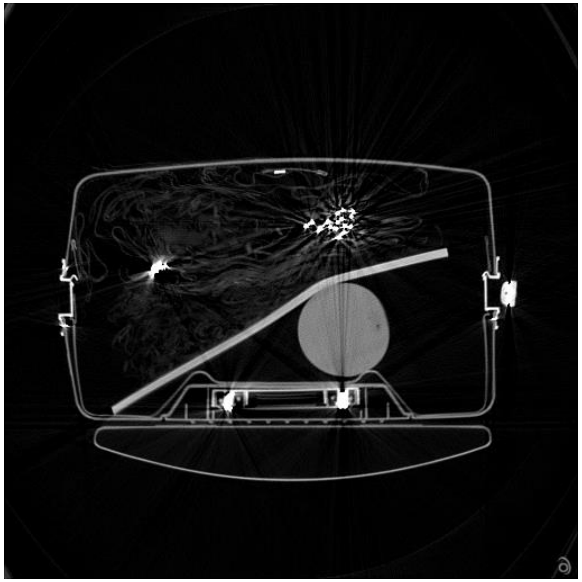
T03 Results Assessment

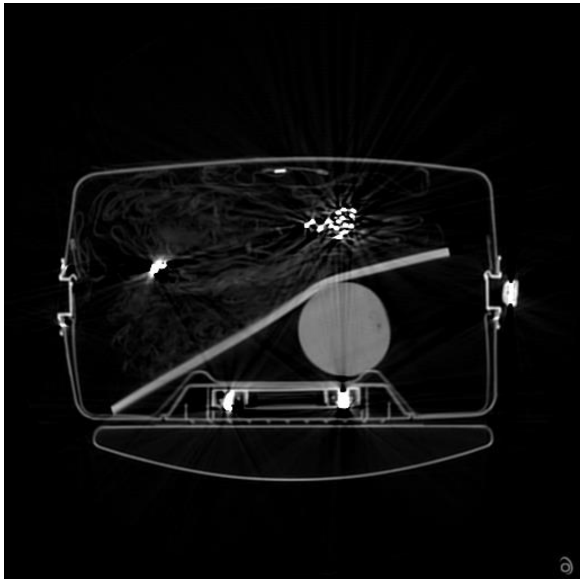
David F. Wiley
Deb Ghosh
Stratovan Corporation

Our Task

Medium_Clutter2 - Slice.175



Xrec - Baseline



Researcher A

Researcher B

Which is *better*? … and why?

(Goal is NOT to rank researchers)

Conclusions – Accuracy Results

	Water				Saline			
	σ	Edge	CCL	Tum	σ	Edge	CCL	Tum
Purdue / Notre Dame								
Harvard								
Tennessee								
UCSD								
Chicago								
Utah								
Boston								
Tufts								

Better		**Insignificant Change**		**Worse**

We are not evaluating rubber sheets due to object philosophy problem.

Conclusions – Compactness Results

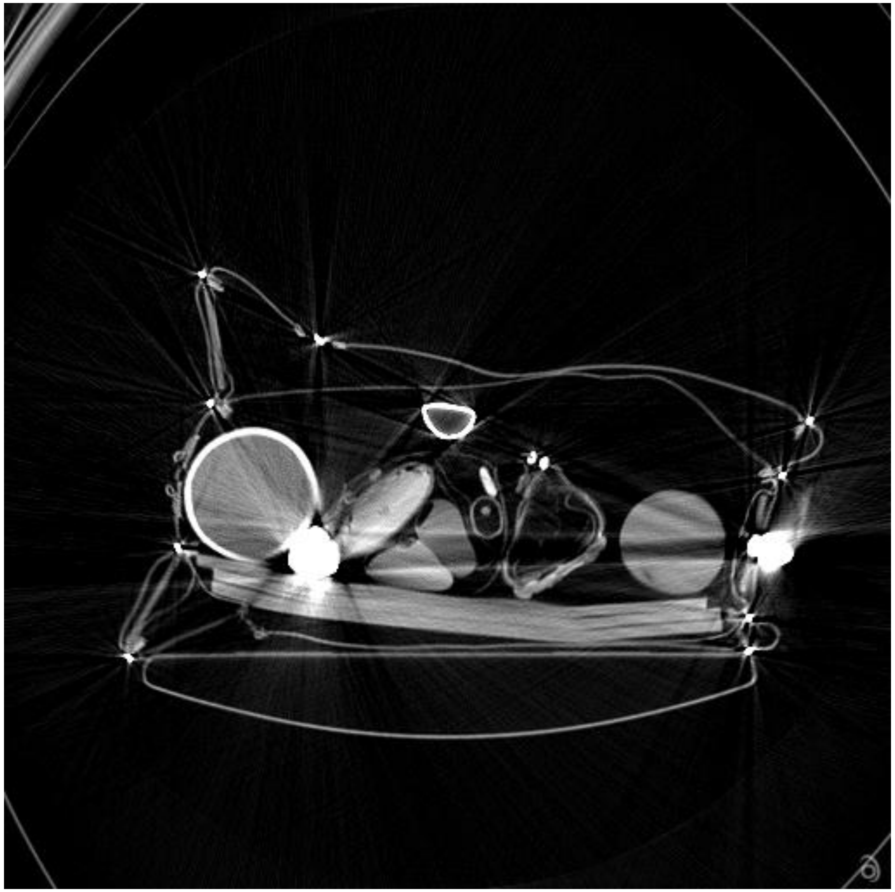
	Water				Saline			
	σ	Edge	CCL	Tum	σ	Edge	CCL	Tum
Purdue / Notre Dame	[orange]	[orange]	[green]	[green]	[blue]	[green]	[blue]	[blue]
Harvard	[yellow]	[yellow]	[green]	[green]	[blue]	[green]	[blue]	[blue]
Tennessee	[orange]	[green]	[green]	[green]	[blue]	[green]	[green]	[green]
UCSD	[yellow]	[yellow]	[yellow]	[yellow]	[blue]	[blue]	[blue]	[blue]
Chicago	[red]	[orange]	[orange]	[orange]	[green]	[yellow]	[orange]	[orange]
Utah	[orange]	[green]	[yellow]	[yellow]				
Boston	[green]	[green]	[green]	[green]	[green]	[green]	[green]	[green]
Tufts	[green]	[green]	[green]	[green]	[green]	[green]	[green]	[green]

Better [blue]	[green]	Insignificant Change [yellow]	[orange]	Worse [red]

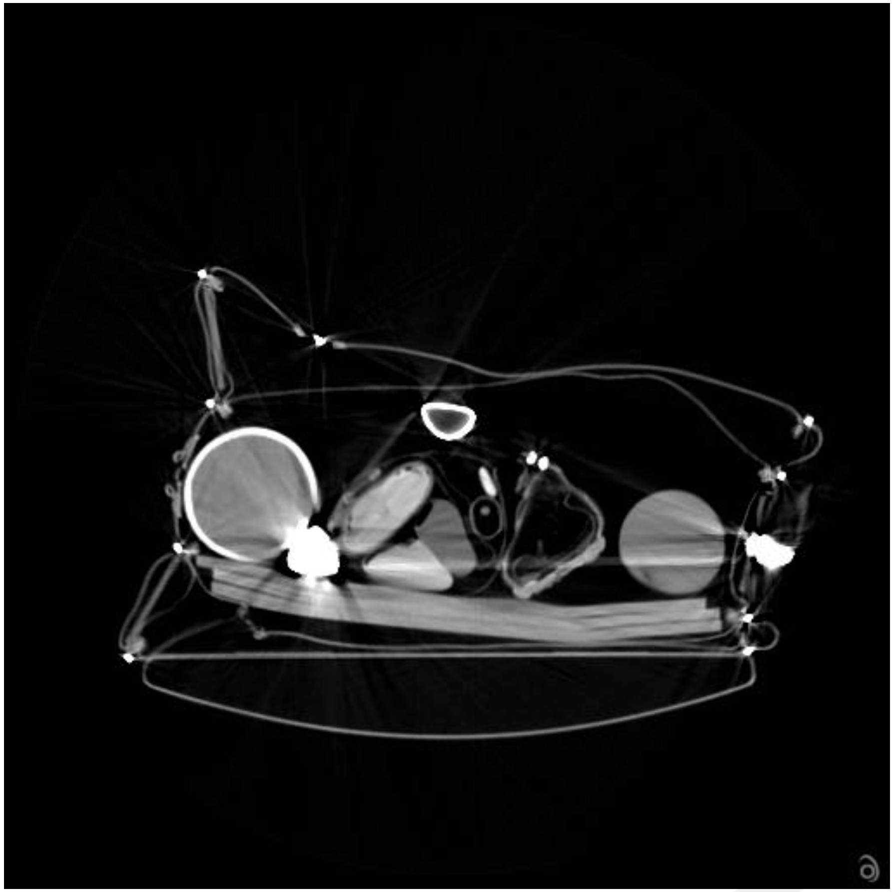
We are not evaluating rubber sheets due to object philosophy problem.

Purdue/Notre Dame – Doped Water (Better)

XRec



Purdue/Notre Dame

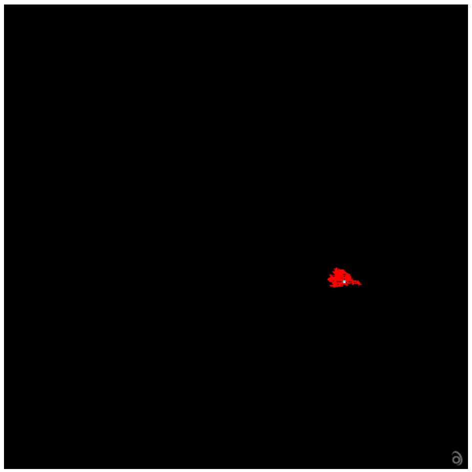


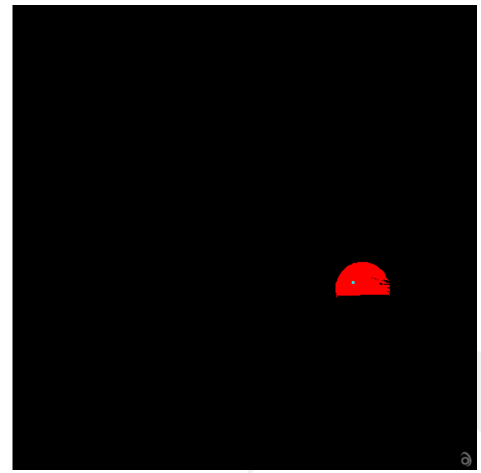
High_Clutter1 Slice.239

Purdue/Notre Dame – Doped Water (Better)

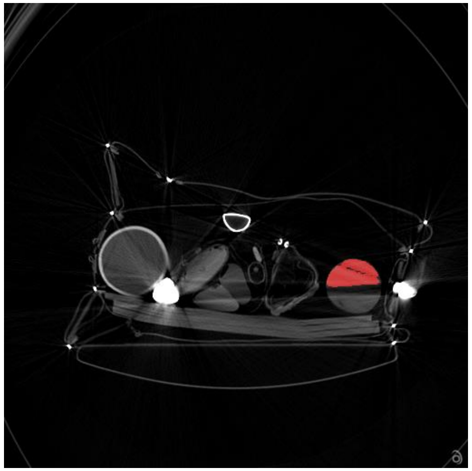
XRec | Purdue/Notre Dame

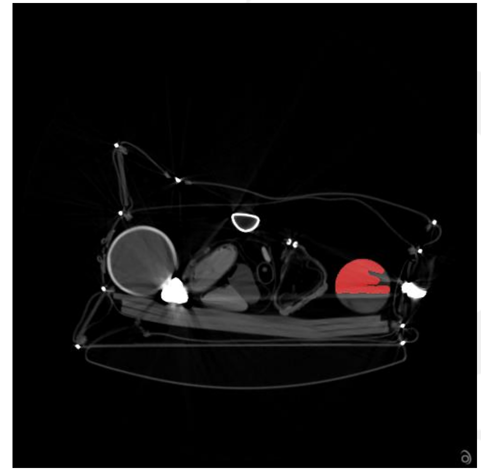
CCL





Tumbler

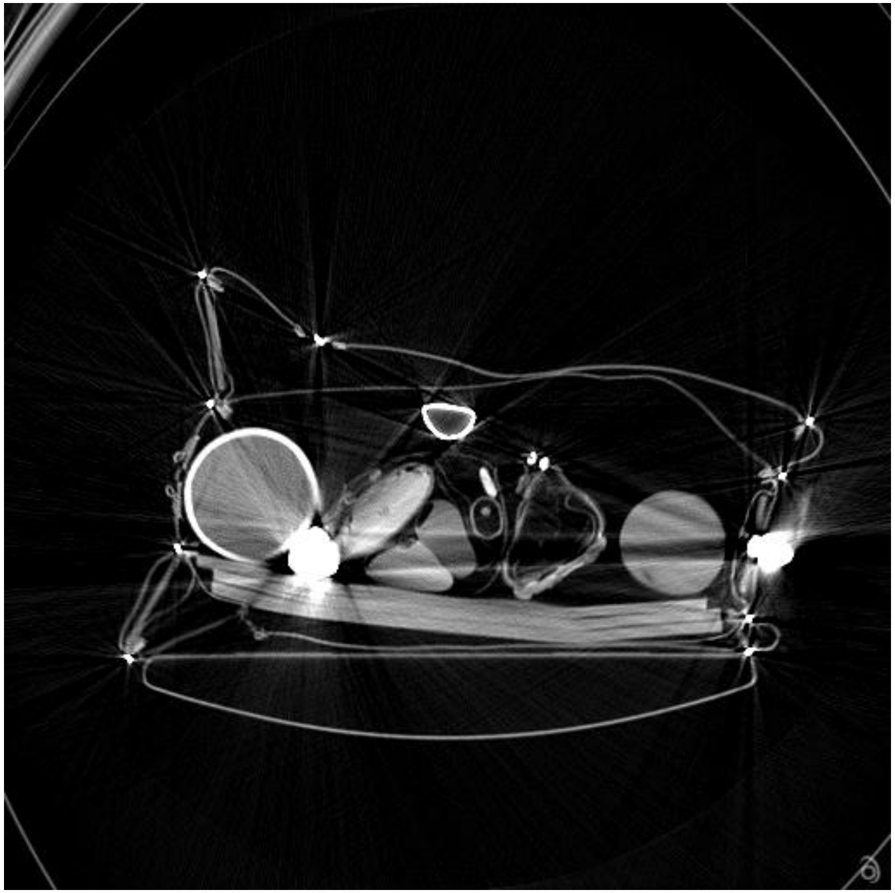




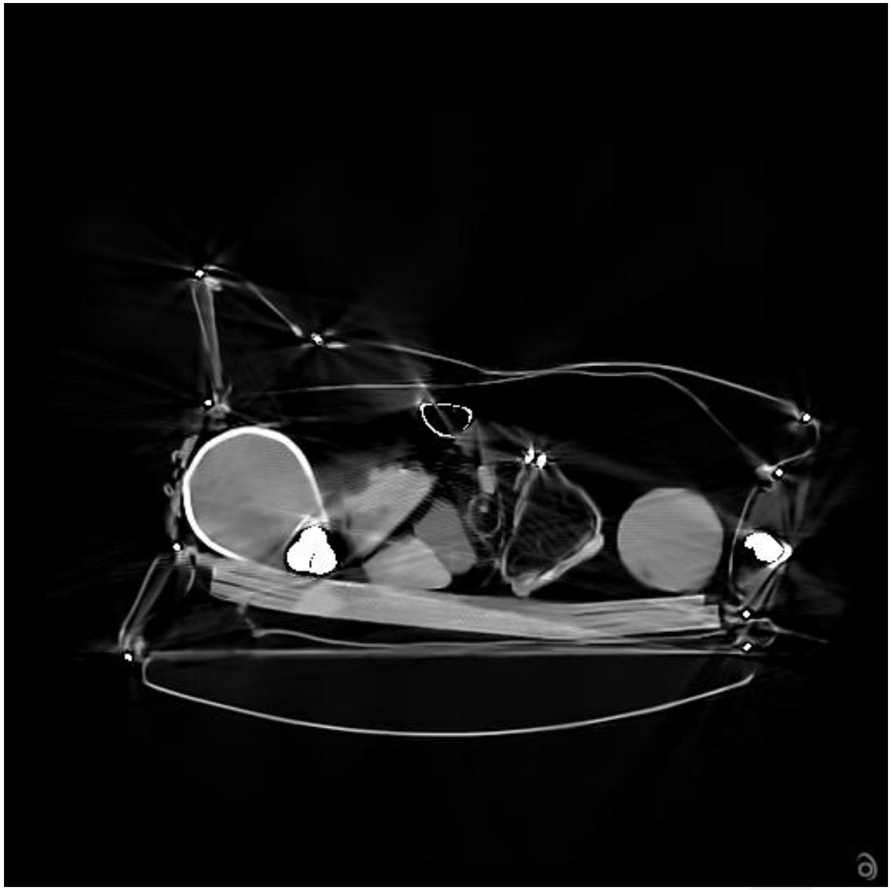
High_Clutter1 Slice.239

Harvard – Doped Water (Better)

XRec



Harvard



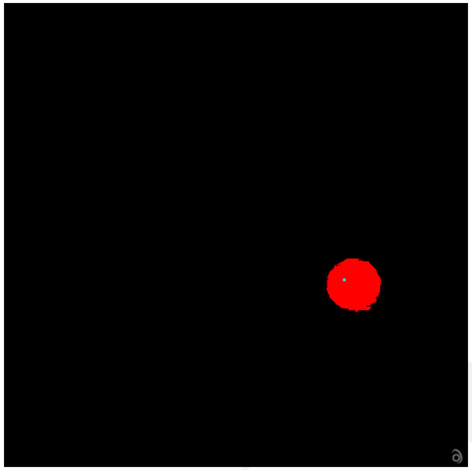
High_Clutter1 Slice.239

Harvard – Doped Water (Better)

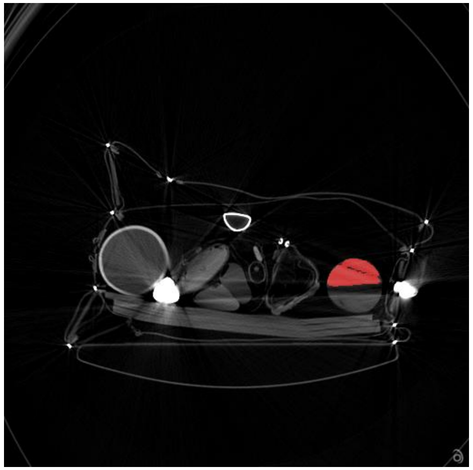
XRec

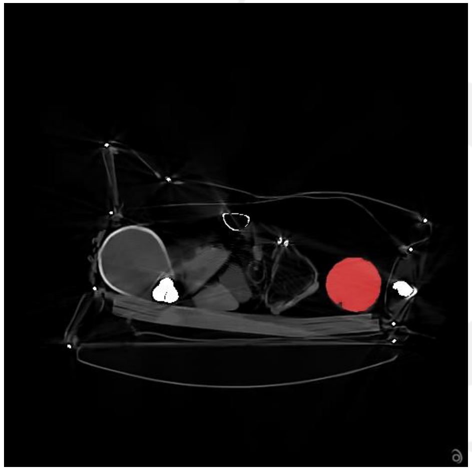
Harvard

CCL



Tumbler

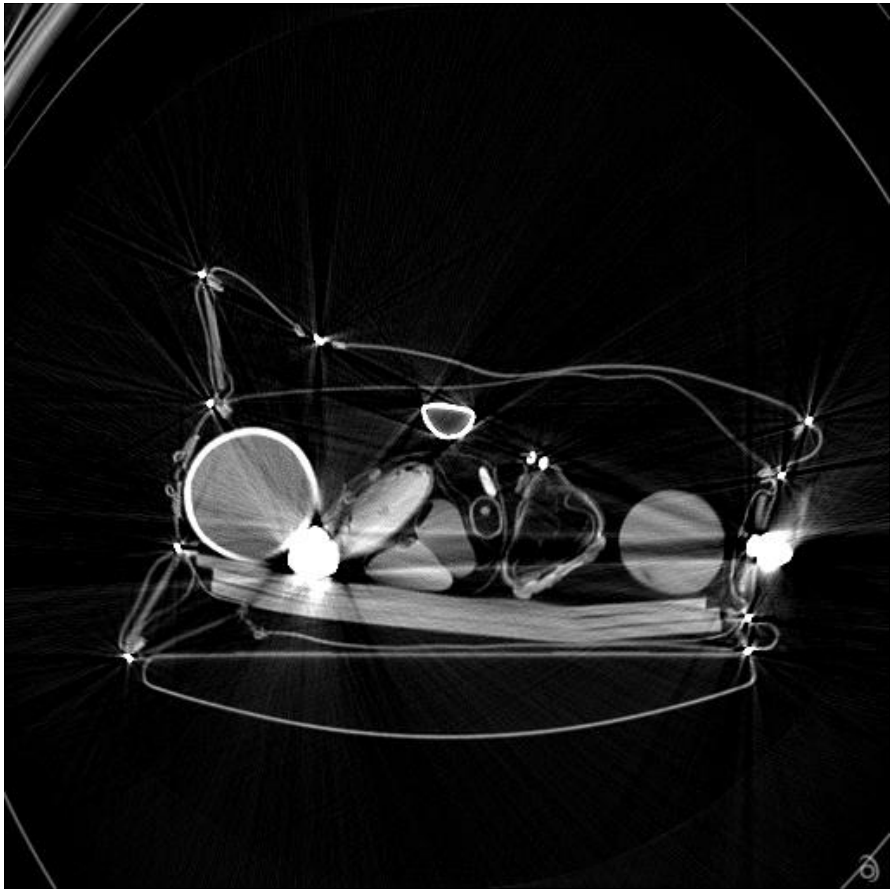




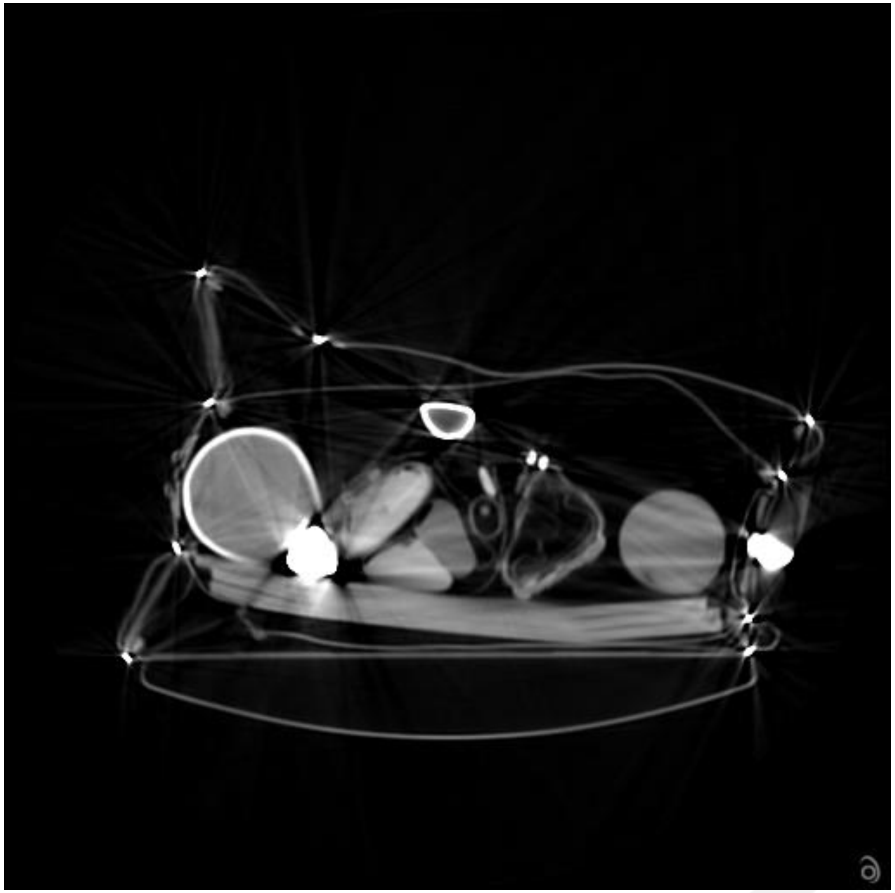
High_Clutter1 Slice.239

Gregor – Doped Water (Better)

XRec



Gregor



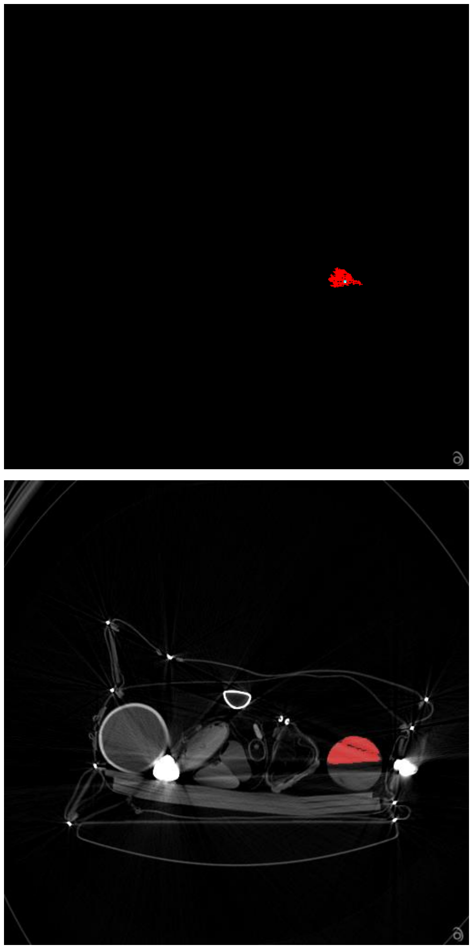
High_Clutter1 Slice.239

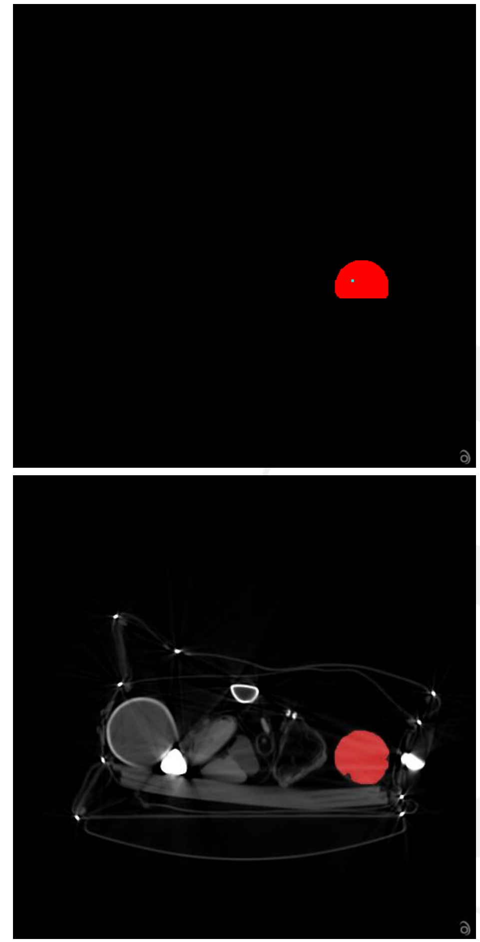
Gregor – Doped Water (Better)

XRec | Gregor

CCL

Tumbler

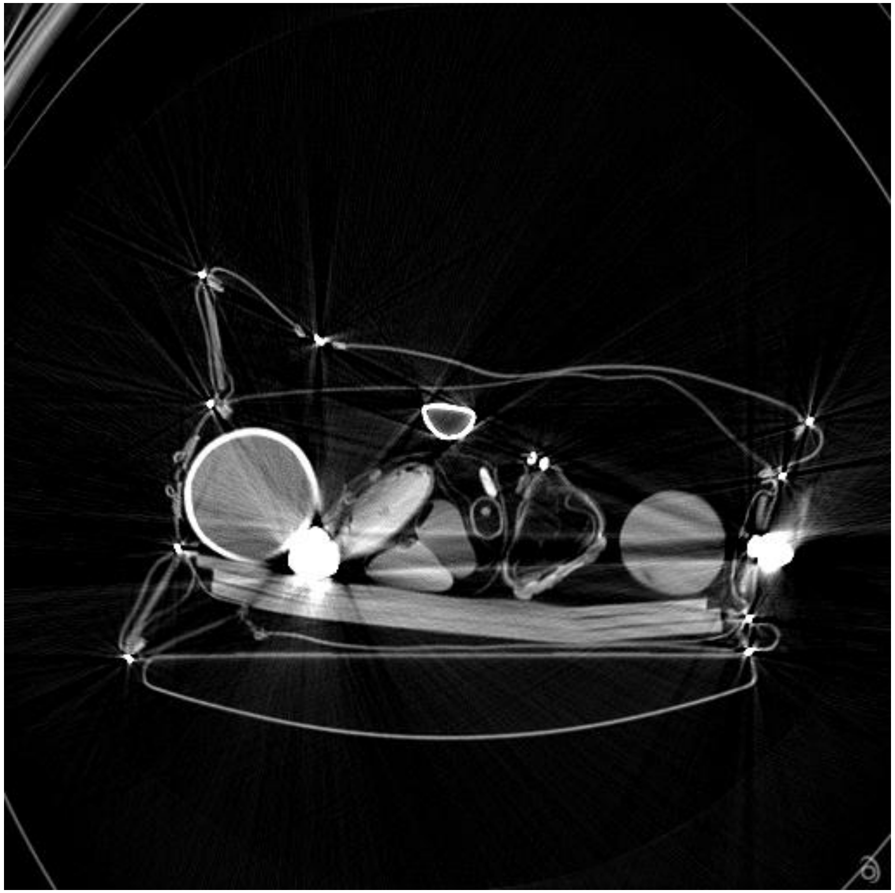




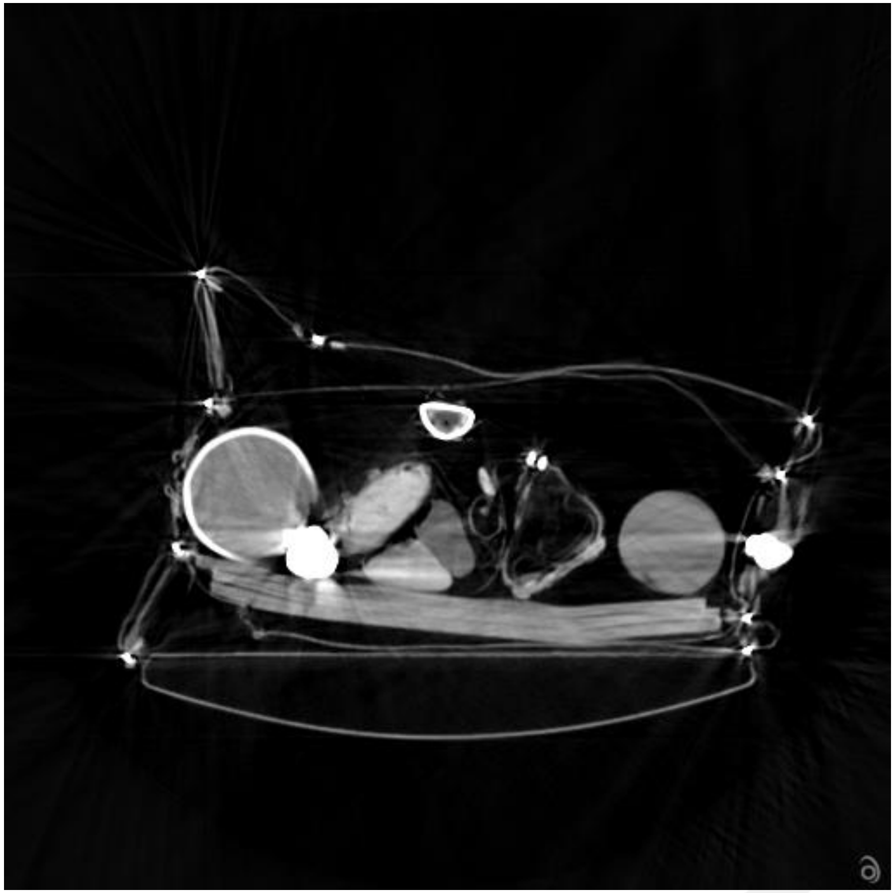
High_Clutter1 Slice.239

UCSD – Doped Water (Better)

XRec



UCSD



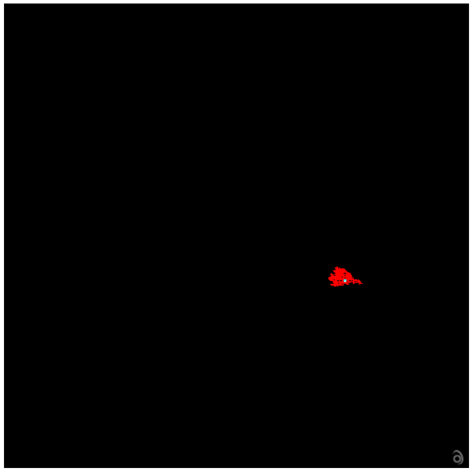
High_Clutter1 Slice.239

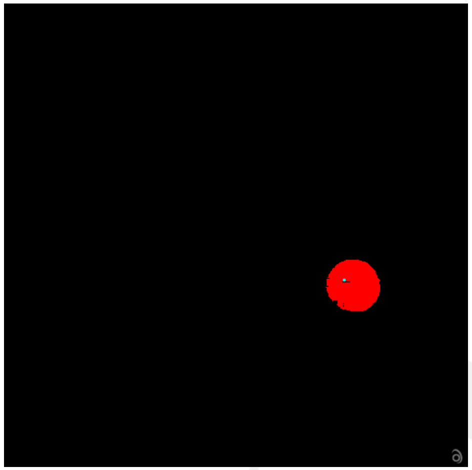
UCSD – Doped Water (Better)

XRec

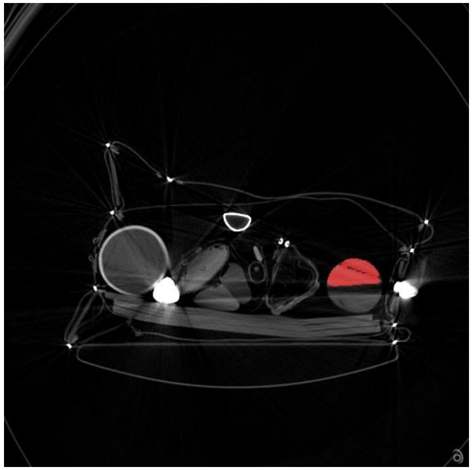
UCSD

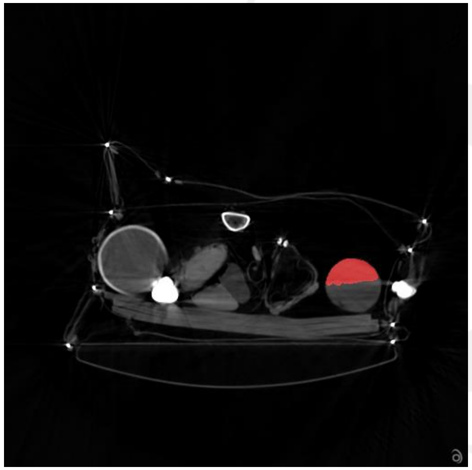
CCL





Tumbler

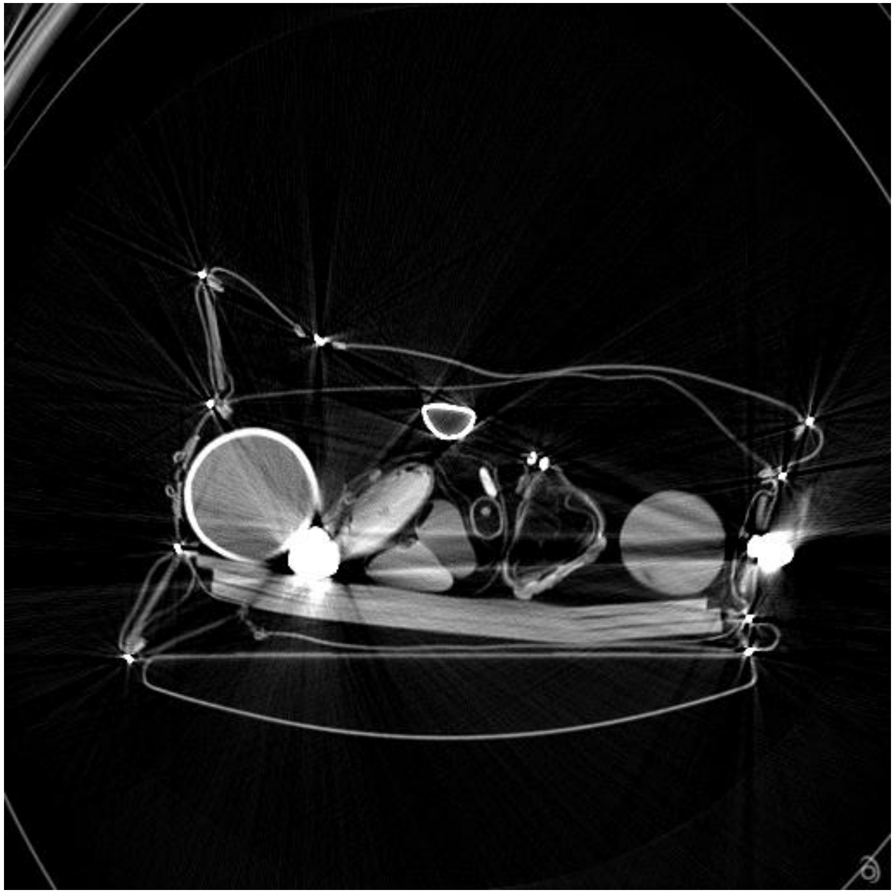




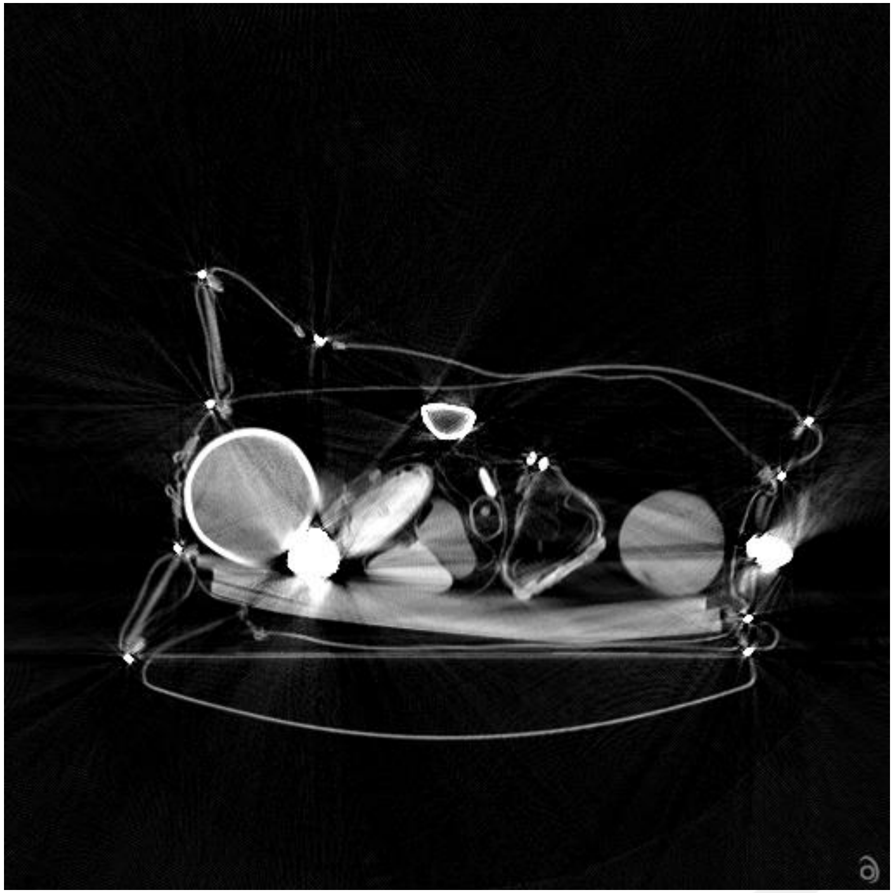
High_Clutter1 Slice.239

Chicago – Doped Water (Better)

XRec



Chicago

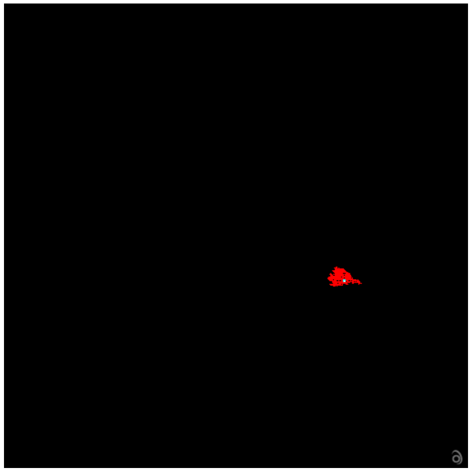


High_Clutter1 Slice.239

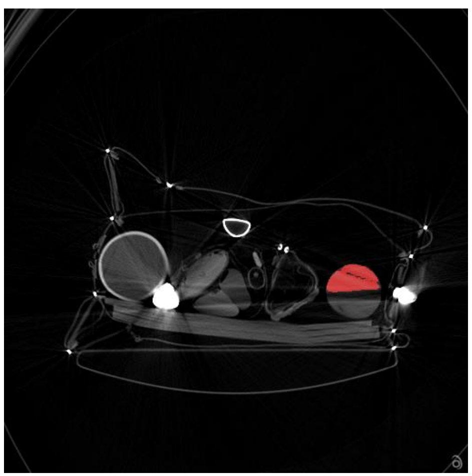
Chicago – Doped Water (Better)

XRec | Chicago

CCL



Tumbler



High_Clutter1 Slice.239

Utah – Water (Better)

XRec

Utah

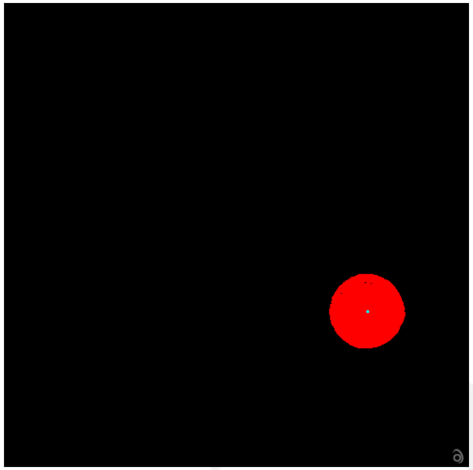
Medium_Clutter1 Slice.231

Utah – Water (Better)

XRec

Utah

CCL



Tumbler

Medium_Clutter1 Slice.231

Impact Relationships

- Improved stddev accuracy seems to reduce edge contrast accuracy
  - algorithms should be sensitive to object edges as well and try to increase contrast

- Reduced edge contrast accuracy did not outweigh gain obtained from improved stddev accuracy
  - segmentations were better or remained the same

Impact Relationships

- Improved stddev accuracy impacted water/saline compactness differently. Why?
  - Water: mean spread out, wider variation
  - Saline: mean compacted, less variation
- Improved stddev compactness correlates to more compact/consistent edge contrast
- Improved edge contrast compactness correlates to improved segmentation consistency

Recommendations for the Future

- Concentrate on reducing stddev (within homogenous objects) while increasing edge contrast
  - This improves segmentation and ultimately feature quality.
- A single bad pixel on an object boundary can cause a segmentation to leak
  - Try to improve the entire object boundary
- Reduced stddev may increase mean spread which can increase cloud size in ATR
  - Look at outliers to find out what's happening.
- Stacked sheets are an object philosophy problem NOT a reconstruction problem

The End

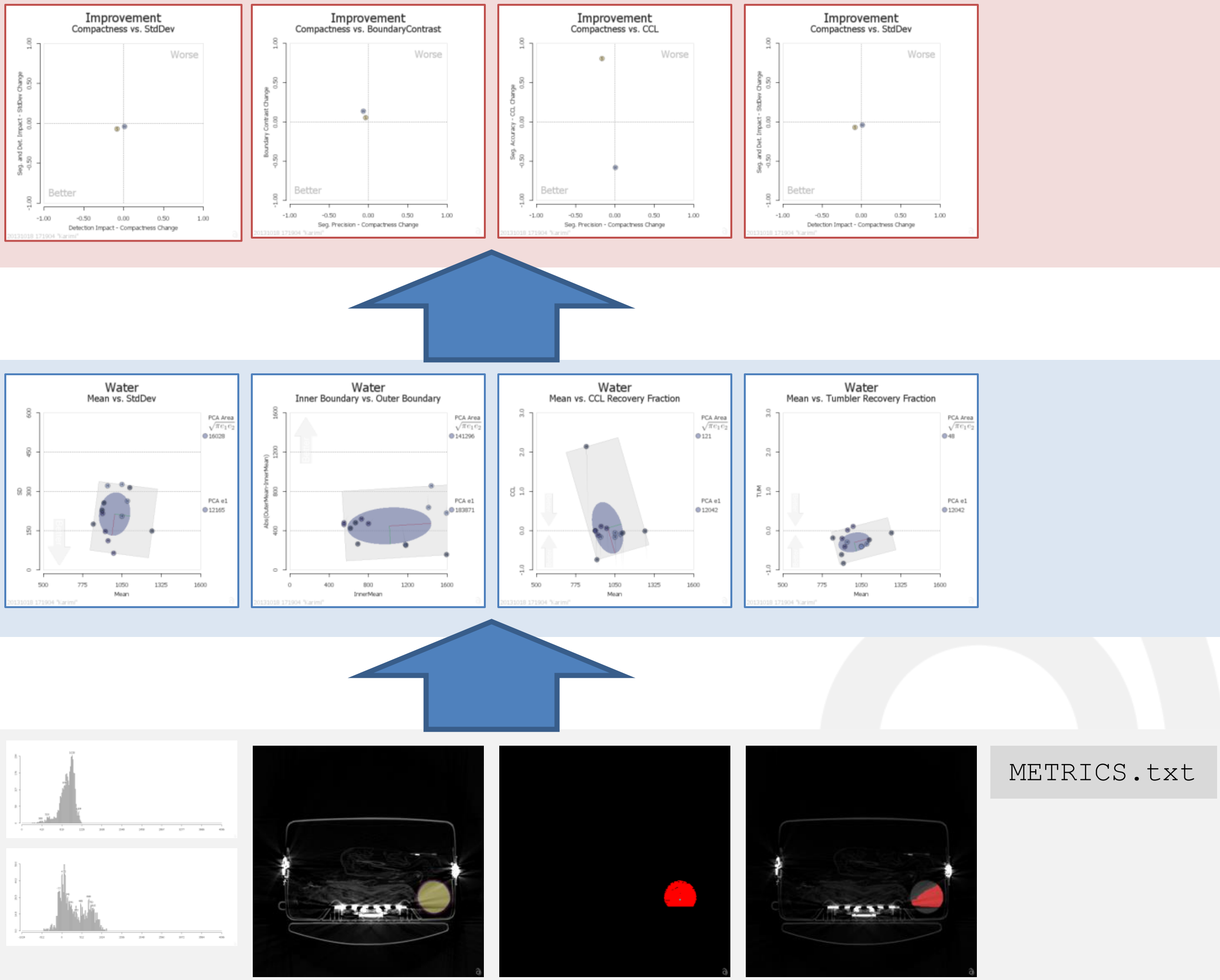
(but there's more slides if you have questions)

Analysis Process - 40GB of Data

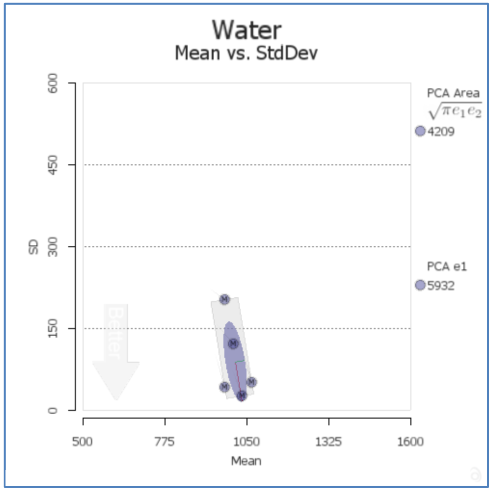
Improvement plots

Cloud object plots

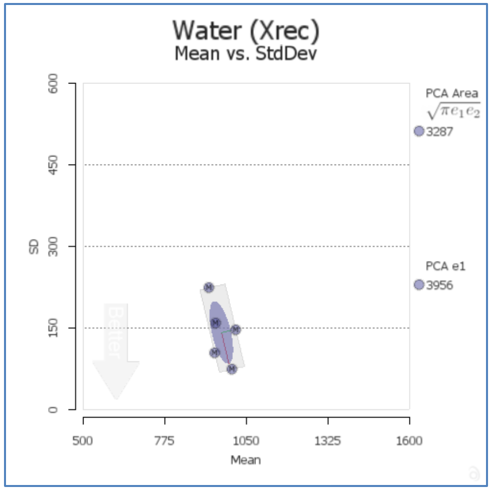
Recon image data



Improvement Over Xrec



−



=

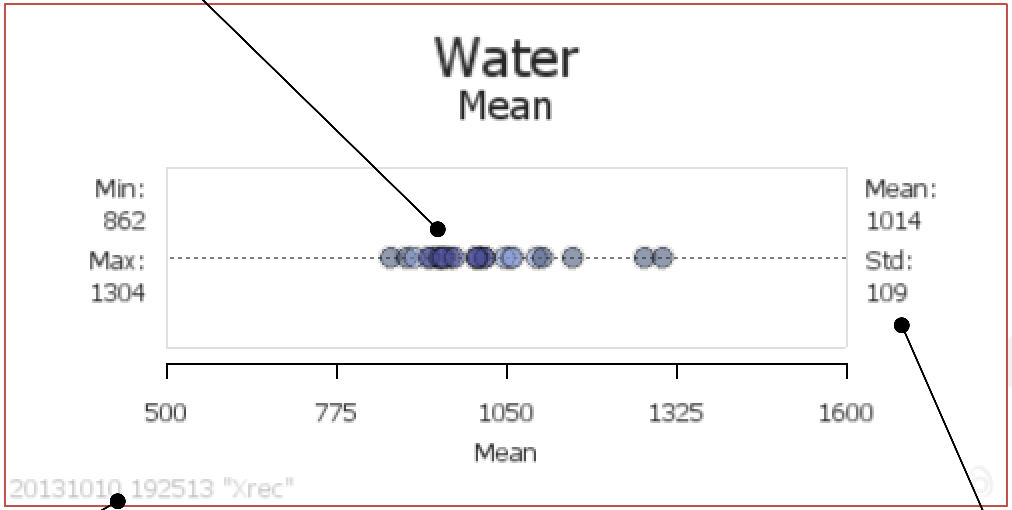
Compactness: $$I_x = \frac{e_1^{res} - e_1^{xrec}}{e_1^{xrec}}$$

Accuracy: $$I_y = \frac{pca_y^{res} - pca_y^{xrec}}{pca_y^{xrec}}$$

We should have used $\sqrt{e_1^2 + e_2^2}$ instead of e_1. We may do this for the final report.

Cloud Results - Mean

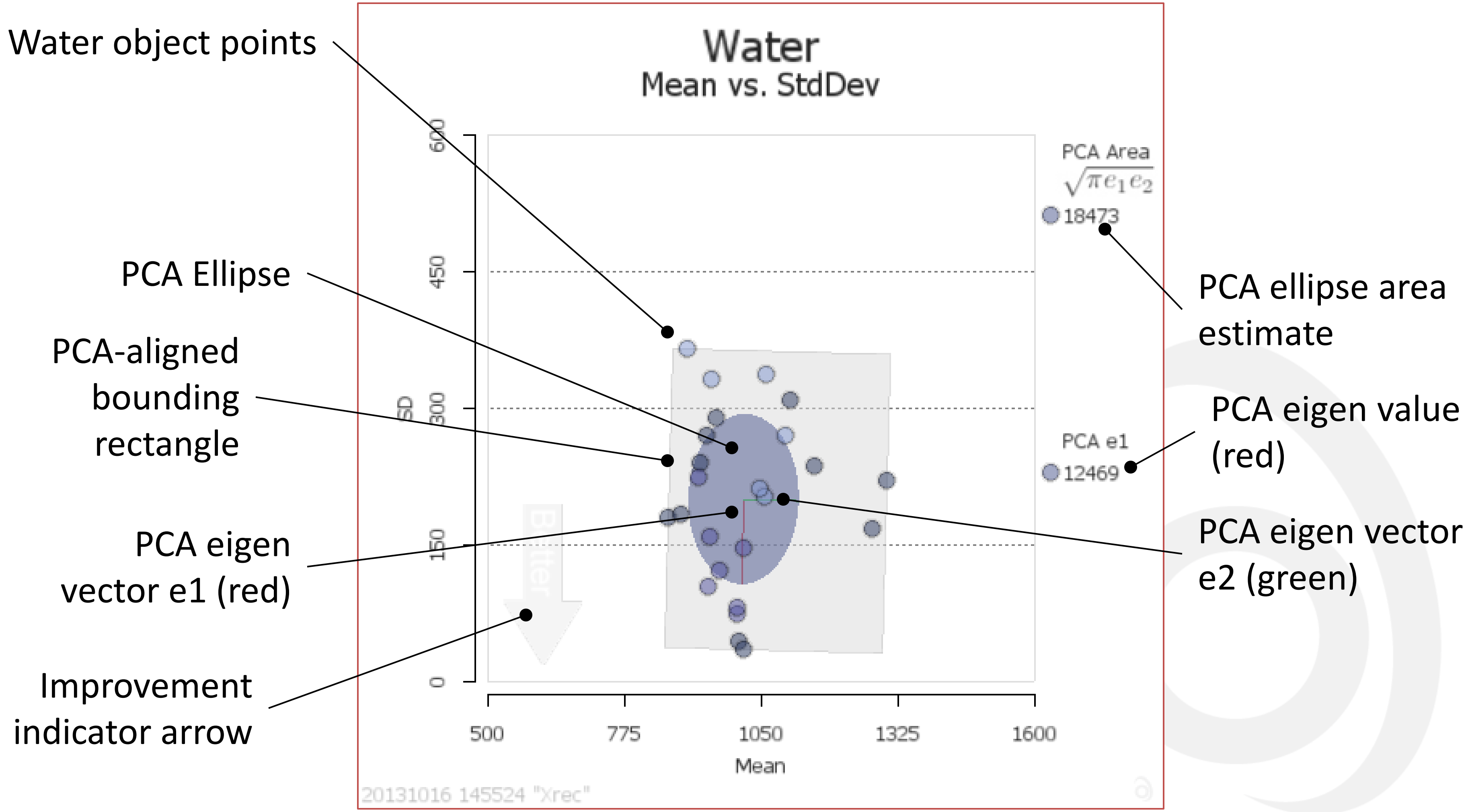
Water object points



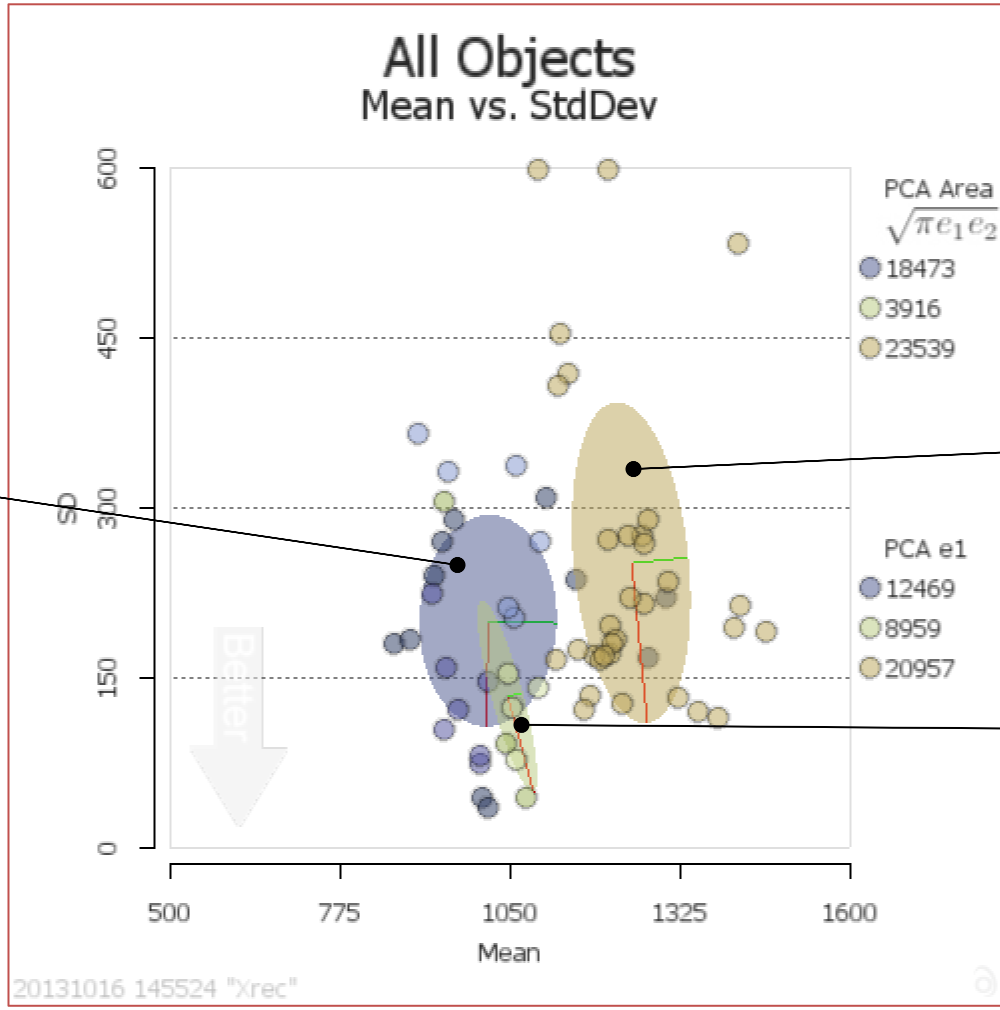
Date: 2013-10-10
24h Time: 19h 25m 13s
Group name: Xrec

In general, improvement is indicated by a lower standard deviation of object mean values.

Cloud Results – Mean vs. Std



Cloud Results – All Objects



Water plot

Rubber sheet plot

Doped water plot

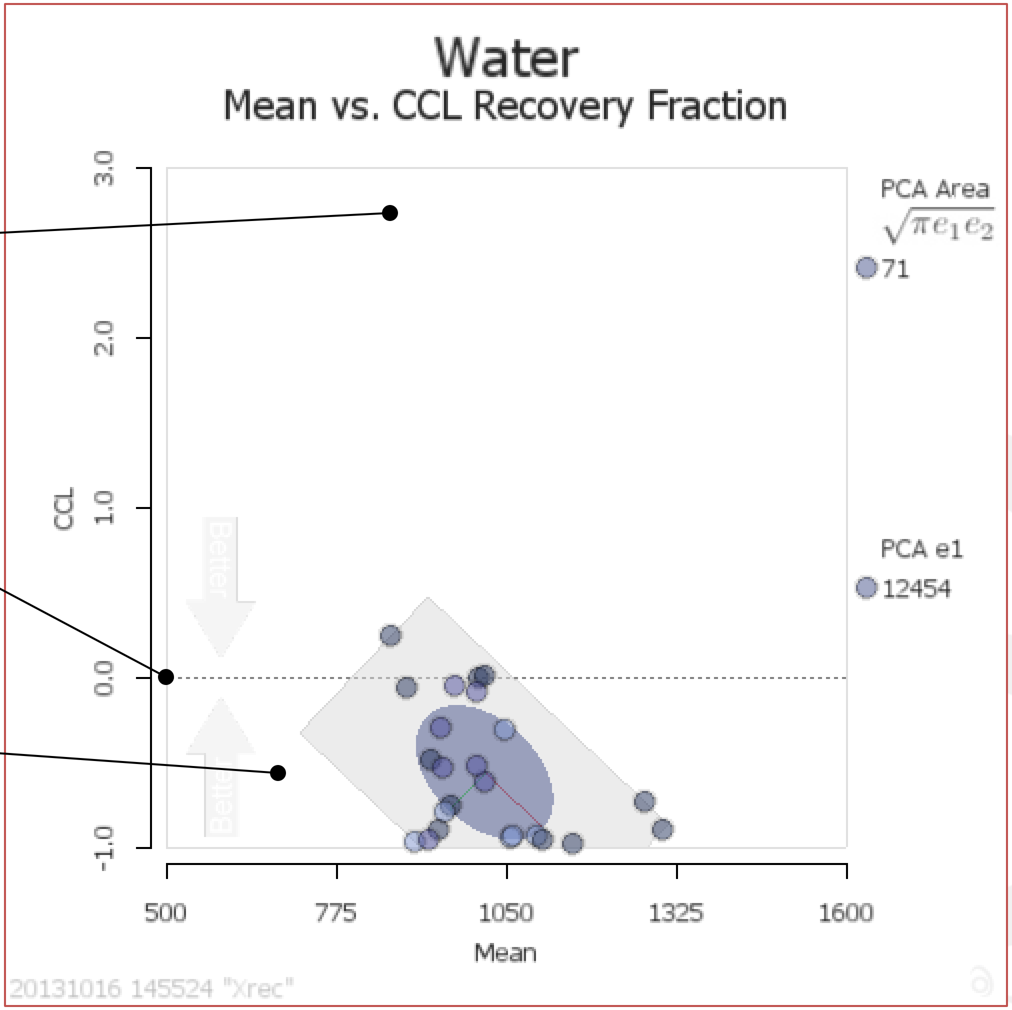
ATR needs to distinguish between these different groups. Significant overlap in this graph implies follow-on challenges in ATR. Good separation implies ATR will be more effective.

Cloud Results – Mean vs. CCL

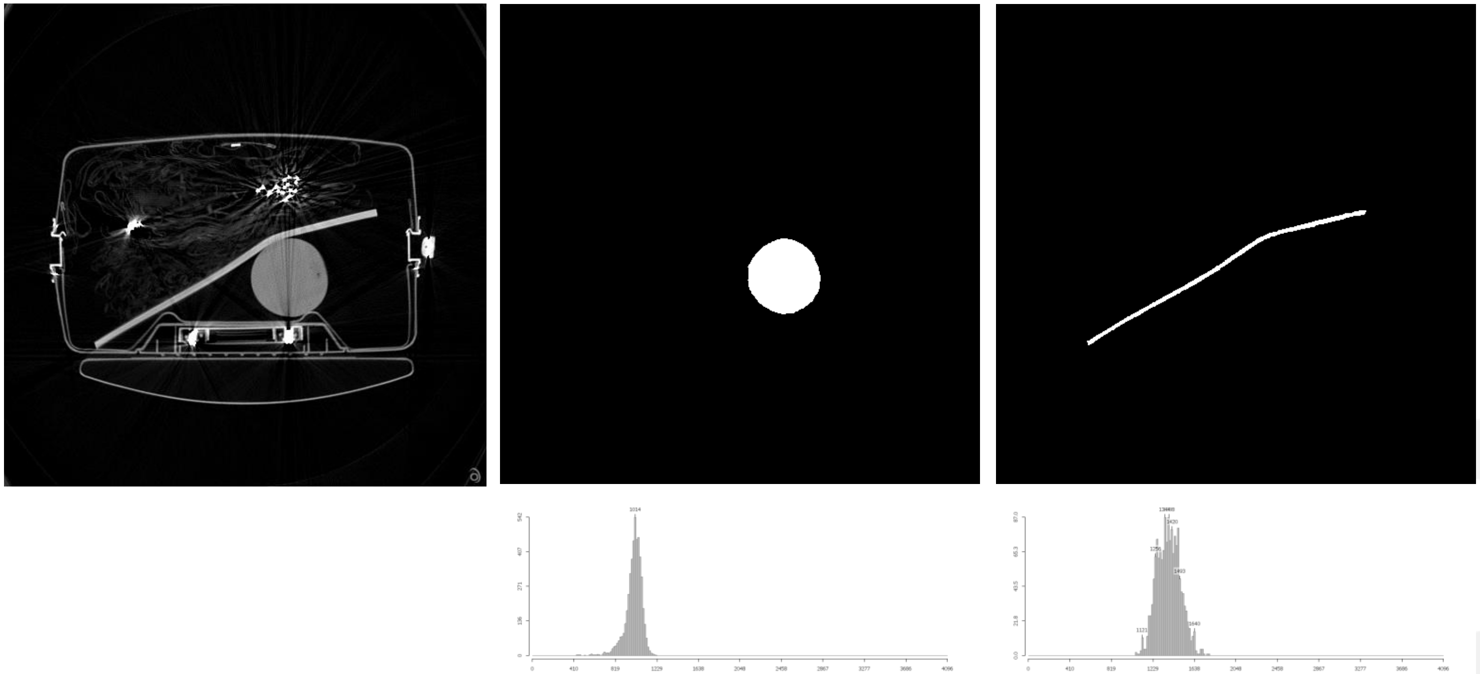
Segmentation was larger than ground truth

Ideal segmentation

Segmentation was smaller than ground truth



What did we measure? – Objects!



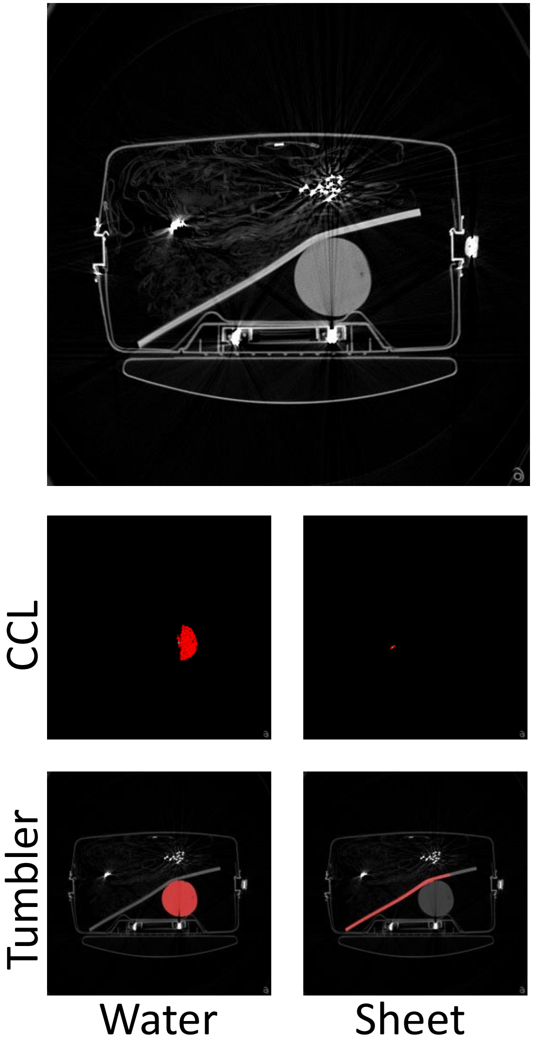
Goal is to accurately segment first *then* compute object characteristics.

Assumption: Homogenous objects should result in a single peak (i.e., stddev = 0)

Implication: Wider peaks make segmentation harder and increase cluster size in detection parameter space

Process: Use the same segmentation mask for all researchers

What did we measure? – Segmentations!



Recovery fraction:

$$R_A = \frac{Pix(A) - Pix(X)}{Pix(X)}$$

Where:

- A is either the CCL of Tumbler segmentation results.
- X is the ground truth segmentation.
- Pix() is simply the number of pixels in the segmentation mask.

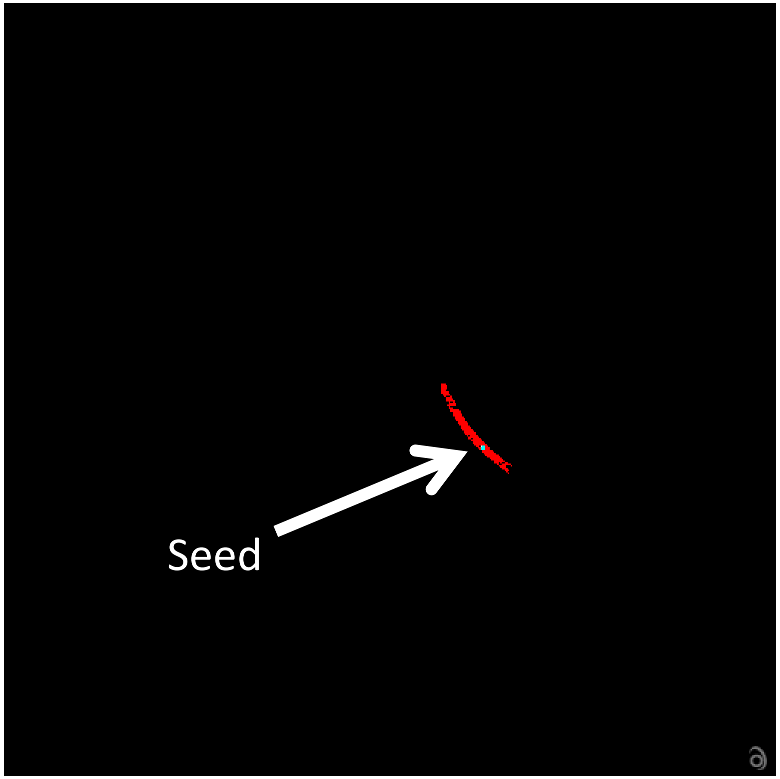
An R-value of **zero is ideal**.

A **negative** value indicates a segmentation **smaller** than the ground truth.

A **positive** value indicates a segmentation **larger** than the ground truth.

CCL Segmentation

Medium_Clutter4.242.fits.SEG_CCL_0_0013.tif



Medium_Clutter4.242.fits.SEG_CCL_ALL.tif

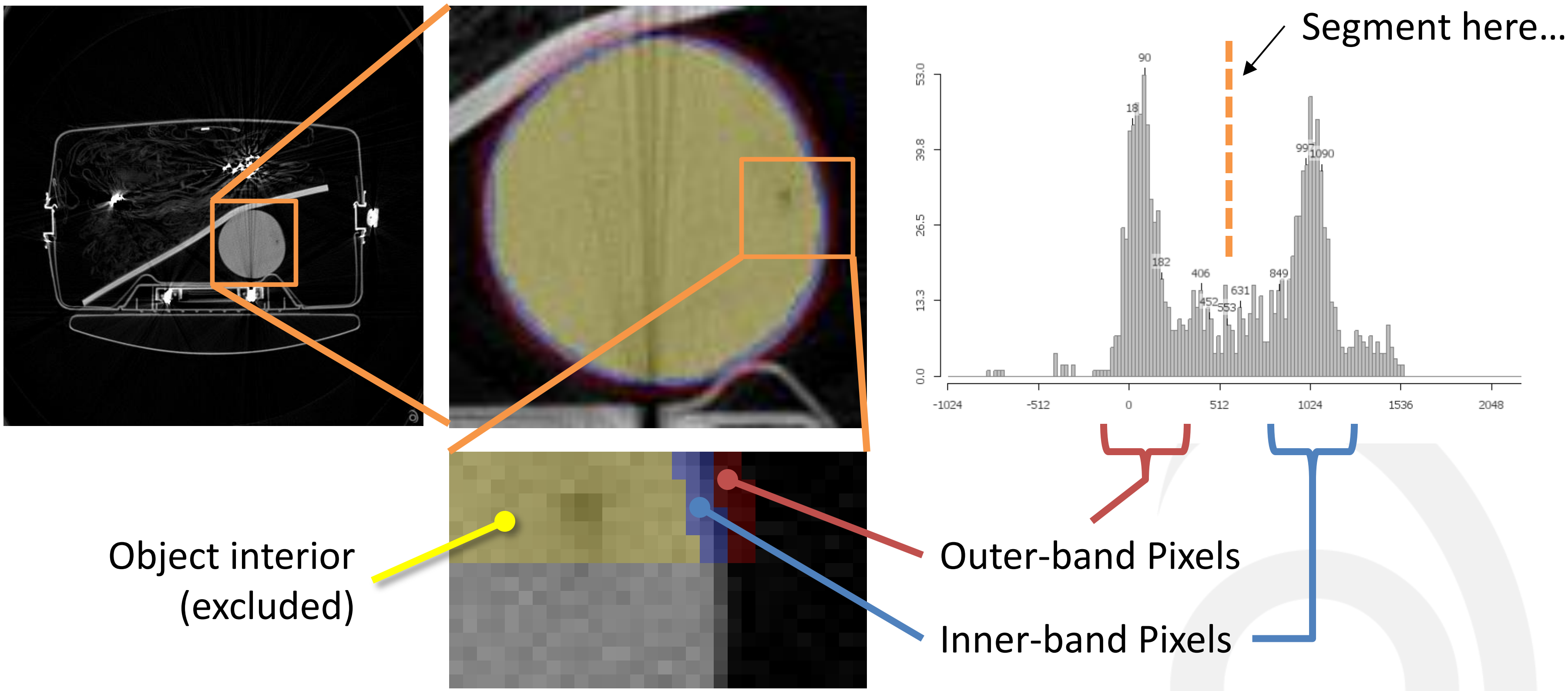
Left-image shows the CCL segmentation (red pixels) from the seed (white dot). In this case, the segmentation only obtains a small fragment of the rubber sheet due to artifacts.
Right-images shows the complete CCL segmentation, for reference only.

Tumbler Segmentation

Medium_Clutter4.242.fits.SEG_DEC_0_0013.tif

Shows Tumbler segmentation results in red pixels. Tumbler uses the same seed point that is used in CCL. In this case, the segmentation gets the lower half of the rubber sheet, but is split by an artifact from the upper portion.

What did we measure? – Boundaries!

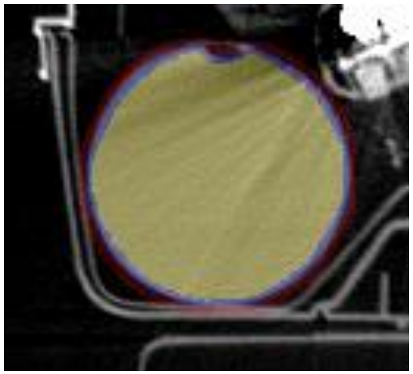


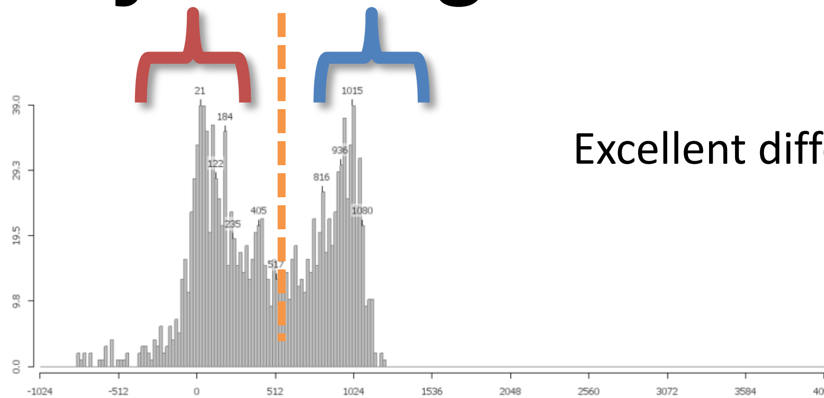
Assumption: Objects should have *crisp* boundaries to enable segmentation

Implication: Low-contrast, poorly defined boundaries, makes segmentation extremely difficult. *Abs(OuterMean-InnerMean)* relates to *"boundary contrast"*

Process: Measure two thin bands of pixels at the object boundary

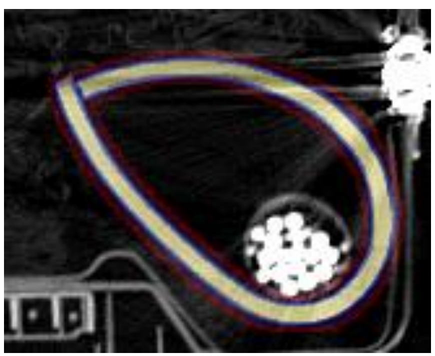
Boundary Histograms

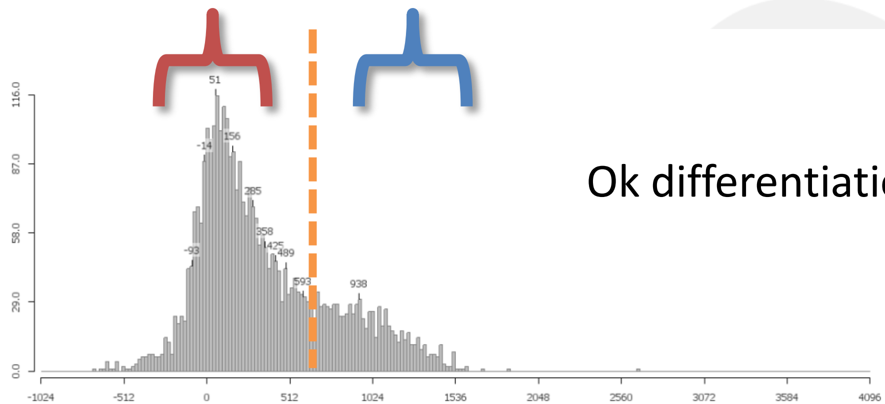




Excellent differentiation

Medium_Clutter4.134.fits.MAN_0_0002_HIST_BOUNDARY.tif



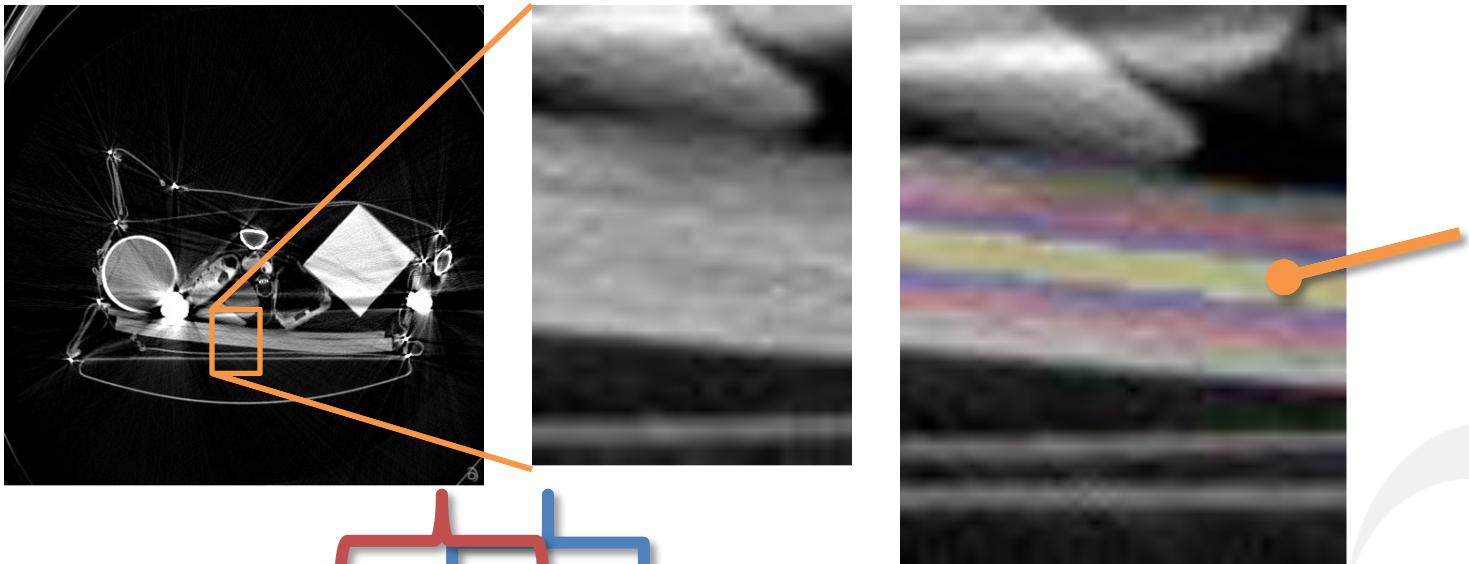


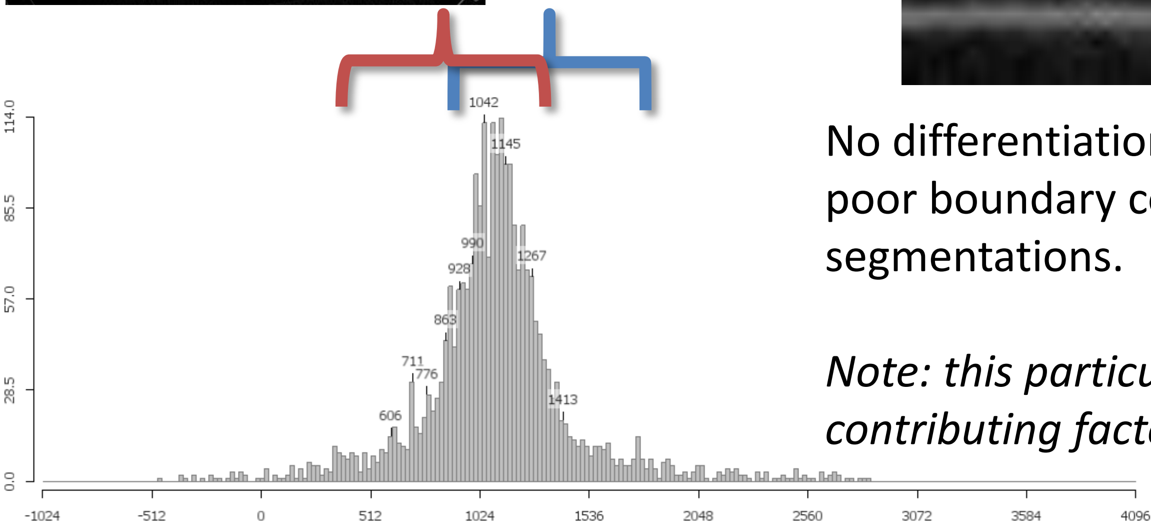
Ok differentiation

Medium_Clutter4.134.fits.MAN_1_0013_HIST_BOUNDARY.tif

Segmentation seeks to identify the boundary between red and blue regions (orange dotted line). **Differentiation between the red and blue histogram peaks directly correlates to impact on segmentation.** Good differentiation yields good segmentation. Poor differentiation yields poor segmentations.

Boundary Histogram - Poor





No differentiation between peaks. This indicates poor boundary contrast and results in poor segmentations.

Note: this particular situation is due to many contributing factors, not just reconstruction.

Metrics

Medium_Clutter4.242.fits.METRICS_MAN_0_0013.txt

```
Mean: 1464.989
  SD: 190.060
*** Metrics with Ideal Value (from Ideals.txt) ***
RMSE: 192.109
PSNR: 26.576
 SNR: 7.708
SSIM: 0.976
*** Metrics with Mean Value ***
RMSE: 190.059
PSNR: 26.669
*** Border Metrics ***
OuterMean: 168.616
InnerMean: 505.100
```

Yellow: object interior

Red: just outside of object boundary

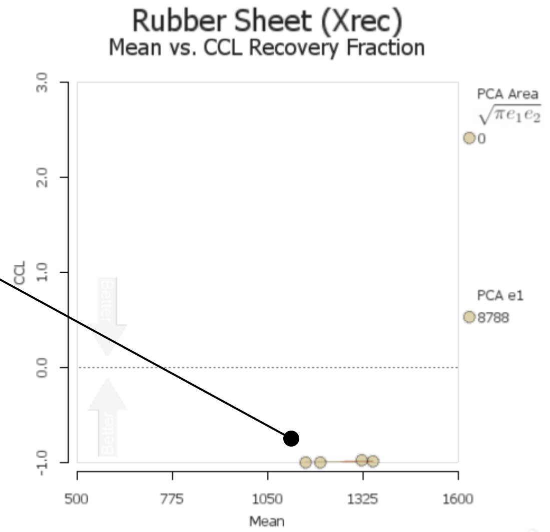
Blue: just inside of object boundary

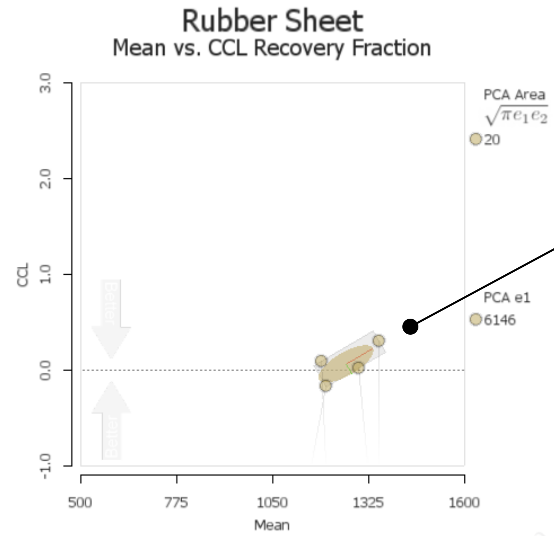
This file records the metrics output for the image slice. These metrics are computed using the cookie-cutter segmentation. **We use the top two metrics (Mean and SD) and the bottom two (OuterMean and InnerMean) in the cloud graphs.**

The difference between OuterMean and InnerMean indicates *boundary contrast*.

Cloud Comparison

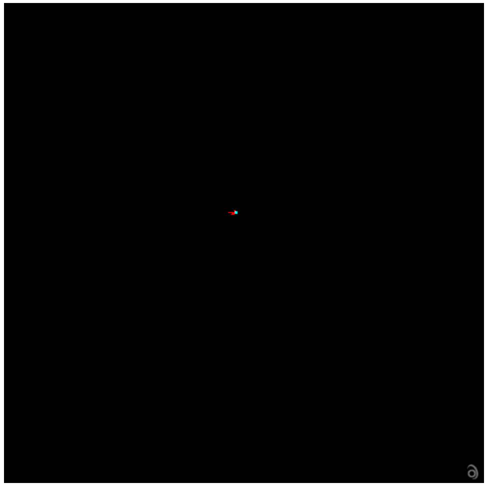
Xrec results are not very good

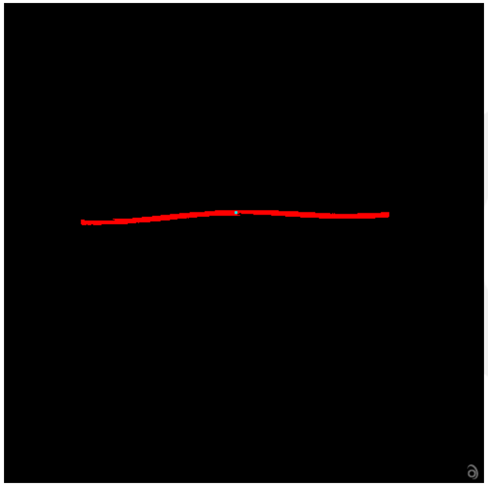




Researcher results are better!

Most of the CCL segmentations miss large portions of the sheet



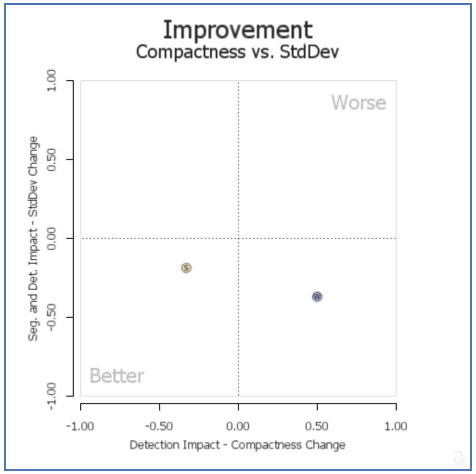


Now, CCL segmentations capture the most of the sheet

Cloud Interpretation

- **Compactness (all clouds)**: this is estimated by the PCA ellipse. Smaller is better. Better compactness improves ATR.
- **ATR Improvement (Mean vs. StdDev clouds)**: you want to see a decrease in standard deviation.
- **Segmentation Improvement (Mean vs. Recovery clouds)**: you want to see object recovery clustered around the vertical 0.
- **Segmentation Improvement (Inner/Outer clouds)**: you want to see good (red/blue) peak separation in boundary histograms.

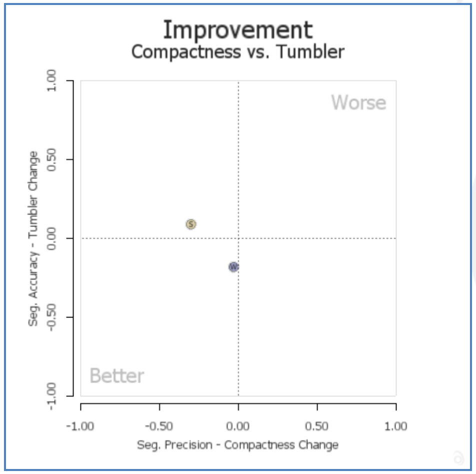
Improvement Interpretation



Improvement relative to standard deviation within an object. Correlates to ATR and segmentation quality.

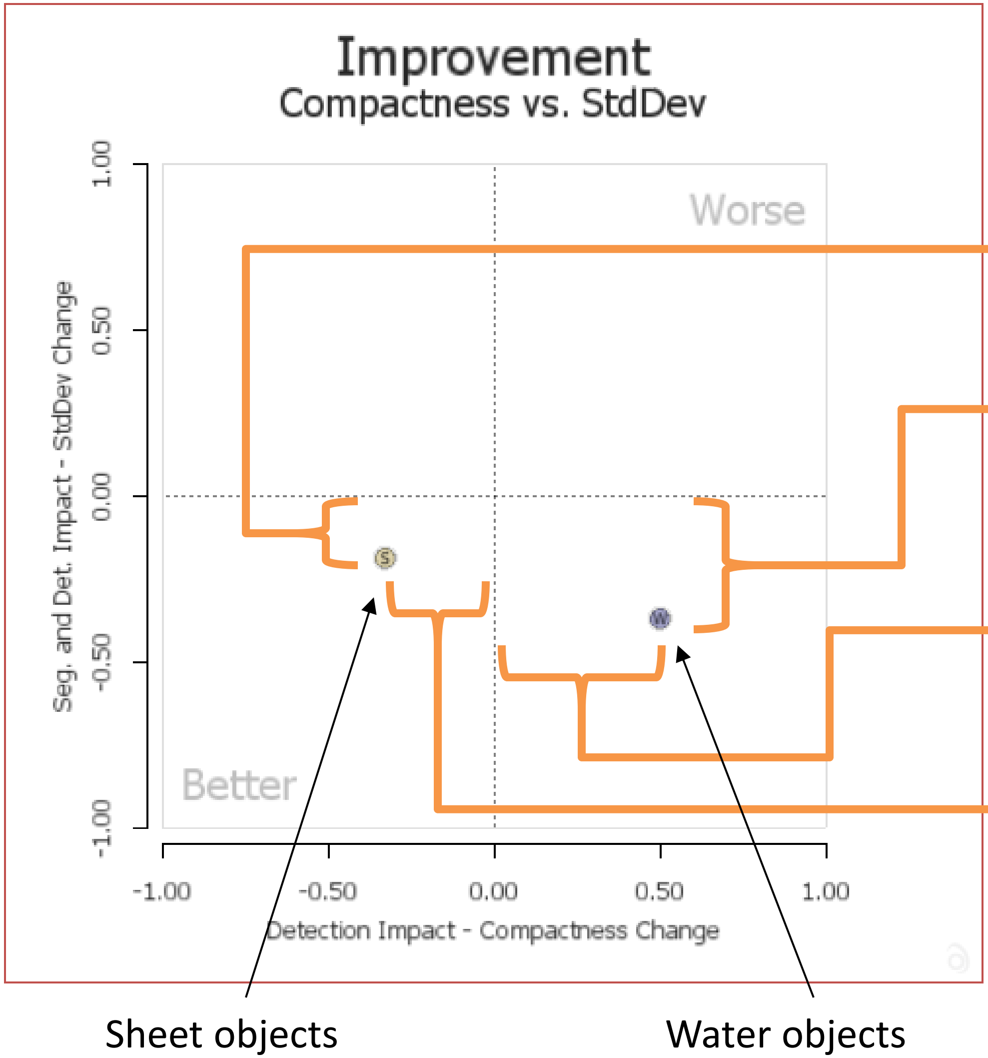
Improvement relative to boundary contrast. Correlates to segmentation quality.

Improvement relative to CCL segmentation. Implies impact on *simple* segmentation algorithms.



Improvement relative to Tumbler segmentation. Implies impact on *sophisticated* segmentation algorithms.

Improvement Interpretation



Better – lower standard deviation for sheet objects.

Better – lower standard deviation for water objects.

Worse – standard deviation cluster increased for water objects. This makes ATR harder.

Better –standard deviation cluster shrunk for sheet objects.

The Results…

In no particular order … same as on FTP site with dual energy groups last.

Purdue

Water: Slightly better standard deviation though less consistent. Slightly reduced and less consistent boundary contrast. Better segmentation accuracy and precision all around.

Saline: Better standard deviation and more consistent. Slightly better boundary contrast. Better segmentation accuracy for CCL. Better segmentation precision.

Sheet: Slightly better standard deviation. No change in boundary contrast. Slightly worse segmentations (stacked sheets problematic).

Bouman\genhuber_mixture_X1\20131019_143014_Cloud_Results

Harvard

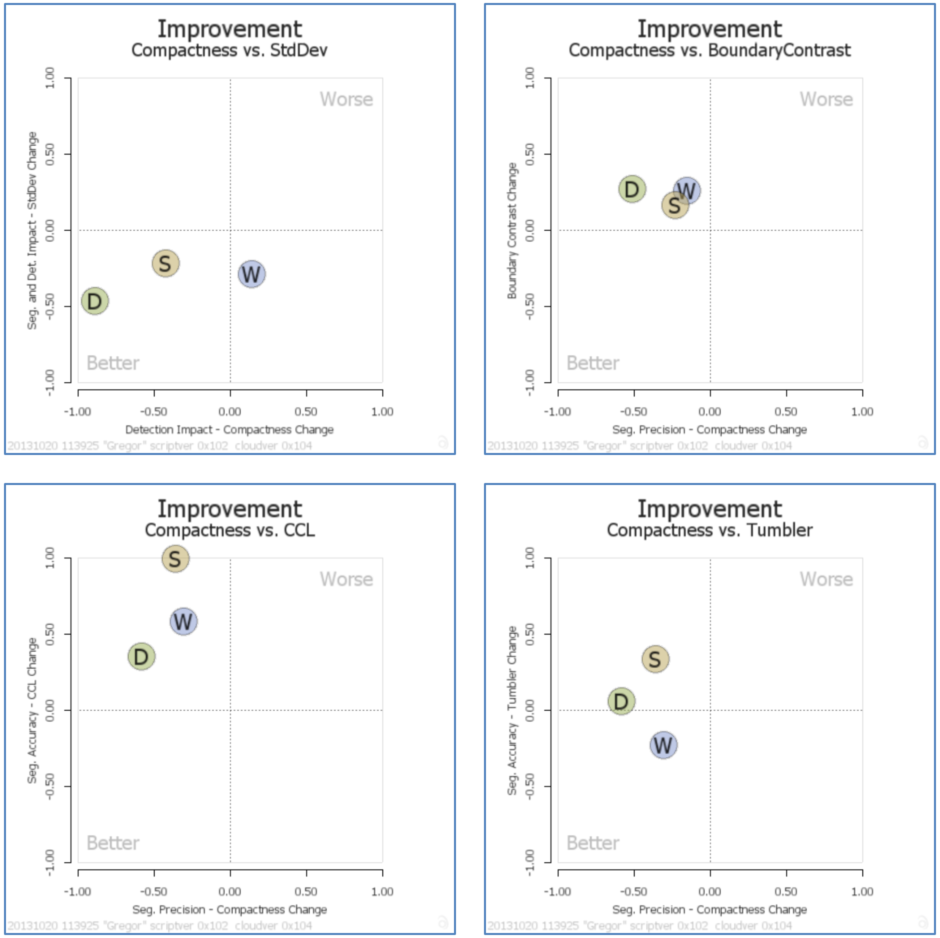
Water: Better standard deviation. Insignificant change to boundary contrast. Slightly better segmentation precision. Slightly better Tumbler accuracy.

Saline: Better standard deviation and more consistent. Reduced boundary contrast but more consistent. No change in segmentation accuracy. Better segmentation precision.

Sheet: Better standard deviation. No change in boundary contrast. Little change in segmentations (stacked sheets problematic).

Do\FITS\SparseRecon\20131018_182230_Cloud_Results

Jens Gregor (Tennessee)



Gregor\Gregor_CGW1B5\20131018_182252_Cloud_Results

Water: Better standard deviation. Reduced boundary contrast but more consistent. Reduced CCL accuracy but more consistent. Better Tumbler accuracy and precision.

Saline: Better standard deviation and more consistent. Reduced boundary contrast but more consistent. Reduced CCL accuracy but more consistent. Better Tumbler precision.

Sheet: Better standard deviation and a bit more consistent. Reduced boundary contrast but more consistent. Worse CCL and Tumbler segmentations (stacked sheets problematic).

UCSD

Water: Better standard deviation. Slightly reduced boundary contrast. Better segmentation accuracy. No change in segmentation precision.

Saline: Too few objects.

Sheet: Better standard deviation and more consistent. Insignificant change in boundary contrast. Worse segmentation accuracy (stacked sheets problematic).

Karimi\mar\20131018_182255_Cloud_Results

Chicago

Water: No change in standard deviation but less consistent. No change in boundary contrast but less consistent. Better CCL accuracy. Less segmentation precision.

Saline: Standard deviation is more consistent. Insignificant change in boundary contrast. Better CCL accuracy. Slightly worse Tumbler accuracy. Less segmentation precision.

Sheet: No change in standard deviation. Insignificant change in boundary contrast. Slightly better segmentation accuracy. Worse segmentation precision (stacked sheets problematic).

LaRiviere2\C111\20131018_175843_Cloud_Results

Utah

Water: Better standard deviation but less consistent. Reduced boundary contrast but more consistent. Better segmentation accuracy. No change in segmentation precision.

Saline: Too few objects.

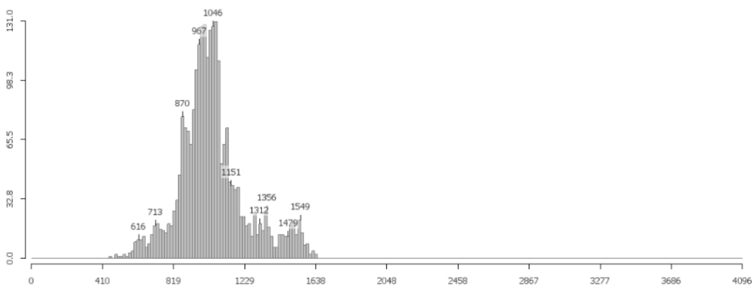
Sheet: (No stacked sheets.) Better standard deviation. Improved boundary contrast consistency. Better CCL accuracy. Little change in Tumbler accuracy. Better segmentation precision (though, no stacked sheets).

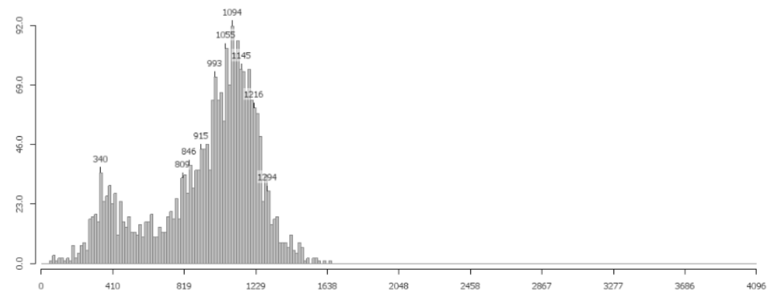
Zeng\ver4\20131018_175851_Cloud_Results

Purdue – Doped Water (Better)

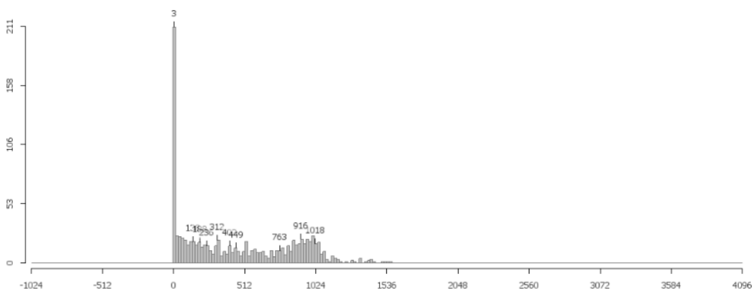
Purdue | XRec

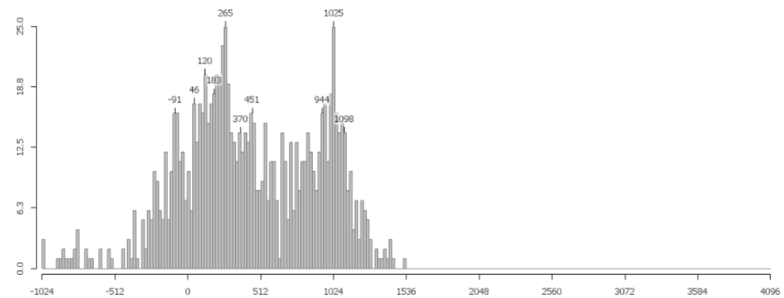
Object interior





Object boundary



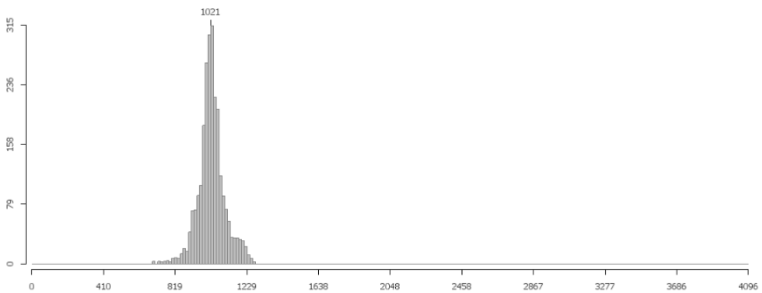


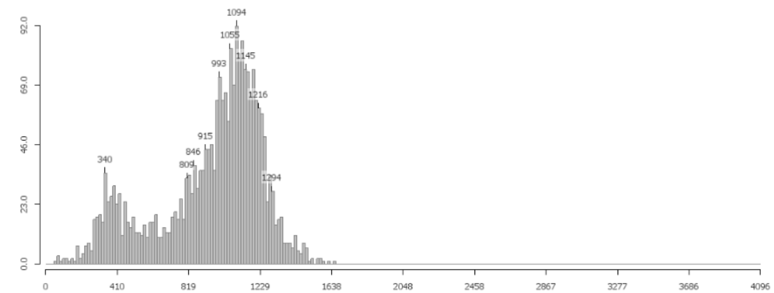
High_Clutter1 Slice.239

Harvard – Doped Water (Better)

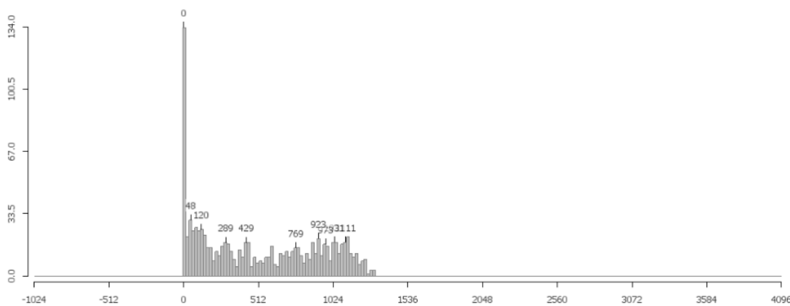
Harvard | XRec

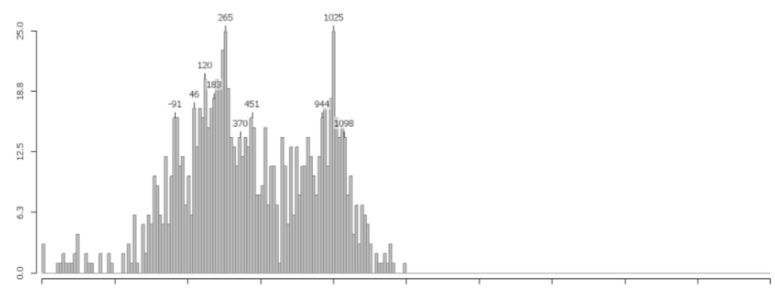
Object interior





Object boundary





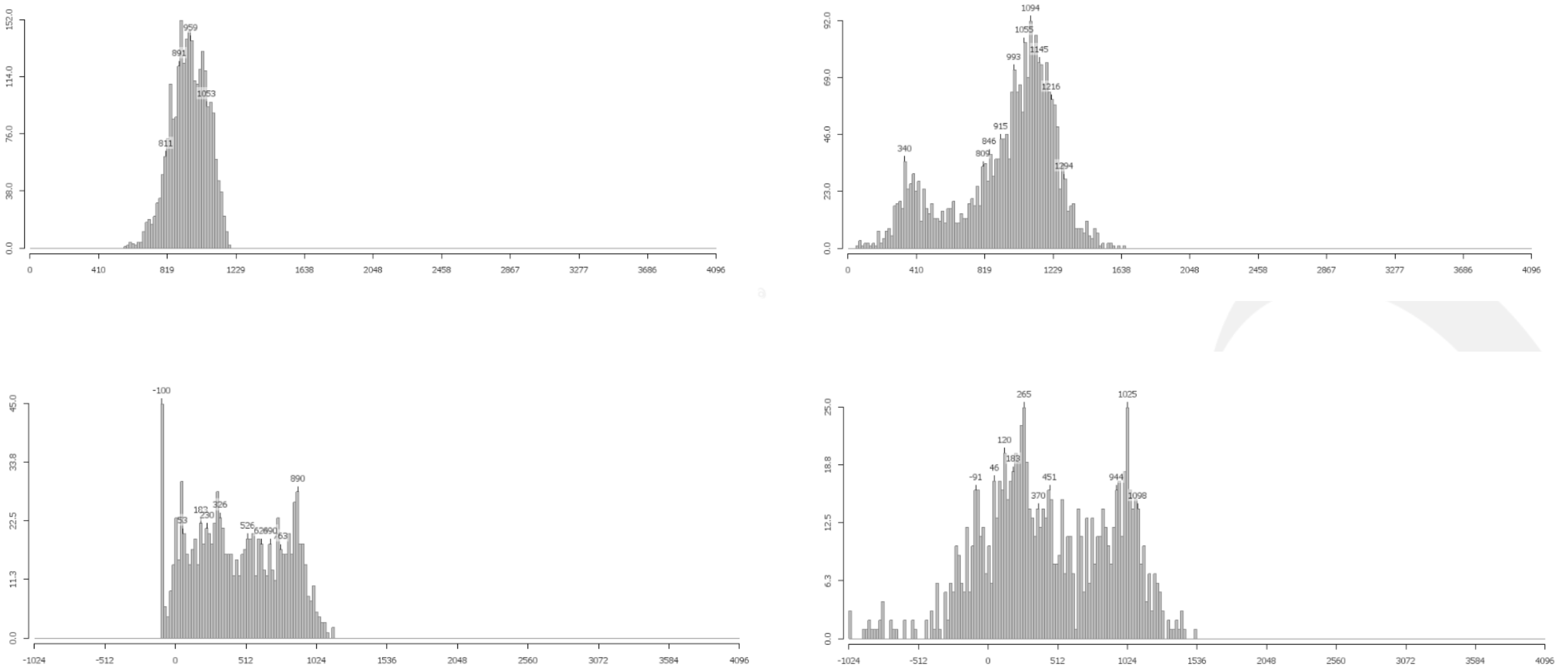
High_Clutter1 Slice.239

Gregor – Doped Water (Better)

Gregor | XRec

Object interior

Object boundary

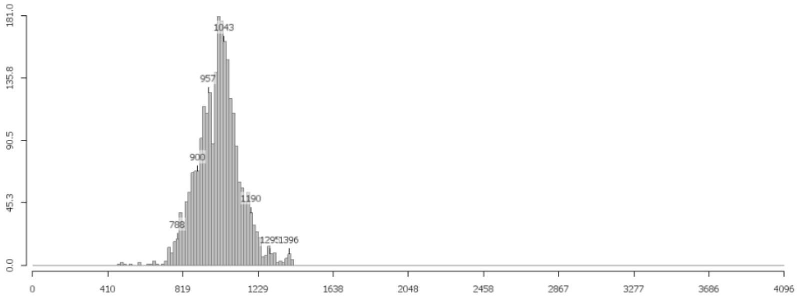


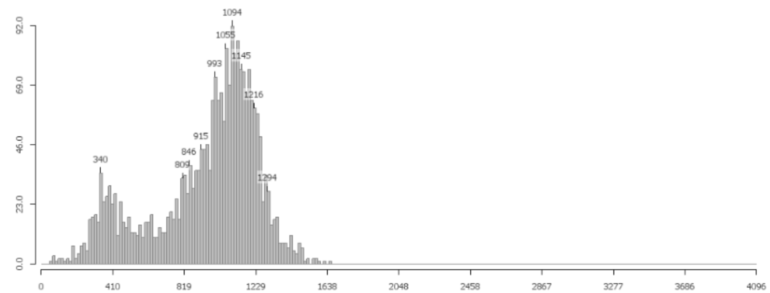
High_Clutter1 Slice.239

UCSD – Doped Water (Better)

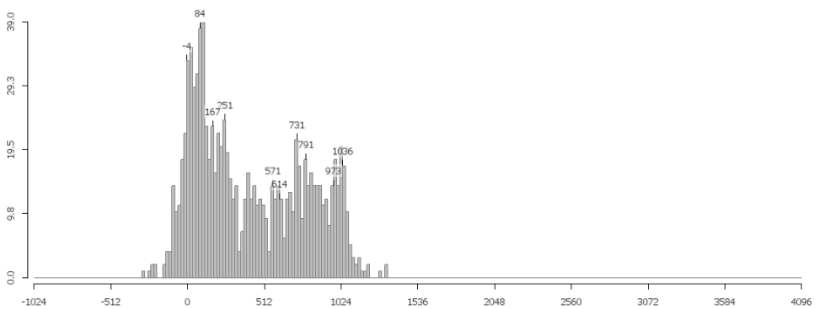
UCSD | XRec

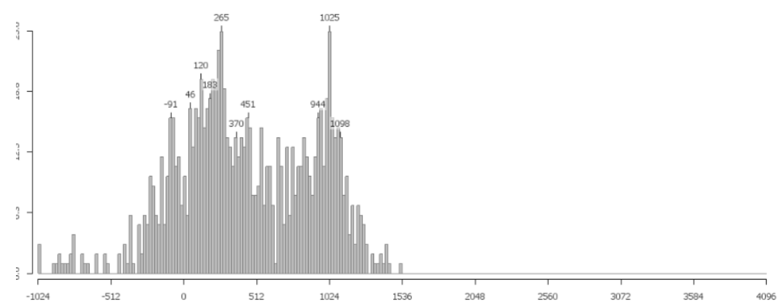
Object interior





Object boundary





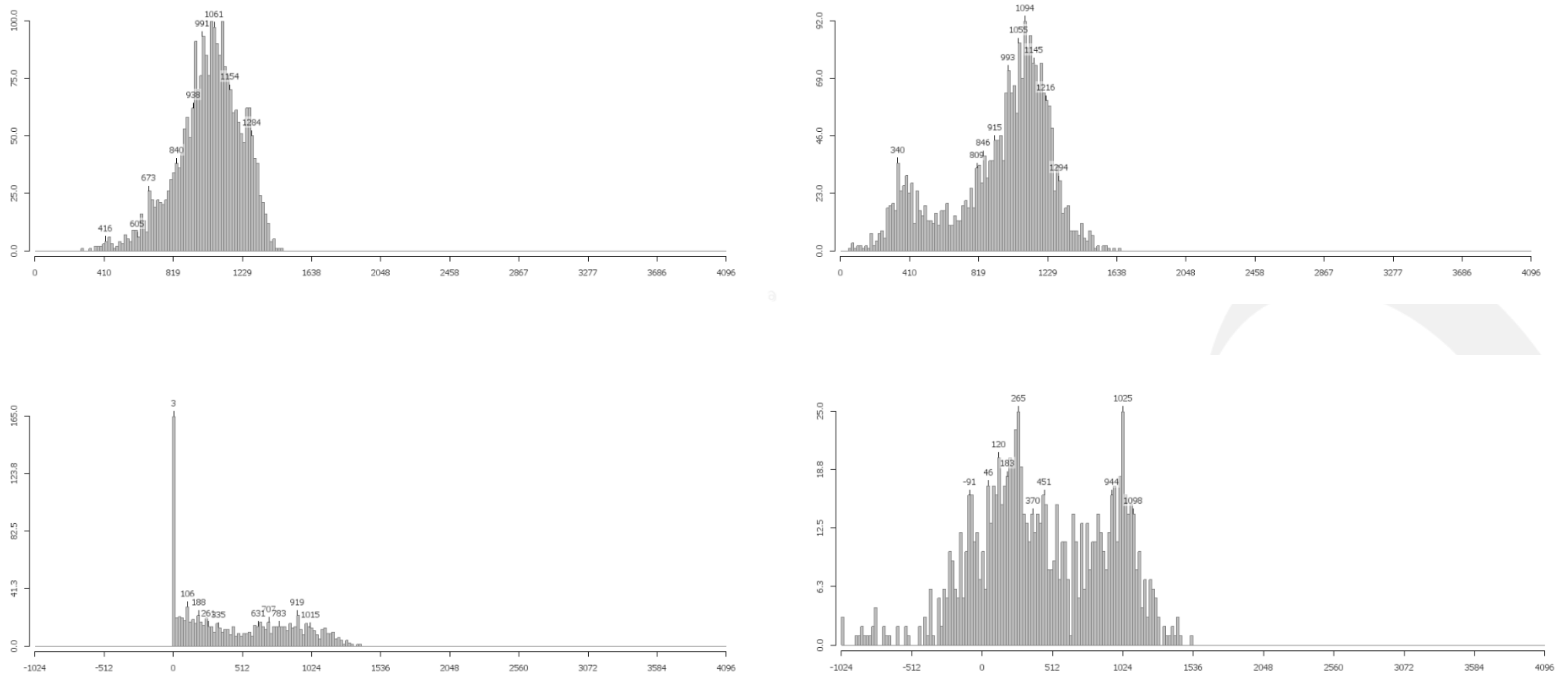
High_Clutter1 Slice.239

Chicago – Doped Water (Better)

Chicago XRec

Object interior

Object boundary

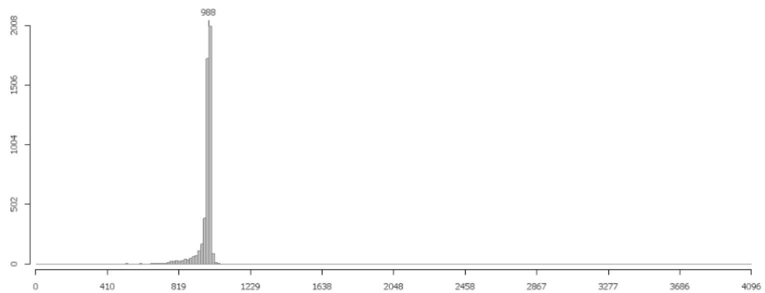


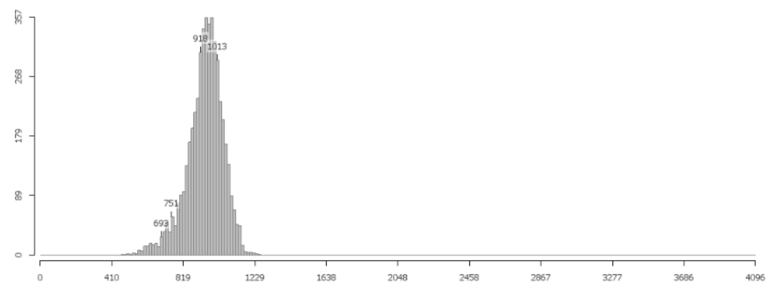
High_Clutter1 Slice.239

Utah – Water (Better)

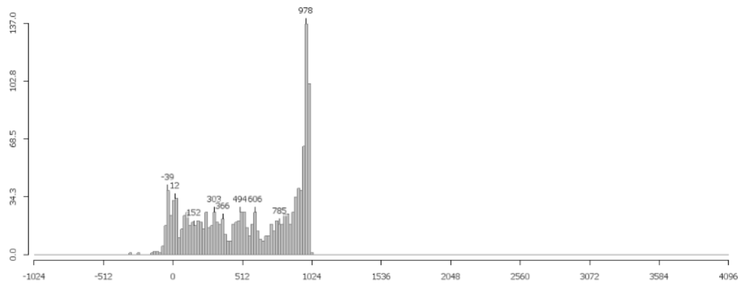
Utah | XRec

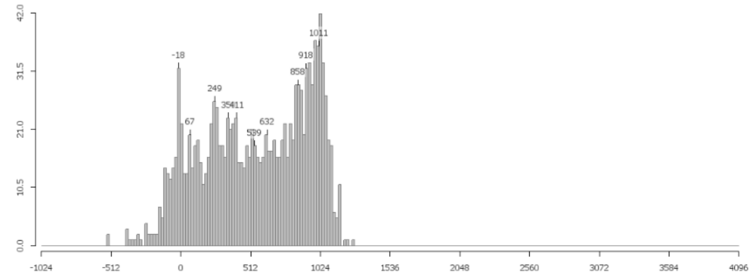
Object interior





Object boundary





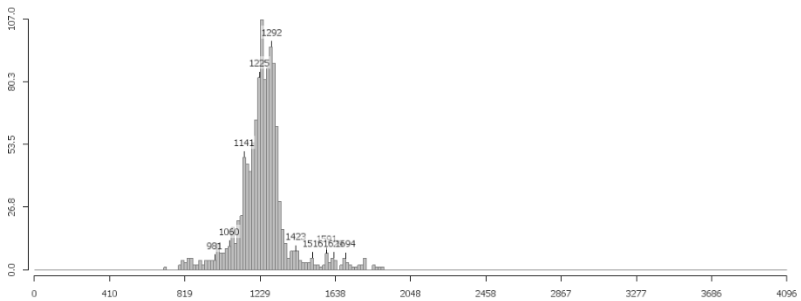
Medium_Clutter1 Slice.231

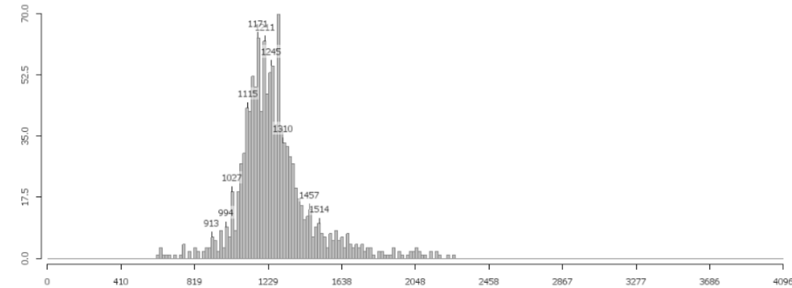
Purdue - Rubber Sheet (Worse)

Purdue

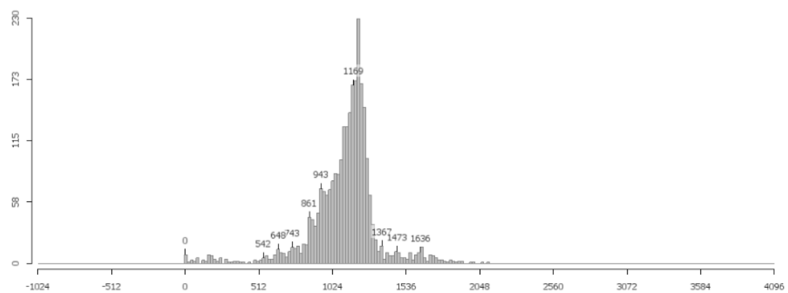
XRec

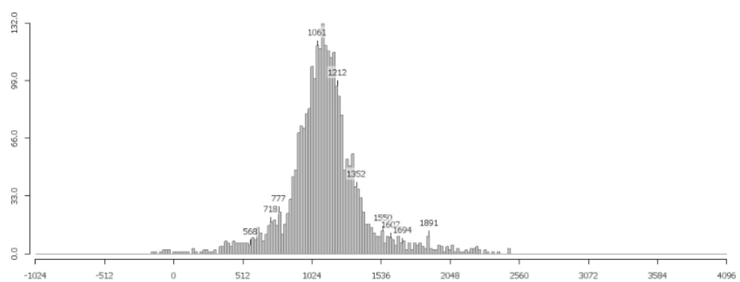
Object interior





Object boundary





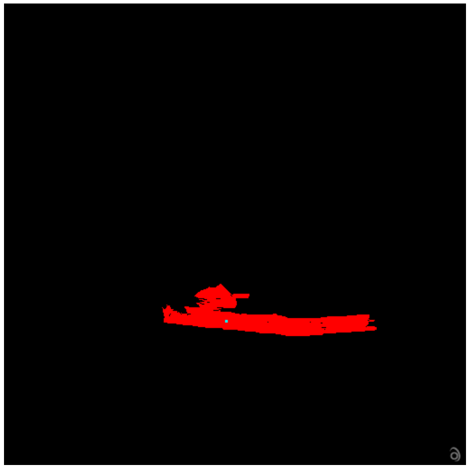
Everyone had trouble with stacked sheets!

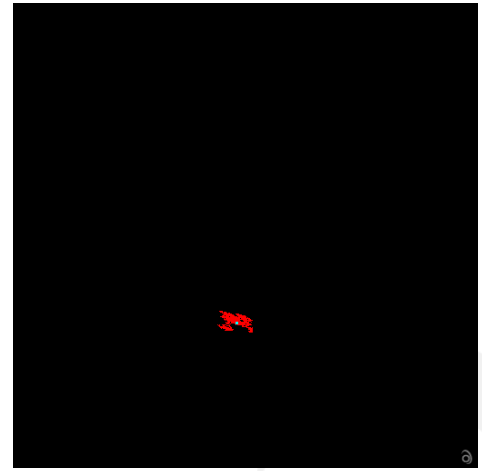
High_Clutter1 Slice.239

Purdue – Rubber Sheet (Worse)

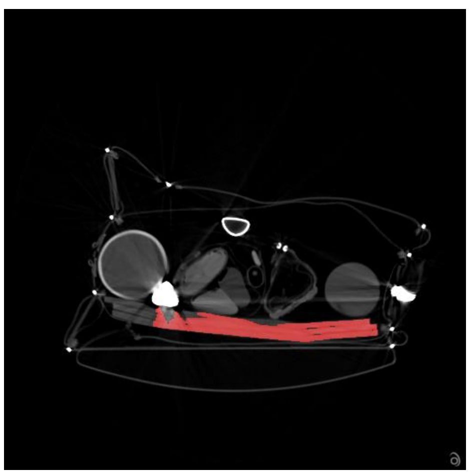
Purdue | XRec

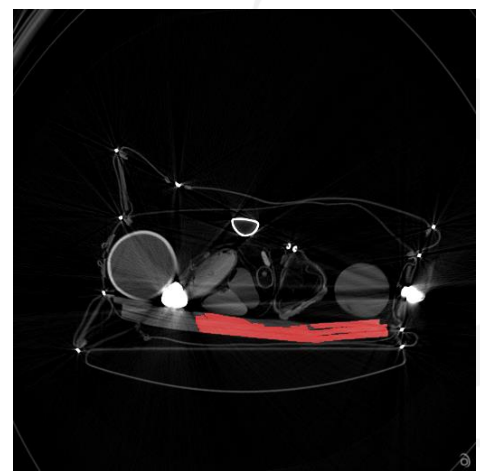
CCL





Tumbler





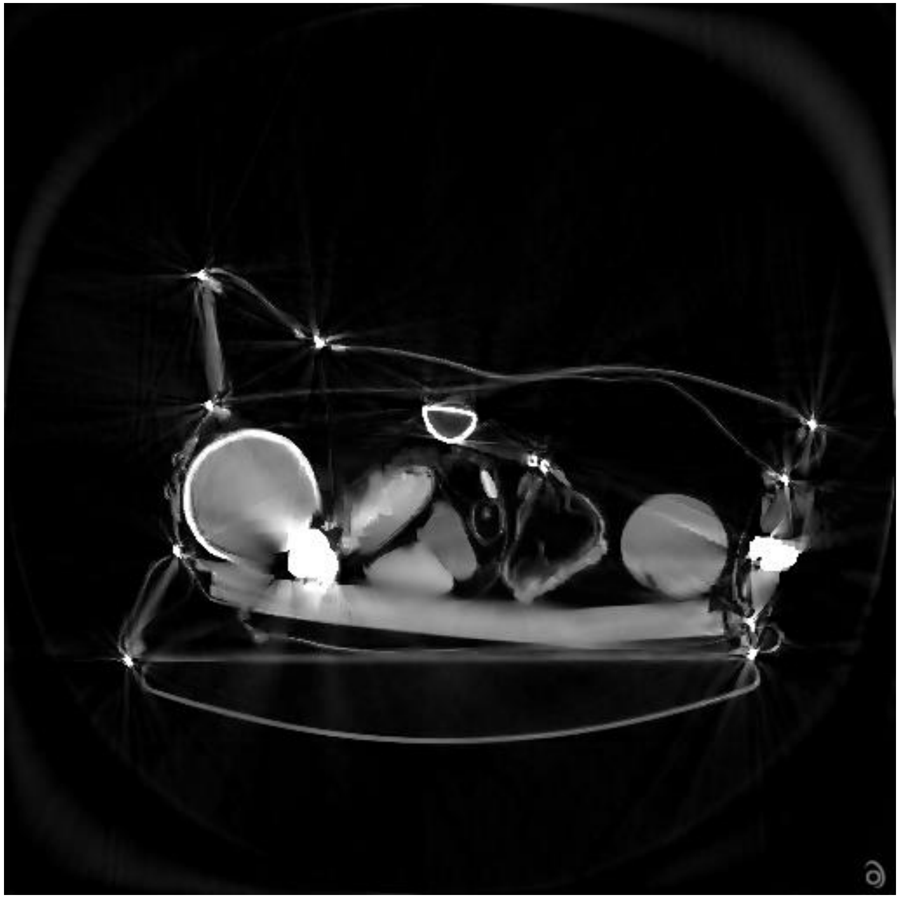
High_Clutter1 Slice.239

Stacked Rubber Sheets

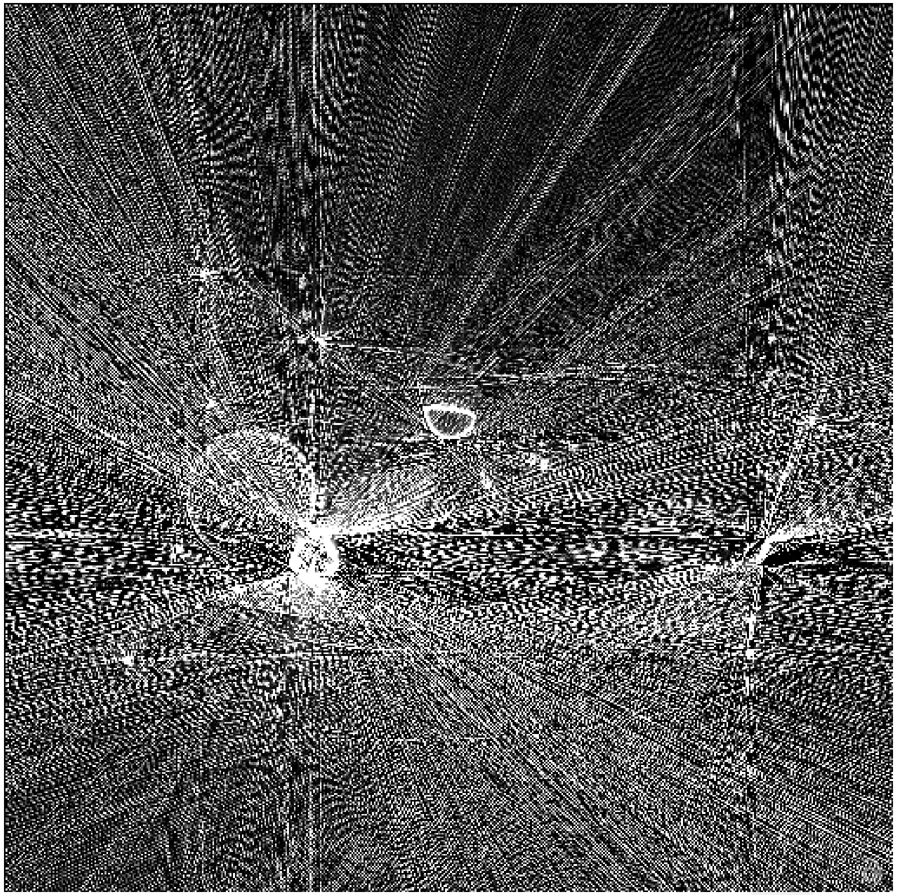
- All groups had trouble with stack sheets
- We won't show stacked sheets results for any more groups since they are all about the same
- Stacked sheets are a resolution problem, not necessarily a recon problem
- All groups did better on a single sheet

Boston – LAC – Doped Water

Boston



YNC



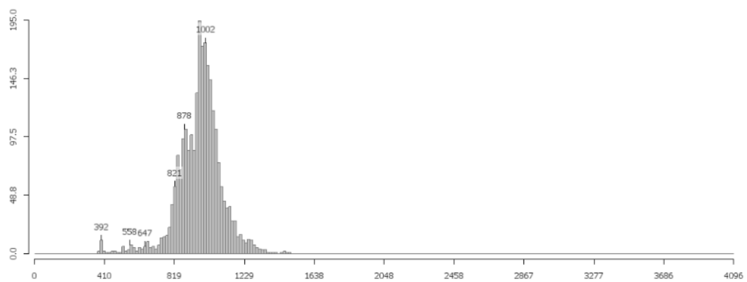
High_Clutter1 Slice.239

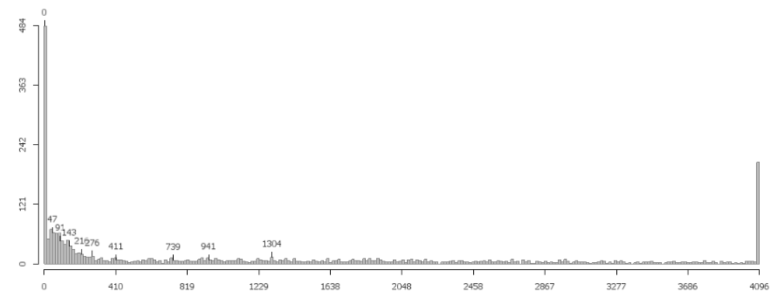
Boston – LAC – Doped Water

Boston

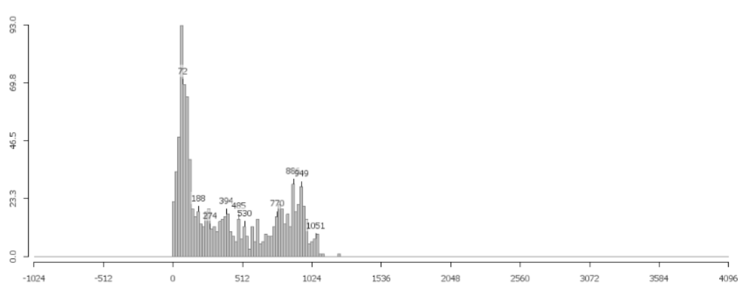
YNC

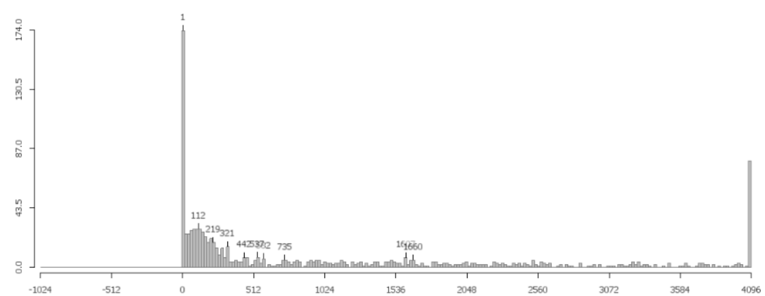
Object interior





Object boundary

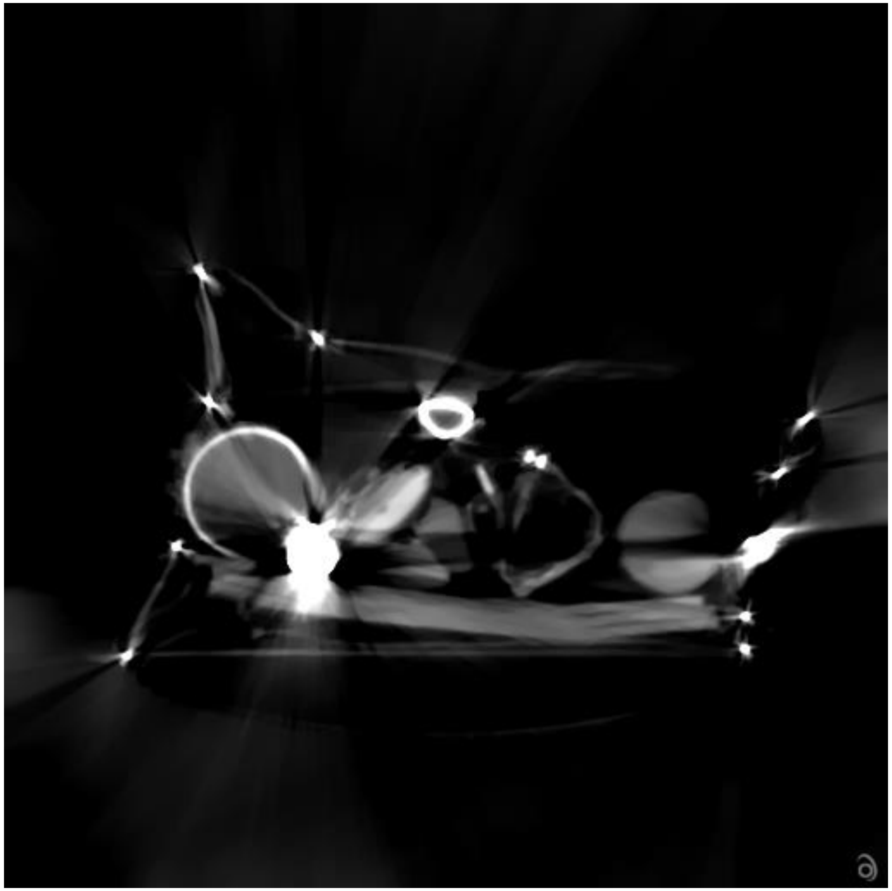




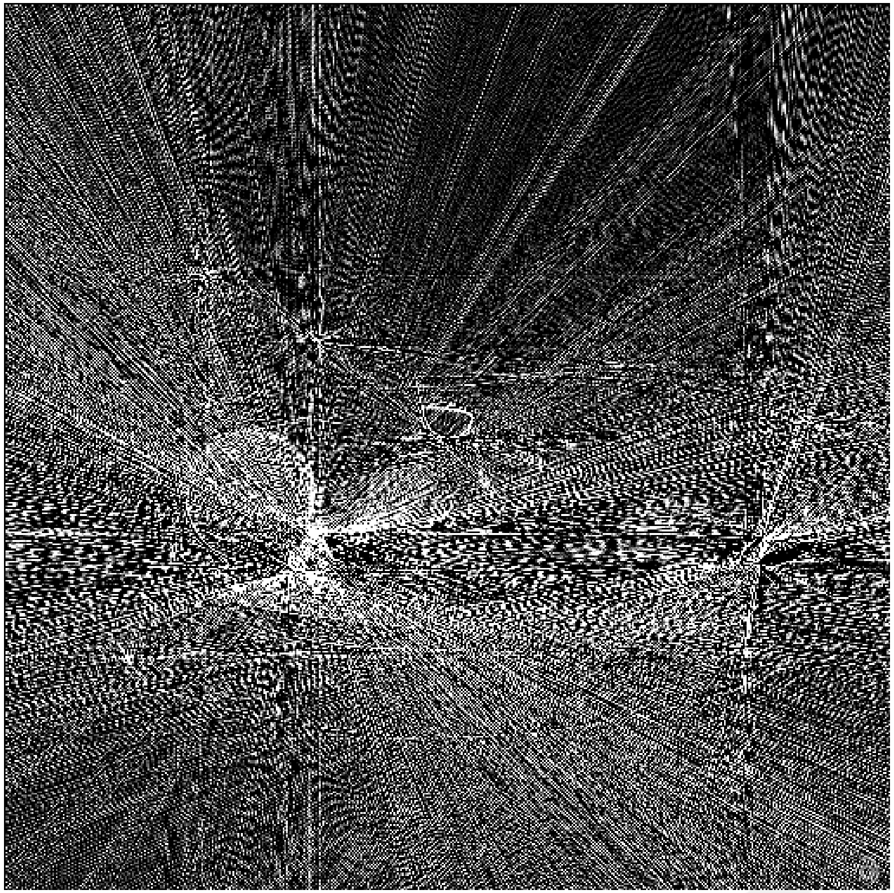
High_Clutter1 Slice.239

Tufts – Compton – Doped Water

Tufts



YNC



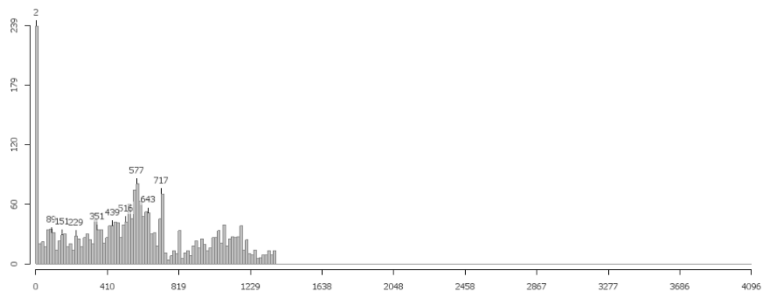
High_Clutter1 Slice.239

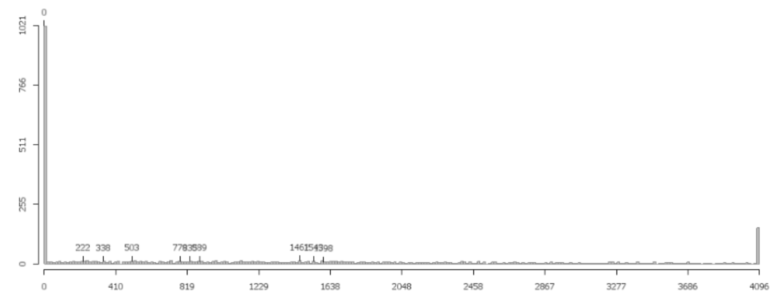
Tufts – Compton – Doped Water

Tufts

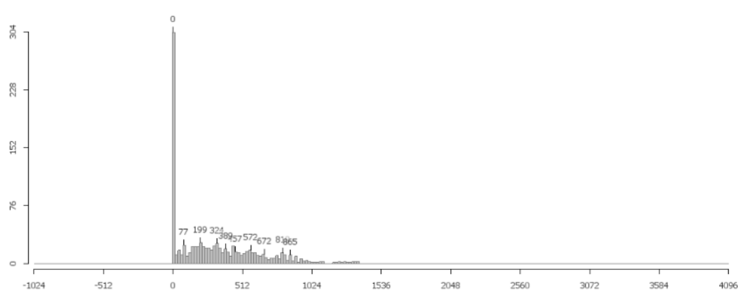
YNC

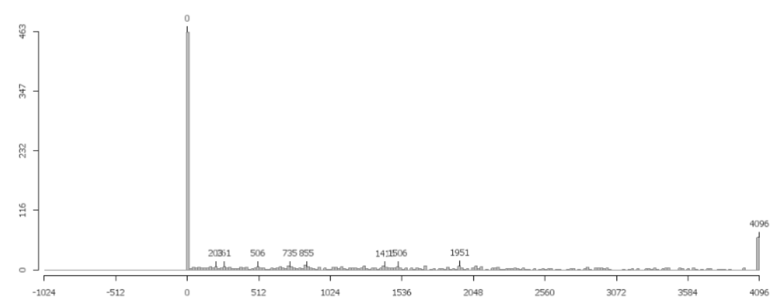
Object interior





Object boundary

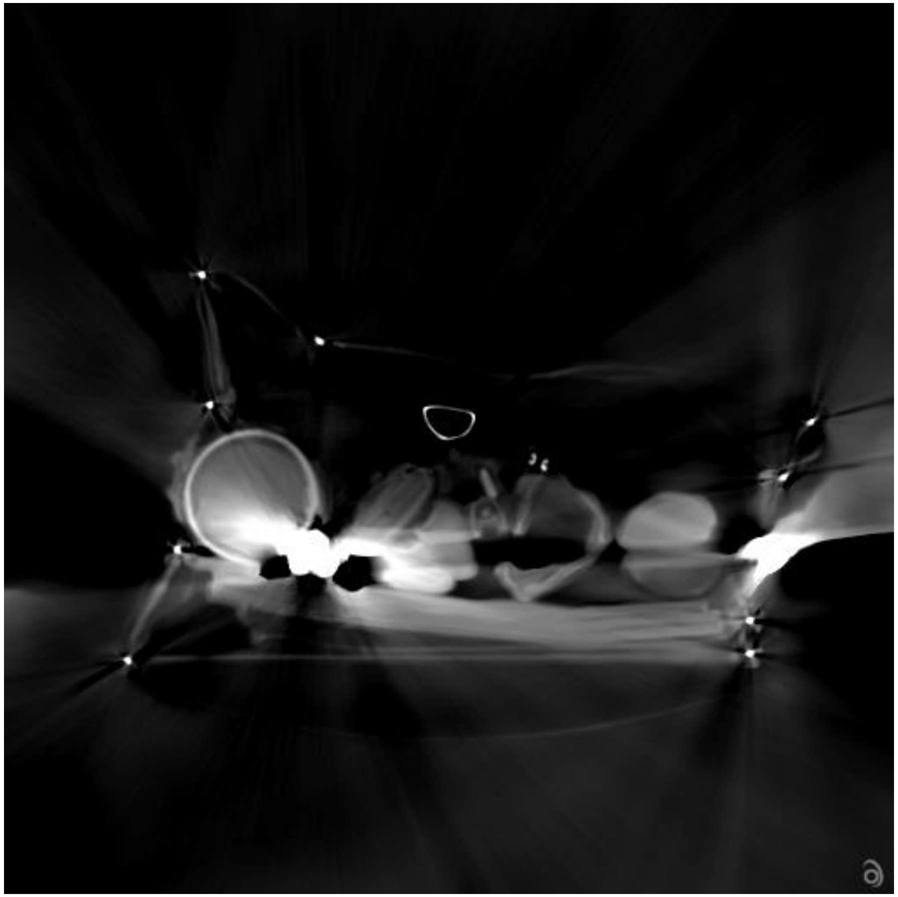


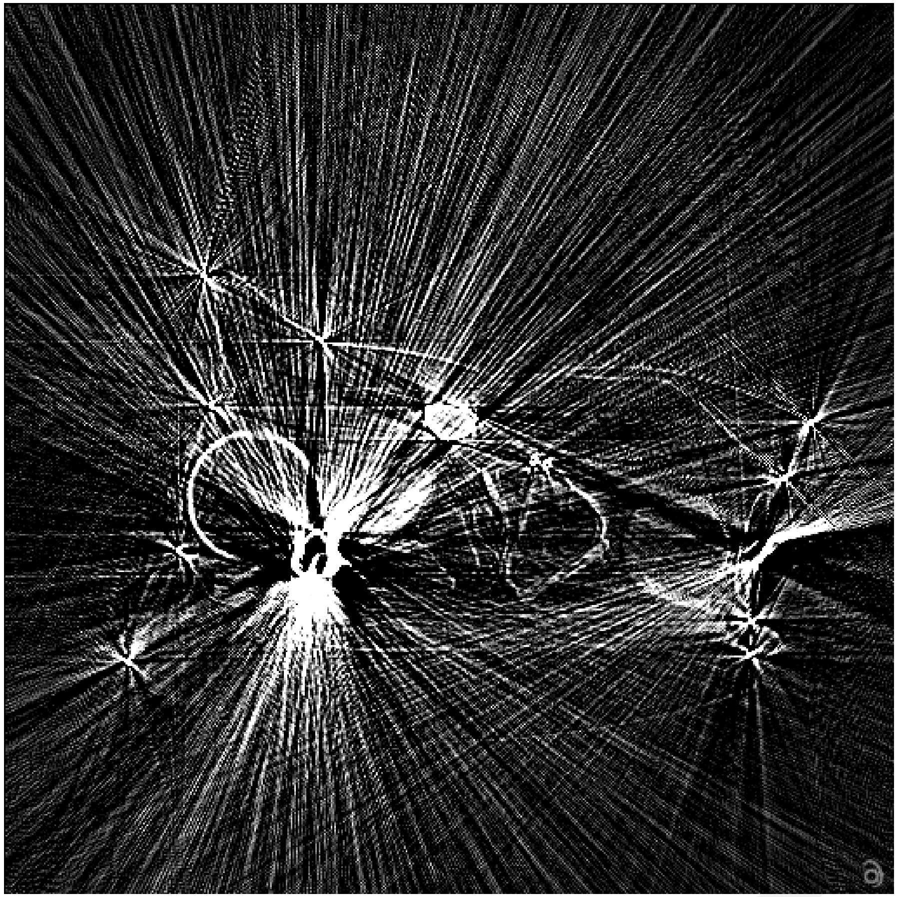


High_Clutter1 Slice.239

Tufts – Photoelectric – Doped Water

Tufts | YNC





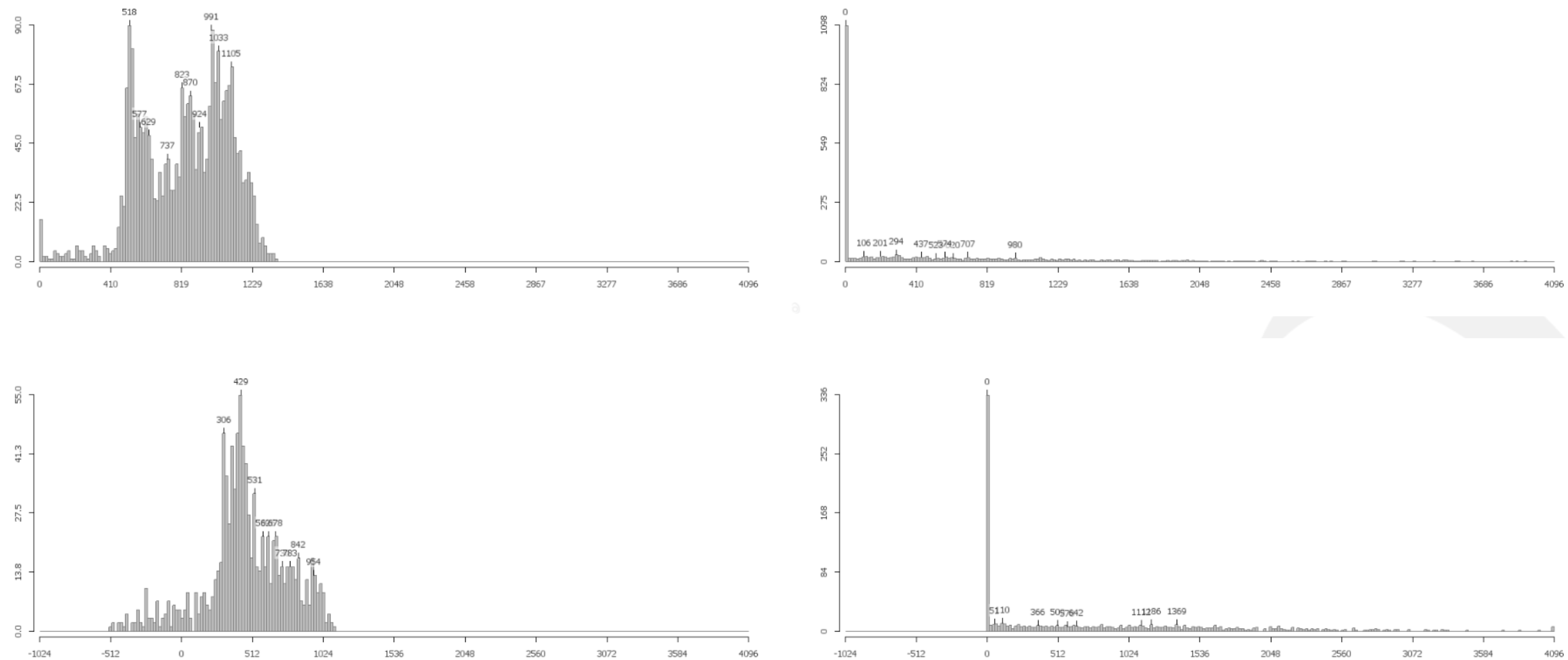
High_Clutter1 Slice.239

Tufts – Photoelectric – Doped Water

Tufts | YNC

Object interior

Object boundary



High_Clutter1 Slice.239

Everyone made progress!