1. Appendix

1.1. Training details

Here we provide additional training details regarding our network. All networks use a single input channel and 48 feature channels in all encode blocks of U-Net. With feature concatenation, all U-Net decode blocks have 96 feature channels. For the classifier, each residual block has 32, 64 and 128 feature channels respectively. Input to the modified sigmoid layer (soft segmentation map) is normalized to $[0, 1]$.

How many training iterations are needed? In general, our framework converges within 300,000 iterations. Due to limited number of training samples, we found out that the model can overfit if the training iterations are too large. Training time takes around 30 minutes to an hour based on the number of iterations used. Training of the model does not require very high computational power. The network takes around 7GB RAM on a NVIDIA TESLA V GPU. It can also work on other NVIDIA GPUs such as 1080Ti.

How to choose α and λ in loss function? We experimentally chose $\alpha = 0.75$ and $\lambda = 0.1$ for our loss function. However, the choice of α can be anywhere between 0.7 to 0.9. It slightly affects the convergence time and the overall performance. However, results usually vary within $1 - 2\%$. For λ, we found out $\lambda = 0.1$ works the best experimentally. In general, our network is able to perform well without too much fine-tuning.

1.2. Evaluation details

Here we provide evaluation details regarding our framework. To obtain particle locations through the segmentation map, we apply non-max suppression with a cut-off threshold of 0.5 and a radius of 16 for the single particle ribosome dataset. For Topaz and CrYOLO, we used a cut-off threshold of -4 and 0.1 respectively. For EMPIAR 10304, we used a cut-off threshold of 0.85 and a radius of 12 for our method, and -6 and 0.05 for Topaz and CrYOLO. For EMPIAR 10499, we used a cut-off threshold of 0.7 and a radius of 20, and -8 and 0.02 for CrYOLO. Note since the output detection probabilities for Topaz is pre-sigmoid layer value, which is the reason for negative values. We used very small values for CrYOLO, as we realized that under extremely low-SNR scenarios, CrYOLO can only detect particles when we have set the confidence level to a very small number. Note all these numbers are chosen after multiple iterations of evaluation to find the optimal performance. In practice, to choose a proper cutoff threshold, we recommend plotting out the distributions of predicted values and set the cutoff threshold to values around top $25\% - 30\%$ of all values.

1.3. Denoising Performance at Different SNR (20% and 60% total frames)

Table 1. Denoising performance (PSNR) for averages obtained from increasing frame fractions compared to the full exposure (without and with application of low-pass filtering).

Compared Against	Method	60%	20%
Full Dose Image	Topaz Denoise	23.24	19.06
	Sequential Learning	23.91	20.04
	Joint Learning w/o Consistency	24.10	21.33
	Joint Learning w/ Consistency	**25.14**	**22.20**
Low Pass-Filtered Image	Topaz Denoise	**25.89**	22.09
	Sequential Learning	24.83	21.13
	Joint Learning w/o Consistency	24.64	20.31
	Joint Learning w/ Consistency	25.30	**22.75**

Note that our performance shows less improvement when SNR increases. This validates our hypothesis that a joint-training framework is more likely to benefit the most when the amount of information is not sufficient; when there is sufficient amount of information already from a single task, the benefits are limited.

1.4. Additional Results

In addition to qualitative results presented in Figure 2 and 3 in the main text, here we present more results for each considered datasets. Figure 1 to 3 are denoising results and Figure 4 to 9 are detection results.
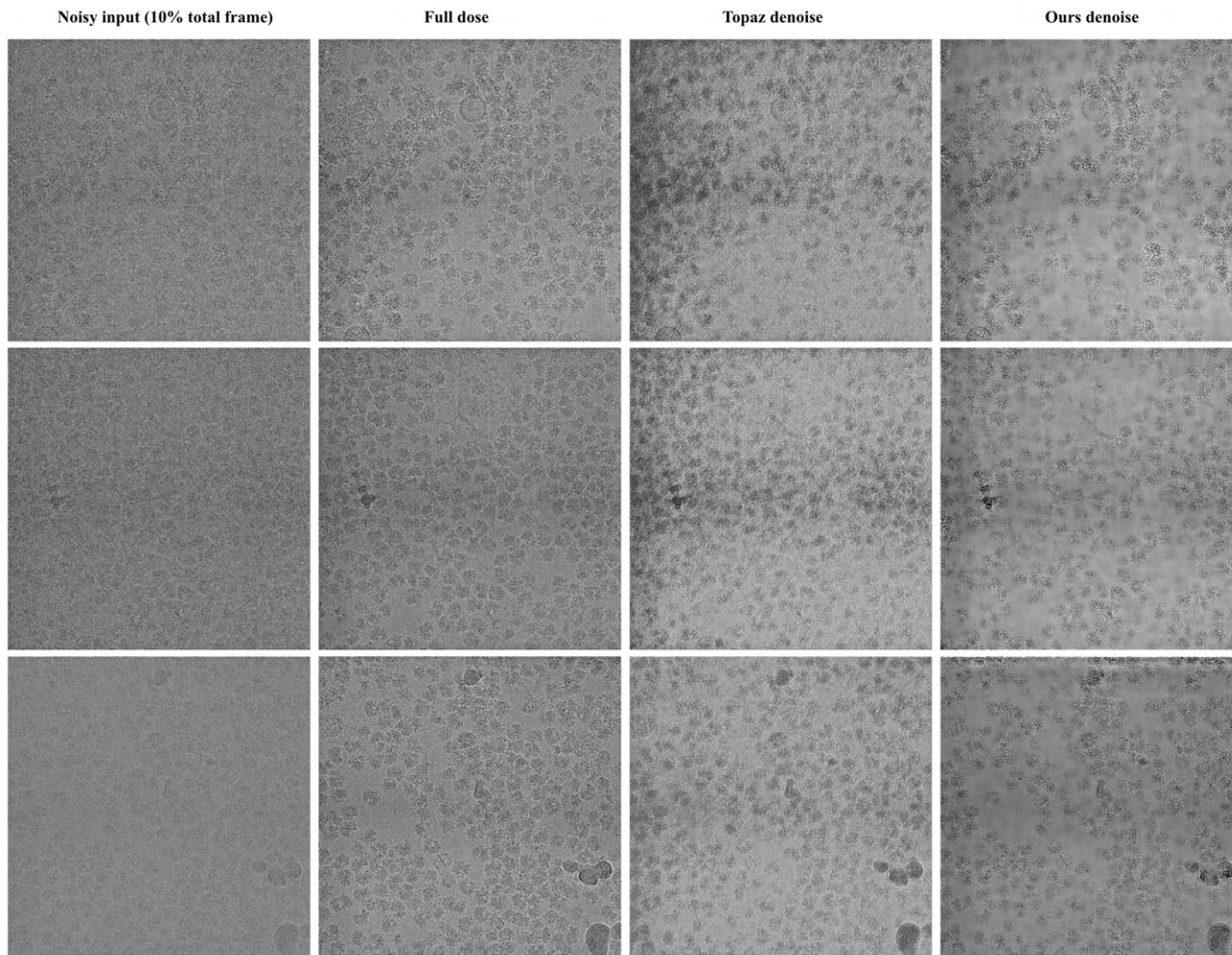

Figure 1. **Additional qualitative results for image denoising of single particle ribosome dataset.** Here we are showing the 10% dose fraction noisy input(1st column), full dose image(2nd column), Topaz denoise(3rd column) and ours denoise(4th column).
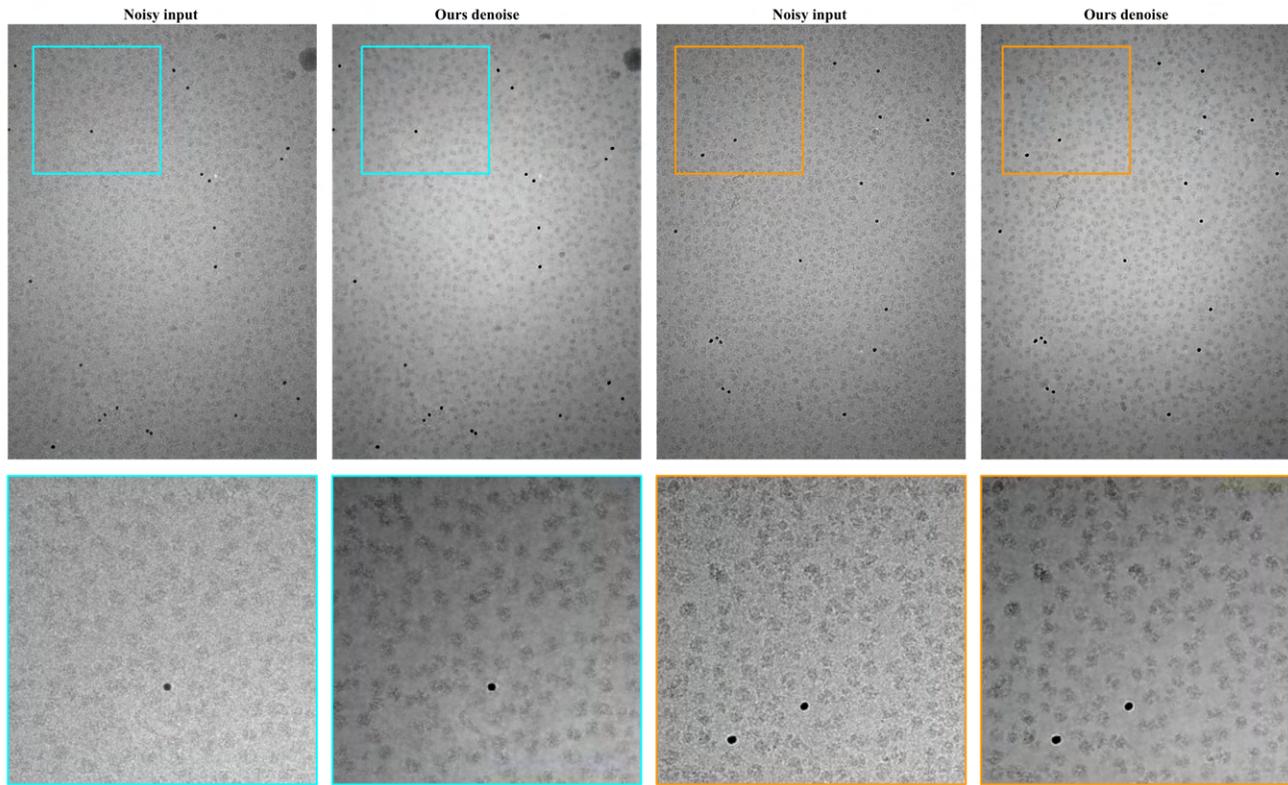
Figure 2. **Additional qualitative results for image denoising of EMPIAR 10304.** Here we are showing the full image. The bottom row is the zoomed-in view of the selected area. In zoomed in views, we can better visualize distinctions between background and particles.
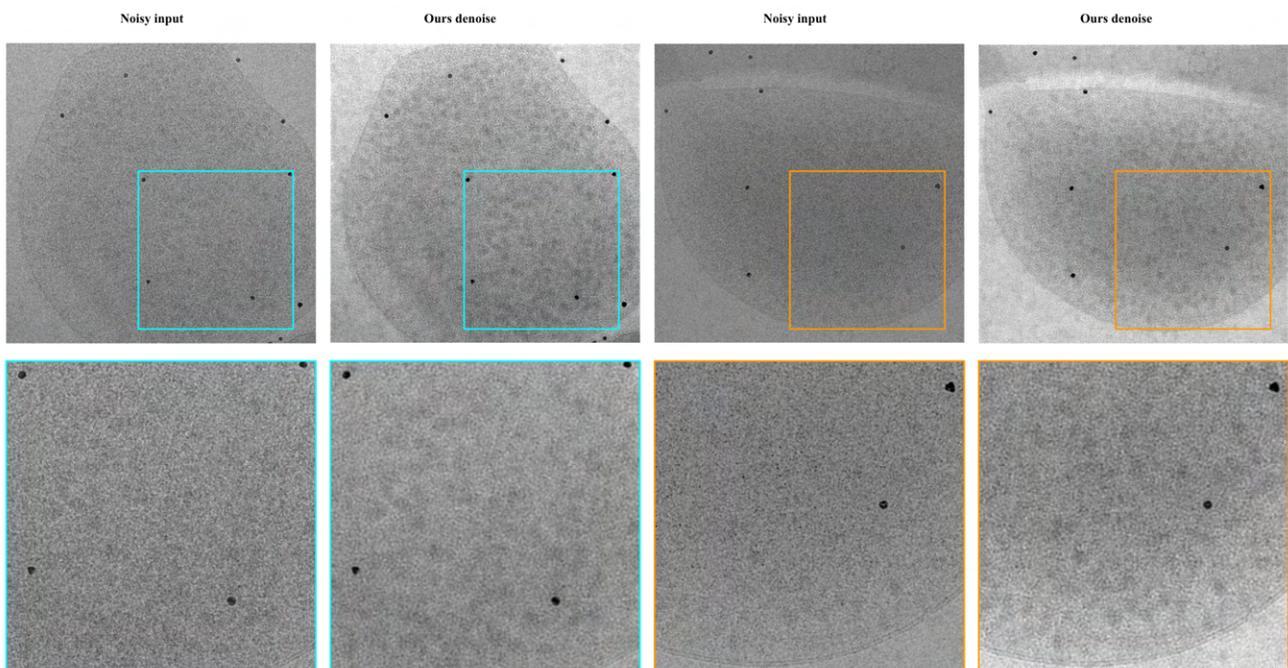

Figure 3. **Additional qualitative results for image denoising of EMPIAR 10499.** Here we are showing the full image. The bottom row is the zoomed-in view of the selected area. In zoomed in views, we can better visualize distinctions between background and particles.

Noisy image **Ours soft segmentation** **Ours w/ consistency**
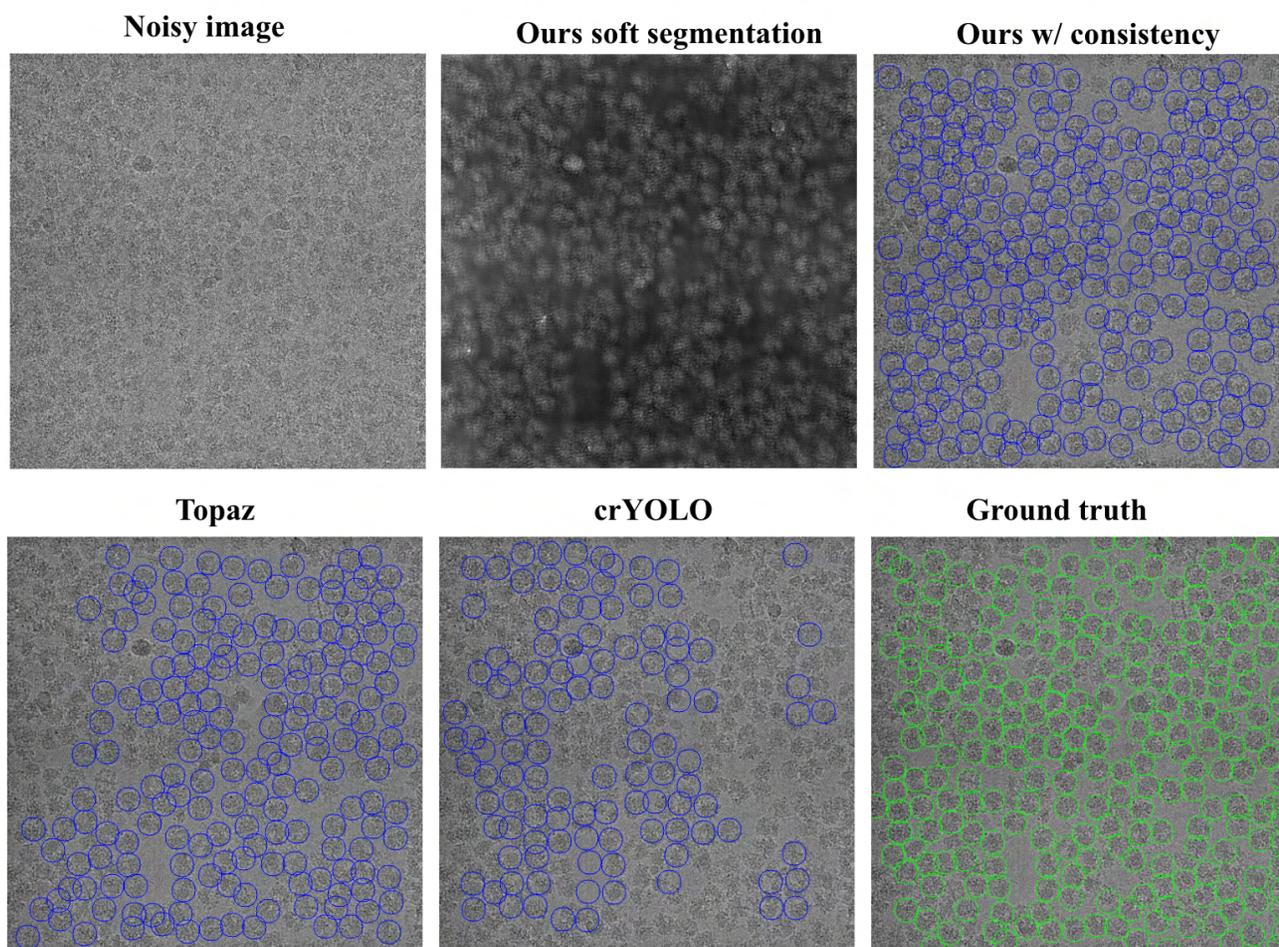
Topaz **crYOLO** **Ground truth**

Figure 4. **Additional qualitative results for particle detection of single particle ribosome dataset.** Here we are showing the segmentation map and our detection output. We also show picking results using Topaz and CrYOLO. Ground truth is obtained by running Topaz and full-dose images.
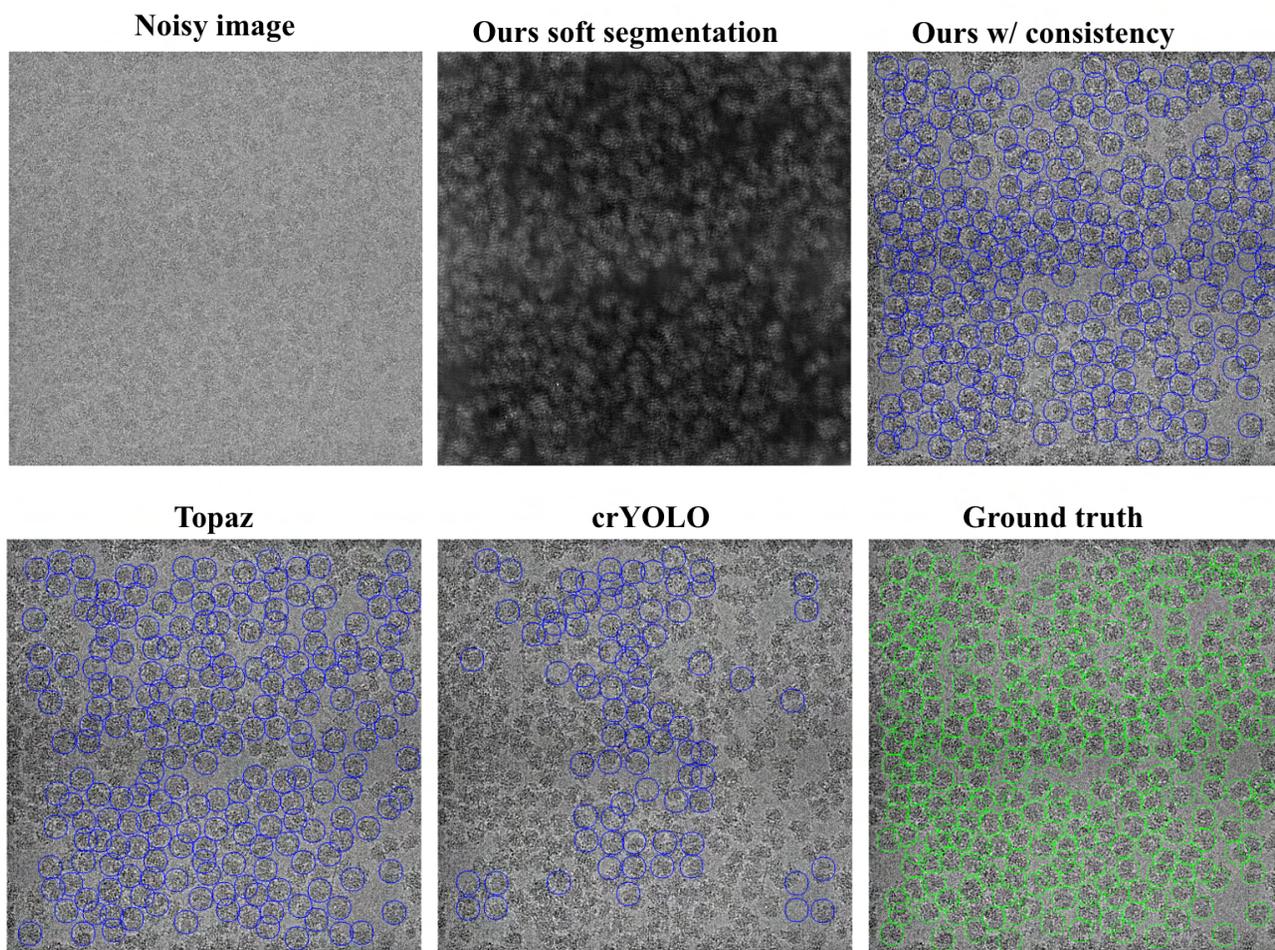
Noisy image   **Ours soft segmentation**   **Ours w/ consistency**

Topaz   **crYOLO**   **Ground truth**

Figure 5. **Additional qualitative results for particle detection of single particle ribosome dataset.** Here we are showing the segmentation map and our detection output. We also show picking results using Topaz and CrYOLO. Ground truth is obtained by running Topaz and full-dose images.
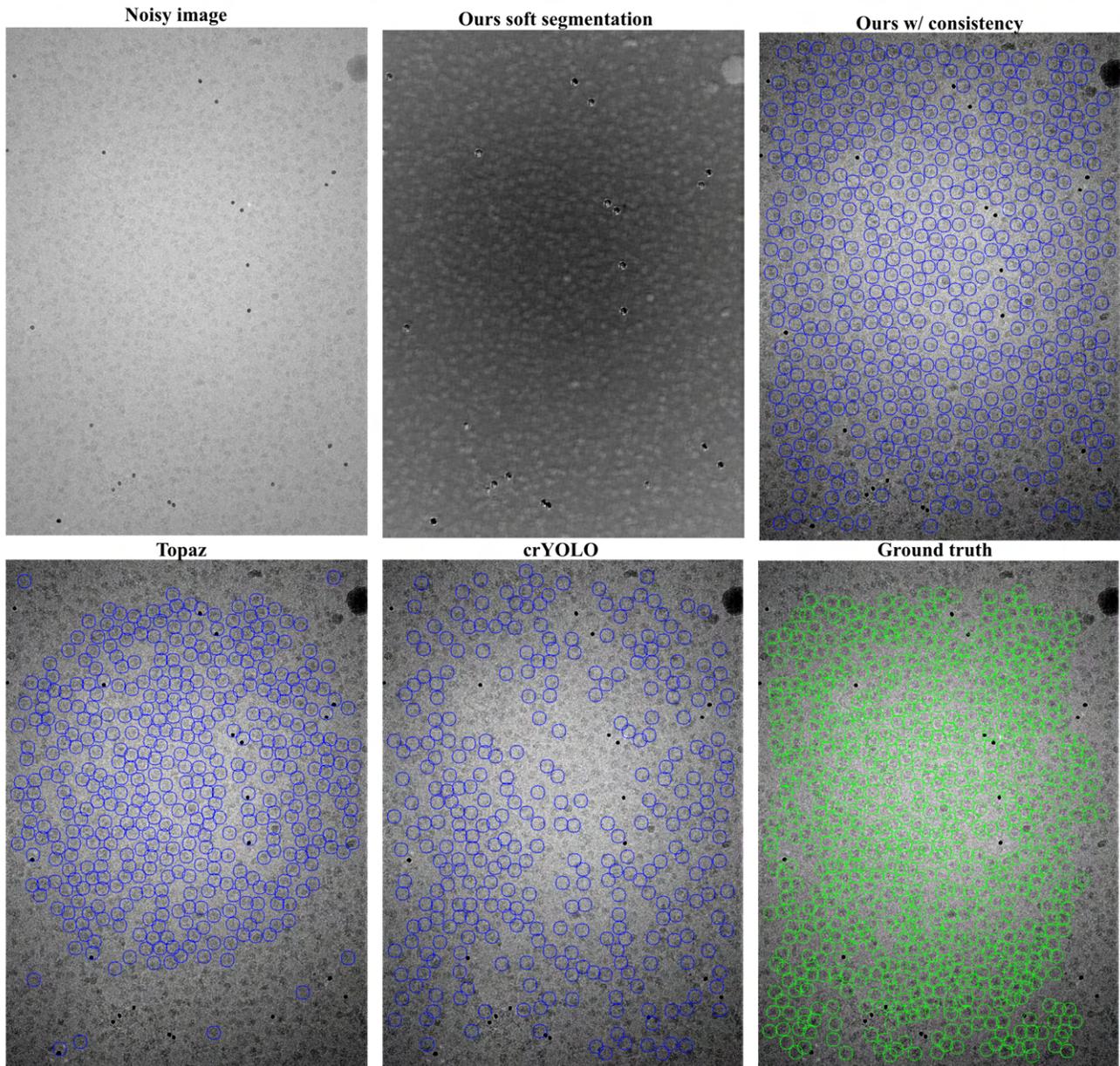
Figure 6. **Additional qualitative results for particle detection of EMPIAR 10304.** Here we are showing the full image with detection result. Note our method is able to avoid contamination areas and gold beads. Note: some of the particles near edges of the image are not getting labeled in manual picking. Our method is able to detect these particles near boundaries.
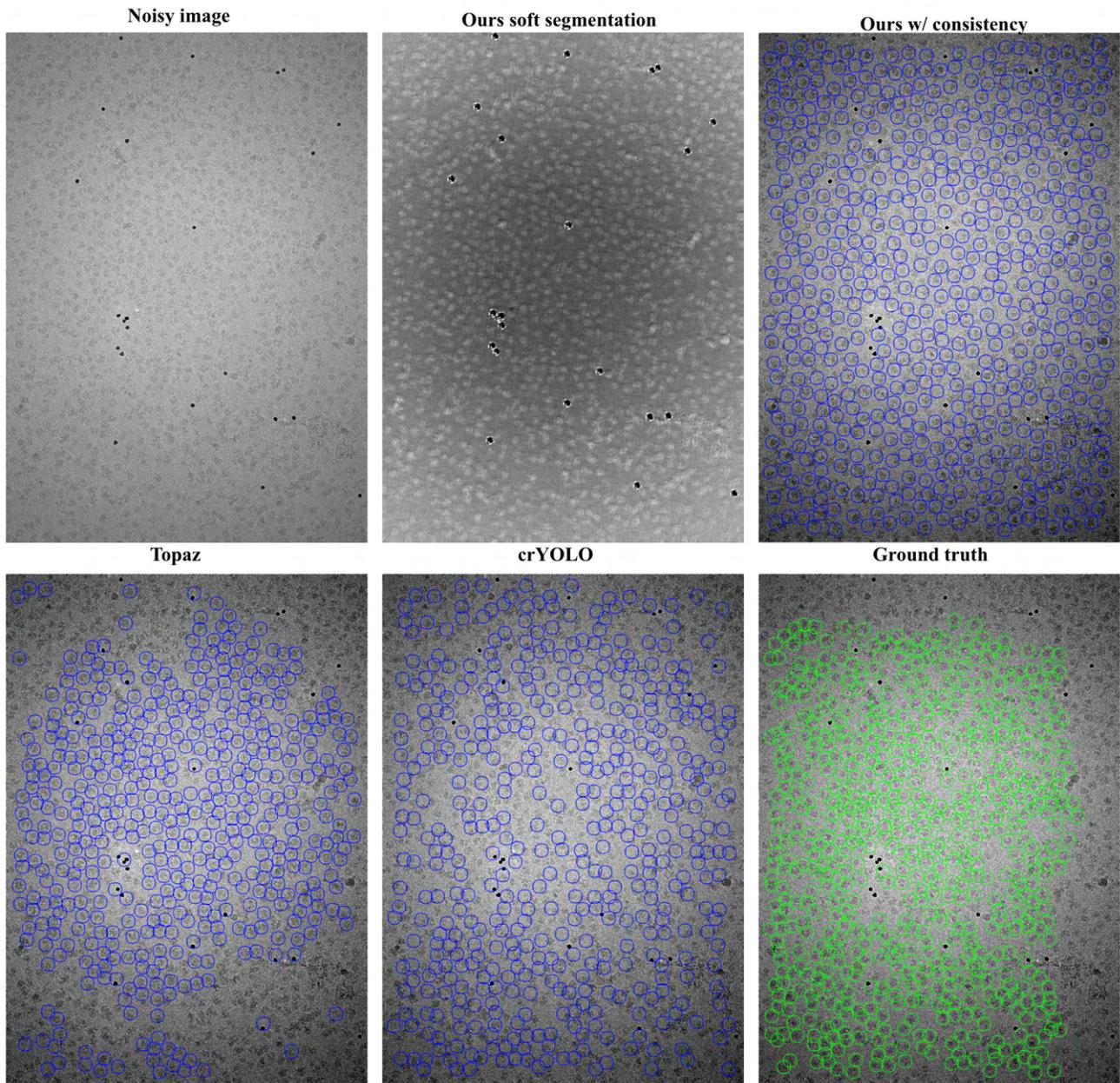
Noisy image

Ours soft segmentation

Ours w/ consistency

Topaz

crYOLO

Ground truth

Figure 7. **Additional qualitative results for particle detection of EMPIAR 10304.** Here we are showing the full image with detection result. Note our method is able to avoid contamination areas and gold beads. Note: some of the particles near edges of the image are not getting labeled in manual picking. Our method is able to detect these particles near boundaries.
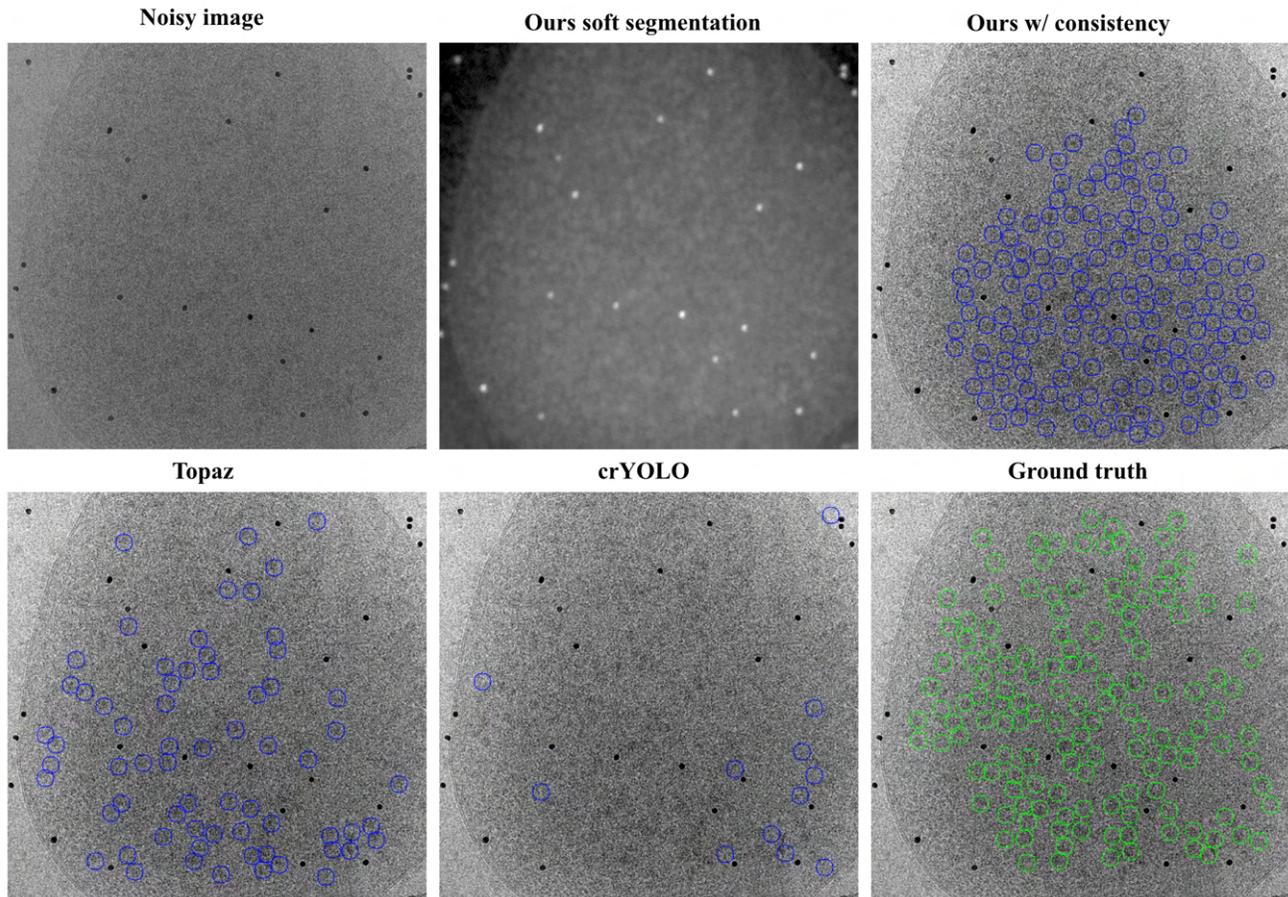
Figure 8. **Additional qualitative results for particle detection of EMPIAR 10499.** Here we are show the soft segmentation map and particle detection outputs using different methods. Our method is able to identify more particles compared to the other two methods.
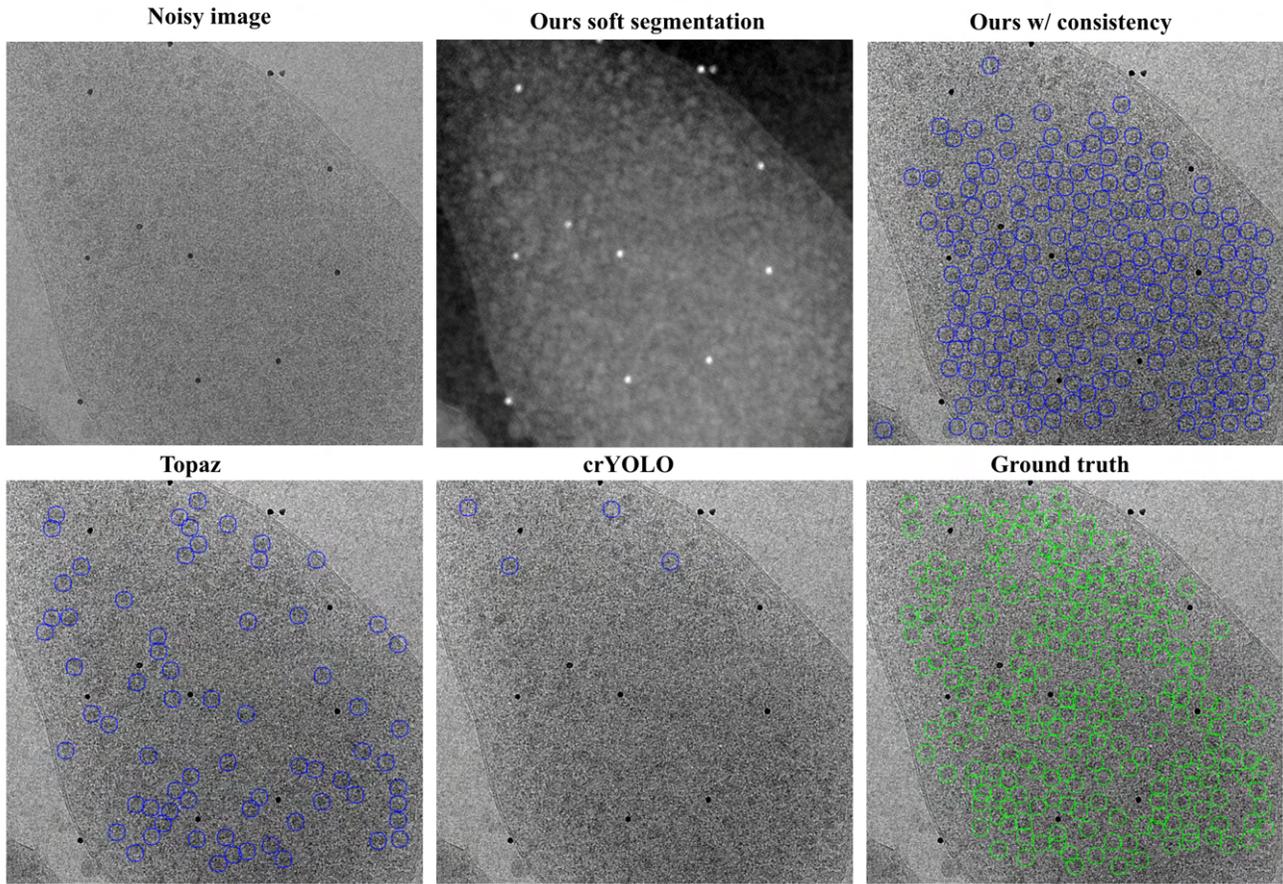
Figure 9. **Additional qualitative results for particle detection of EMPIAR 10499.** Here we are show the soft segmentation map and particle detection outputs using different methods. Our method is able to identify more particles compared to the other two methods.