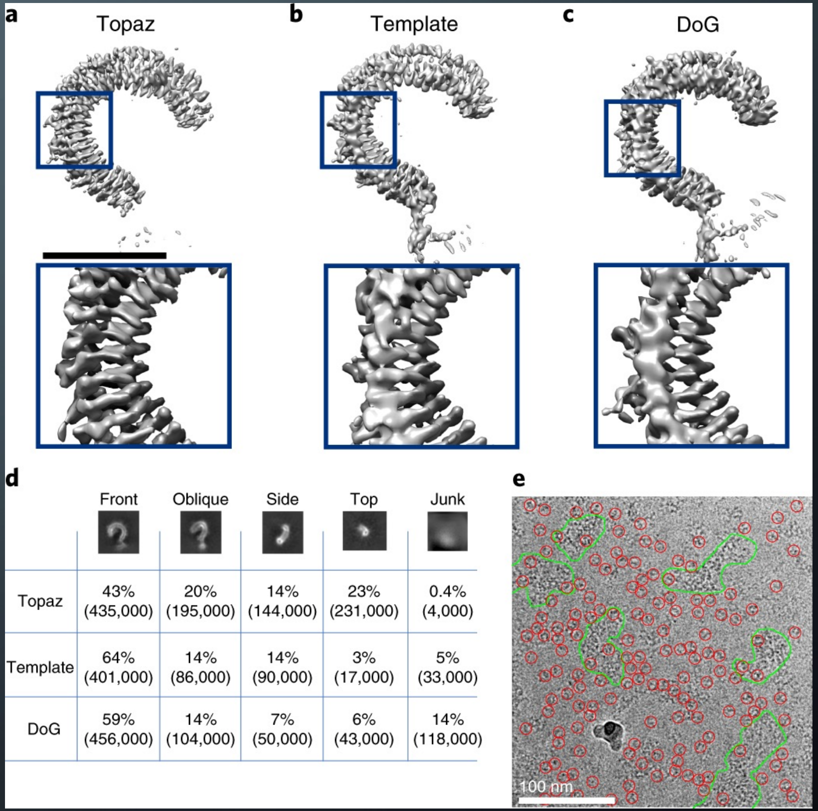
Display:



New Job in P1 W1

topaz





	Front	Oblique	Side	Top	Junk
Topaz	43% (435,000)	20% (195,000)	14% (144,000)	23% (231,000)	0.4% (4,000)
Template	64% (401,000)	14% (86,000)	14% (90,000)	3% (17,000)	5% (33,000)
DoG	59% (456,000)	14% (104,000)	7% (50,000)	6% (43,000)	14% (118,000)

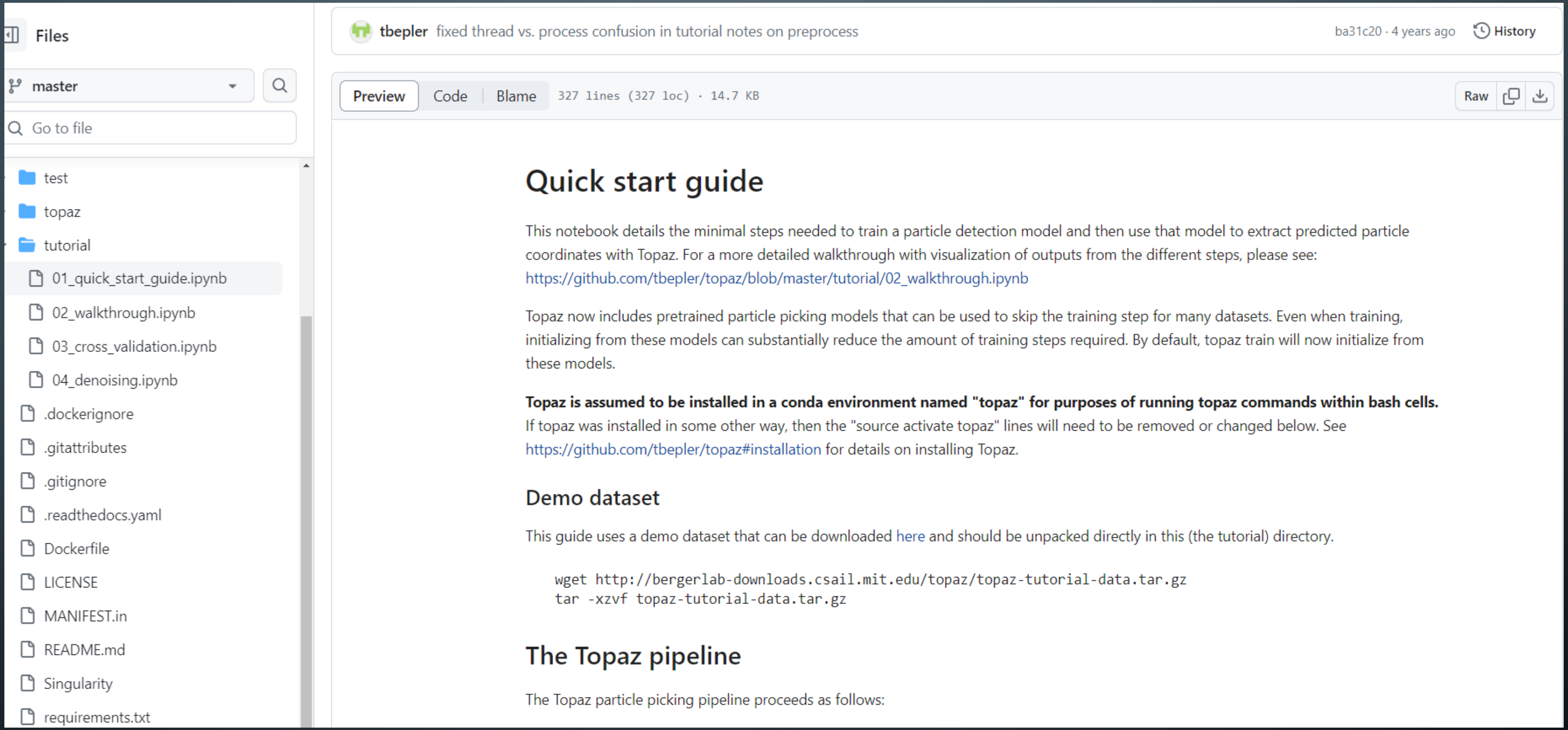
https://www.ebi.ac.uk/empiar/EMPIAR-10261/

TOPAZ ISN'T MAGIC

- What Topaz can do
  - Denoise the mics to enhance contrast
  - Extract particles using own models
  - Get trained on your data
  - Extract particles using trained models
  - Great on single GPU

- What Topaz can't do
  - Filter particles for you
  - Cannot box ptcls, you have to EXTRACT
  - MAGIC

TRY IT YOURSELF



Quick start guide

This notebook details the minimal steps needed to train a particle detection model and then use that model to extract predicted particle coordinates with Topaz. For a more detailed walkthrough with visualization of outputs from the different steps, please see: https://github.com/tbepler/topaz/blob/master/tutorial/02_walkthrough.ipynb

Topaz now includes pretrained particle picking models that can be used to skip the training step for many datasets. Even when training, initializing from these models can substantially reduce the amount of training steps required. By default, topaz train will now initialize from these models.

Topaz is assumed to be installed in a conda environment named "topaz" for purposes of running topaz commands within bash cells. If topaz was installed in some other way, then the "source activate topaz" lines will need to be removed or changed below. See https://github.com/tbepler/topaz#installation for details on installing Topaz.

Demo dataset

This guide uses a demo dataset that can be downloaded here and should be unpacked directly in this (the tutorial) directory.

```
wget http://bergerlab-downloads.csail.mit.edu/topaz/topaz-tutorial-data.tar.gz
tar -xzvf topaz-tutorial-data.tar.gz
```

The Topaz pipeline

The Topaz particle picking pipeline proceeds as follows:

Install it from: https://github.com/tbepler/topaz/blob/master/README.md

https://github.com/tbepler/topaz/blob/master/tutorial/01_quick_start_guide.ipynb