SAND2020-7836PE

Deep Learning for Characterization of Geometric Uncertainty

Presented by: Cari Martinez (9323)

Sandia MLDL Workshop, August 5, 2020

Sandia National Laboratories

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Outline

CAMI LDRD project motivation

Automated CT Segmentation

Geometric Uncertainty Characterization for DL Segmentation Models

- Monte Carlo Dropout Network
- Bayesian Convolutional Neural Network

Can we predict the behavior of as-built parts with error bars?

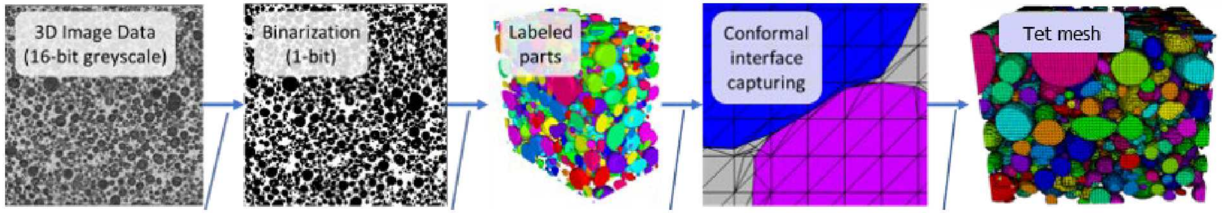
Credible Automated Meshing of Images (CAMI) LDRD, Scott Roberts (1513), PI

Hypothesis: We can develop an automated and credible image-to-mesh technology that can demonstrate the physics impact of per-unit variability on material, component, or system performance

Objective: We seek to develop a methodology for **automatically**, **efficiently**, and **reproducibly** creating **conformal** finite element meshes from **3D tomography** with **quantified uncertainty**.

Research thrusts – primary science questions:

1. Automatic CT segmentation via Machine Learning (ML)
2. Automatic conformal tetrahedral mesh creation (ATM)
3. Uncertainty quantification and propagation (UQ)



Uncertainty quantification and propagation

Deep learning algorithms
- Image segmentation
- Part identification

Automatic tetrahedral meshing
- Conformal interfaces
- Feature-governed mesh resolution

Physics solve
- Finite element method predictions

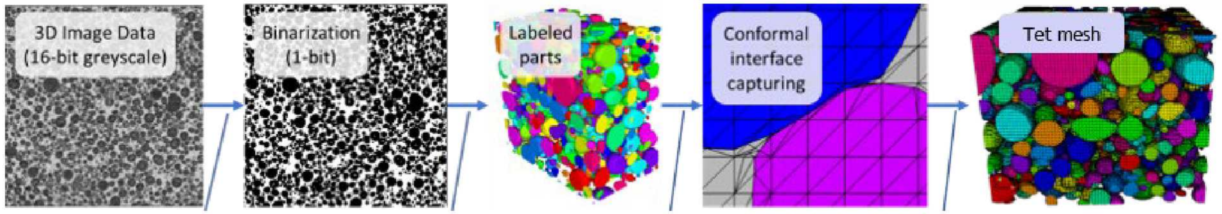
Can we predict the behavior of as-built parts with error bars?

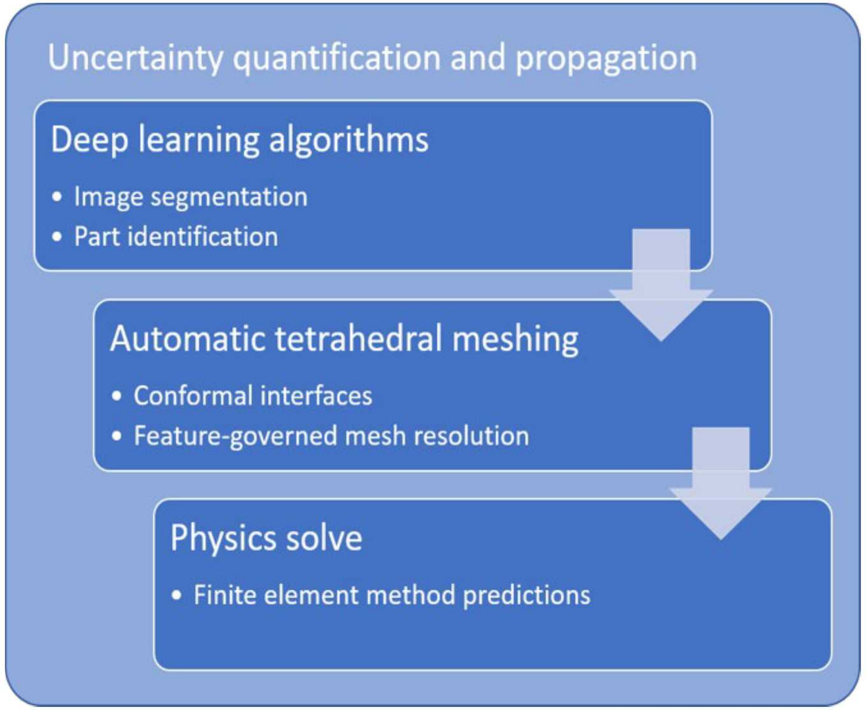
Hypothesis: We can develop an automated and credible image-to-mesh technology that can demonstrate the physics impact of per-unit variability on material, component, or system performance

Objective: We seek to develop a methodology for **automatically**, **efficiently**, and **reproducibly** creating **conformal** finite element meshes from **3D tomography** with **quantified uncertainty**.

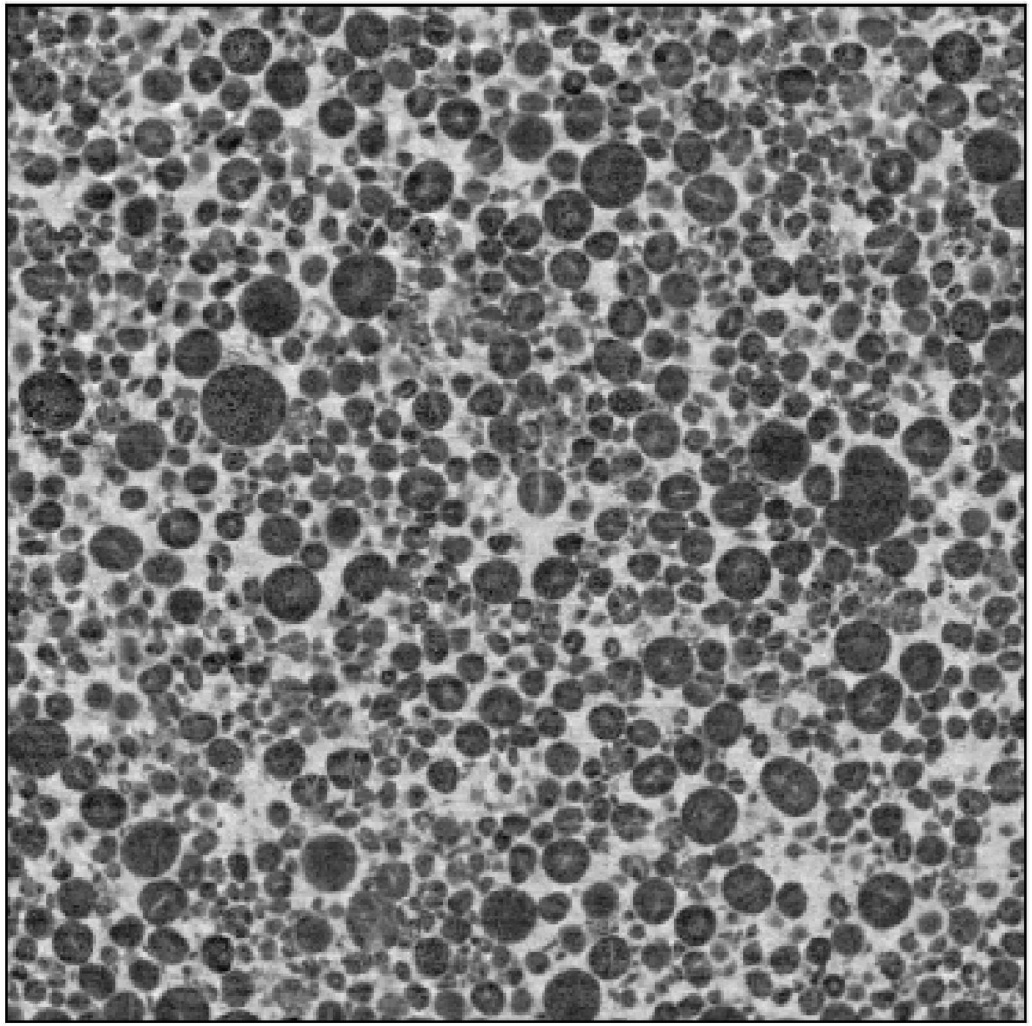
Research thrusts – primary science questions:

1. Automatic CT segmentation via Machine Learning (ML)
2. Automatic conformal tetrahedral mesh creation (ATM)
3. Uncertainty quantification and propagation (UQ)

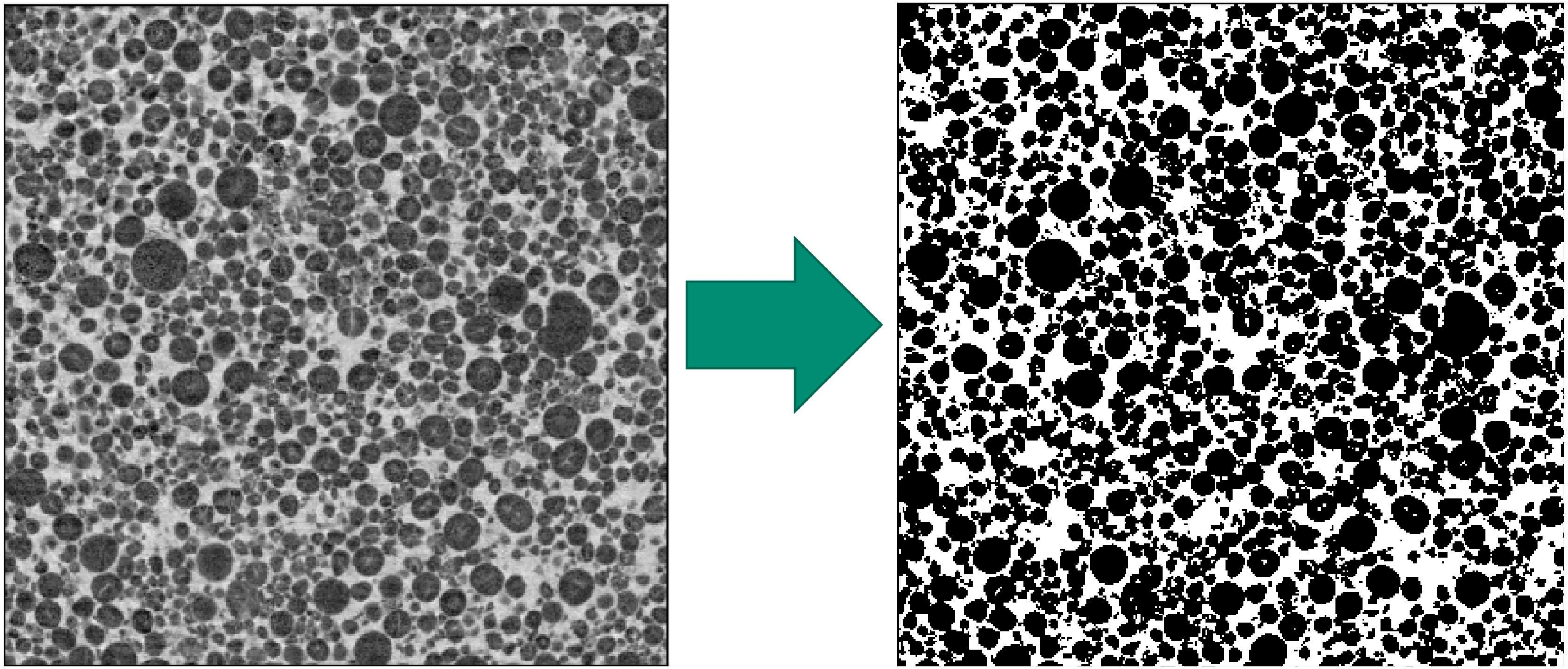




CT produces raw images unsuitable for simulation

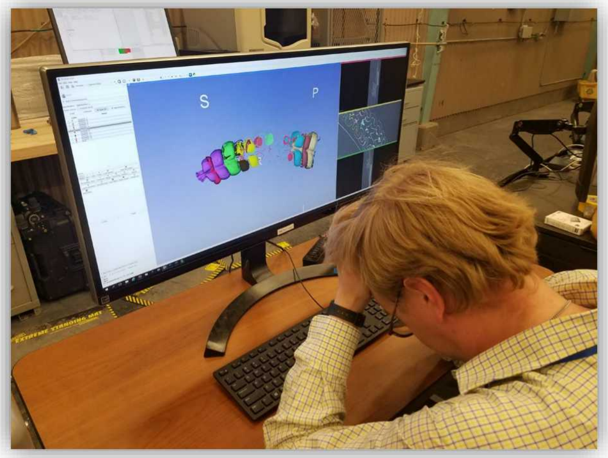


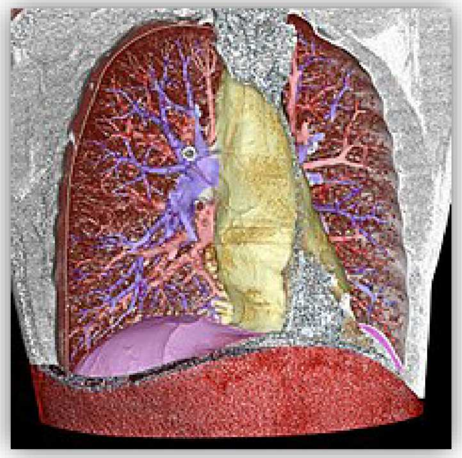
In order to use in simulations, we need to segment the image

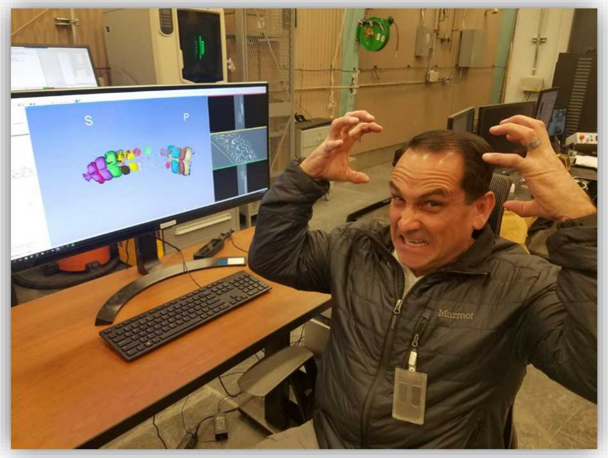


CT Segmentation is hard for humans (and painful)

CT scans must be labeled by component for simulations



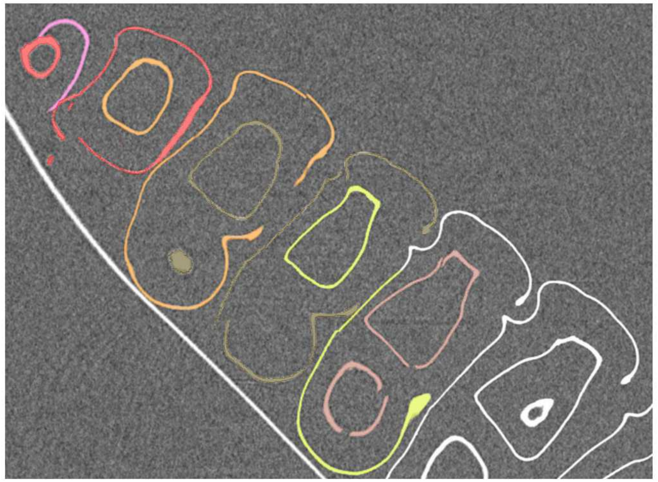


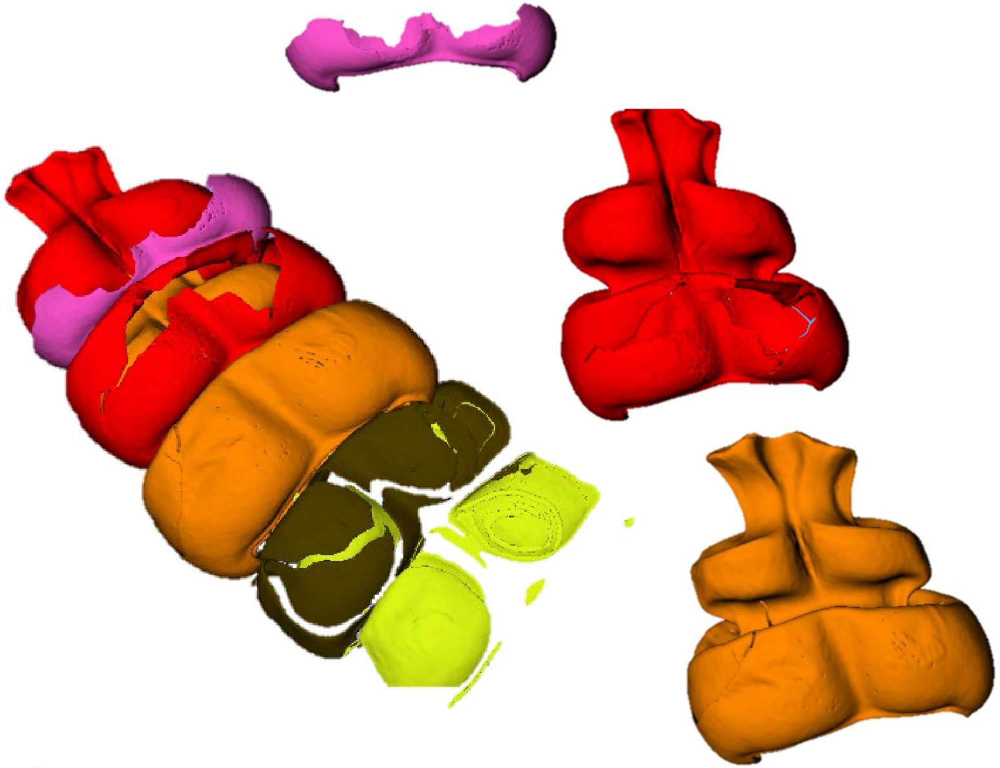


https://en.wikipedia.org/wiki/Image_segmentation

Labeling by hand does not scale

Deep Learning algorithms can segment CT scans automatically
- Find each component of the object by material
- Find any anomalies
- Pass this to a usable form for numerical simulations



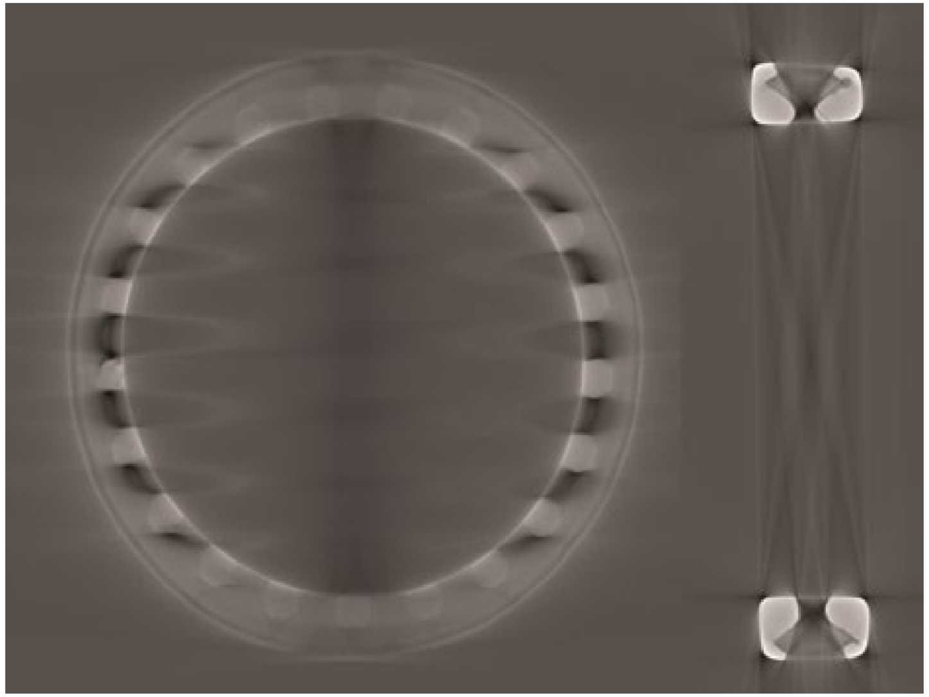


Using ML to save time and effort while improving accuracy

AUTOMATED CT SEGMENTATION

Thresholding is not enough to get good segmentations of dense materials

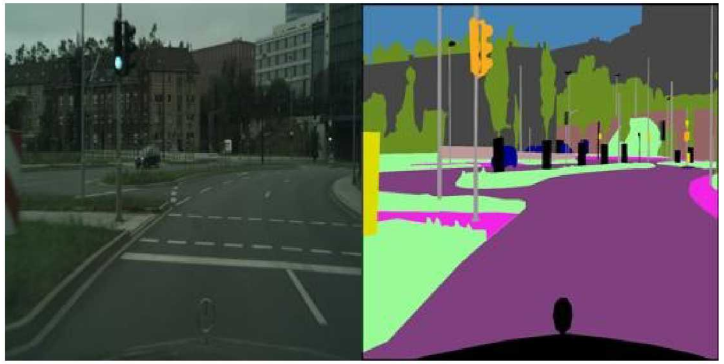
High density material can cause bright and dark "shadow" artifacts in the CT reconstruction



Segmentation is a classic computer vision problem

Image segmentation is well studied

- Small files
- Large training sets



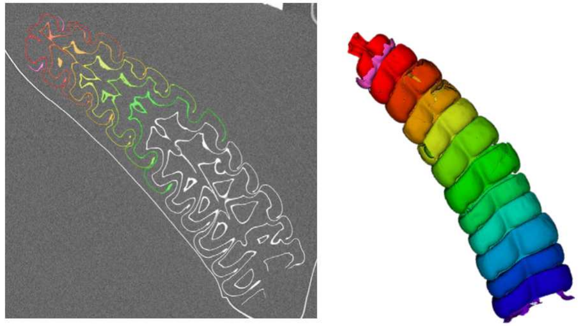
https://www.cityscapes-dataset.com/

Cityscape
(~1e5 pixels)

Volumetric segmentation is different

- Big data
- Class imbalance (lots of background)
- Small training sets with "bad" human labels
- Humans can't label a billion voxels without taking big shortcuts

Medical researchers are leading this work toward Deep Learning solutions



Rattlesnake Tail
(~1e9 voxels)

Mitigating challenges with CS/ML/statistics techniques

CT scans are large

- Used chunks of the volume
- Optimized our model for GPU memory usage on GPU cluster

Class Imbalance

- Adjusted loss function that guides training

Artifacts and noise

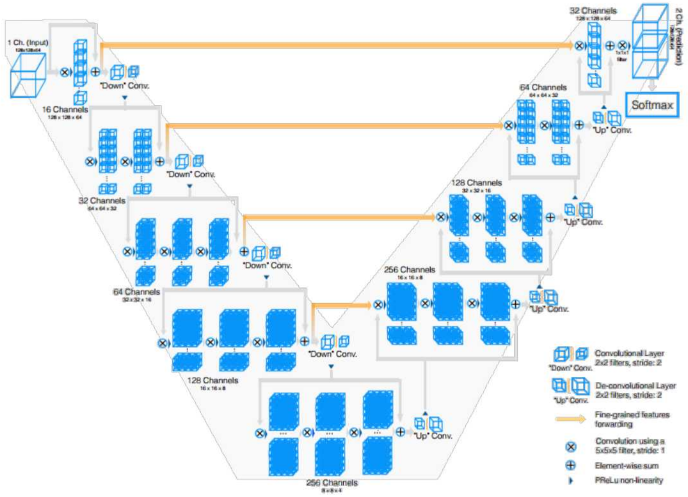
- Selected Convolutional Neural Network (CNN) architecture with strength in shape recognition

Deep learning has a wide set of existing solutions

"Deep learning is the first class of algorithms that is scalable... performance just keeps getting better as you feed them more data"

Andrew Ng, Founder of Google Brain

- U-net, a big advance in biomedical segmentation
  - Olaf Ronneberger, Philipp Fischer, Thomas Brox , "U-Net: Convolutional Networks for Biomedical Image Segmentation", in Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer, LNCS, Vol.9351: 234--241, 2015
- V-net follows as a natural extension
  - F. Milletari, N. Navab, and S. A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in 2016 Fourth International Conference on 3D Vision (3DV), Oct 2016, pp.565–571

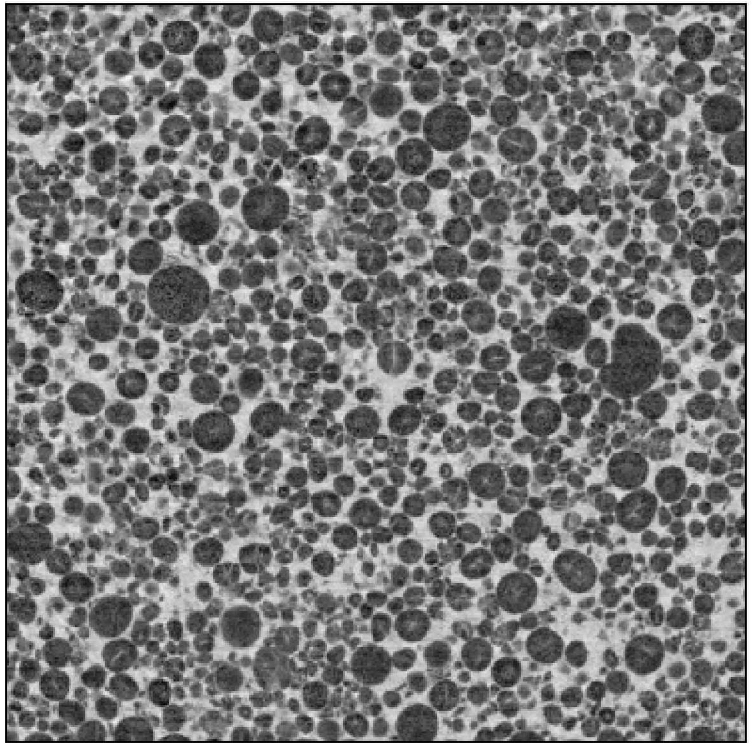


V-Net architecture for segmenting volumetric data (2016)

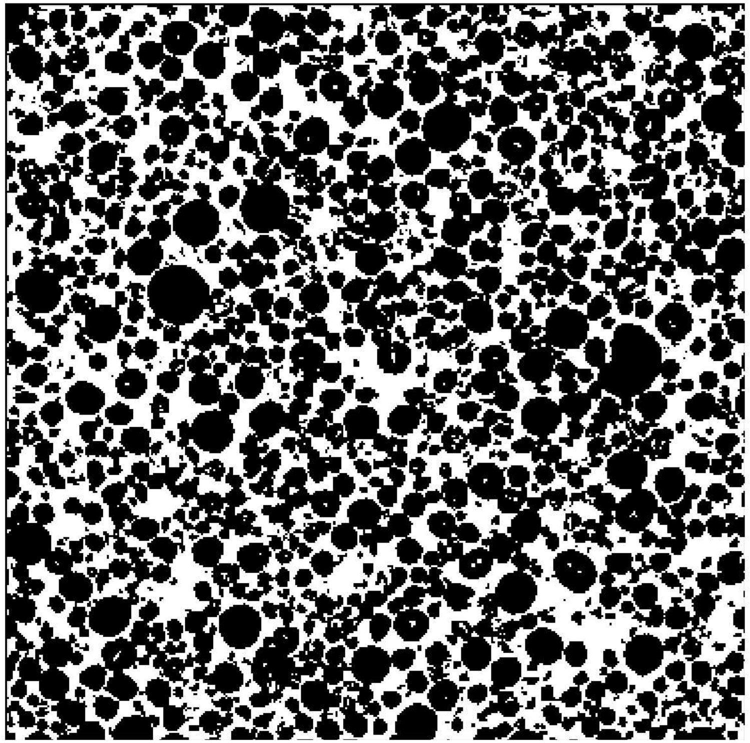
We started with a V-Net and made improvements as necessary

Volumetric battery segmentation achieves high accuracy

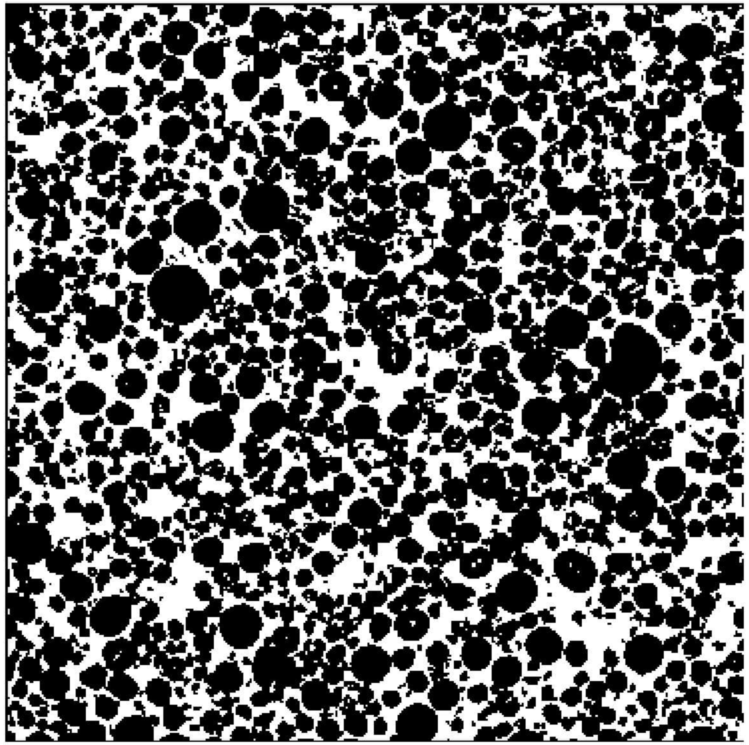
Slice of 3D Image



Human label



ML prediction

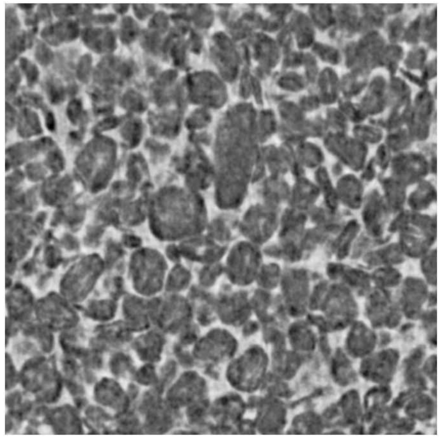


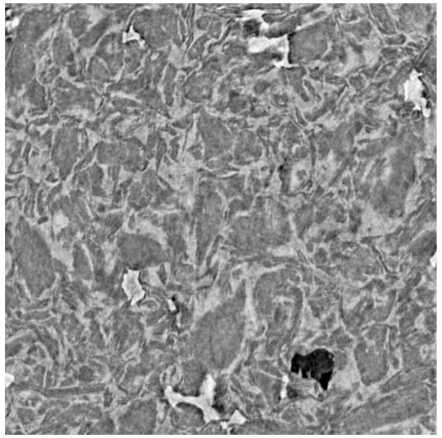
Averaged 99.7% accuracy compared to human labels over held out test set

Extending segmentation model to different battery types seems challenging

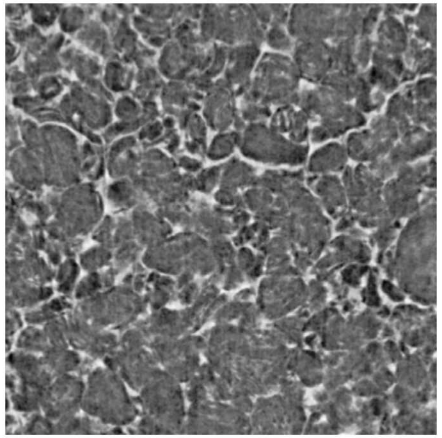
DOMAIN NAME	ACCURACY
E35	0.984
Tesla	0.973
Litarion	0.966
25R6	0.955
Electrode_I_1	0.948
Electrode_III_1	0.945
GCA400	0.928
Electrode_IV_1	0.917
Electrode_II_2	0.902
GCA2000	0.900
Electrode_I_2	0.892
Electrode_III_2	0.773
Electrode_IV_3	0.748
Electrode_IV_2	0.745
Electrode_II_3	0.699
Electrode_III_3	0.668
Mean	**0.8714375**

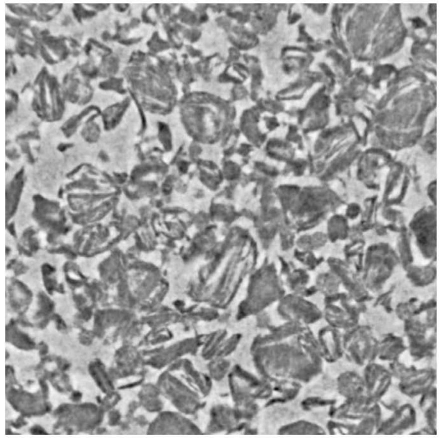
TRAINING SET

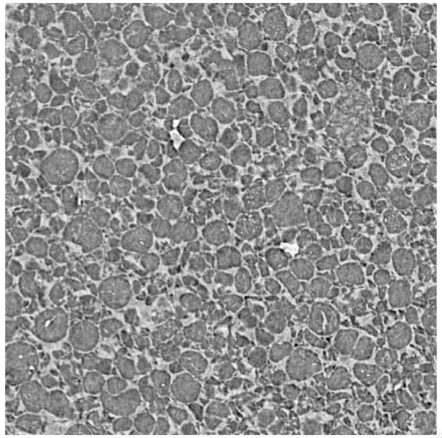

Litarion


Electrode IV_1

TEST SET

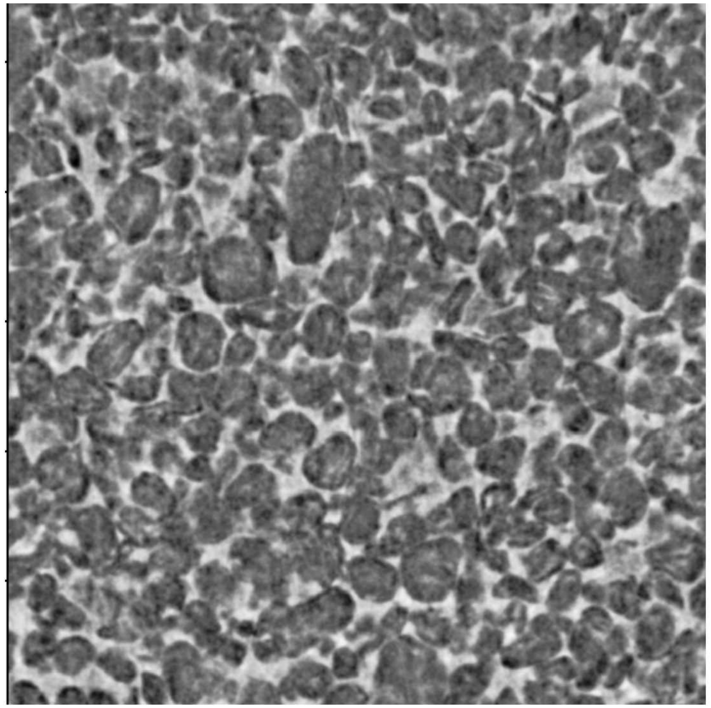

E35


GCA400

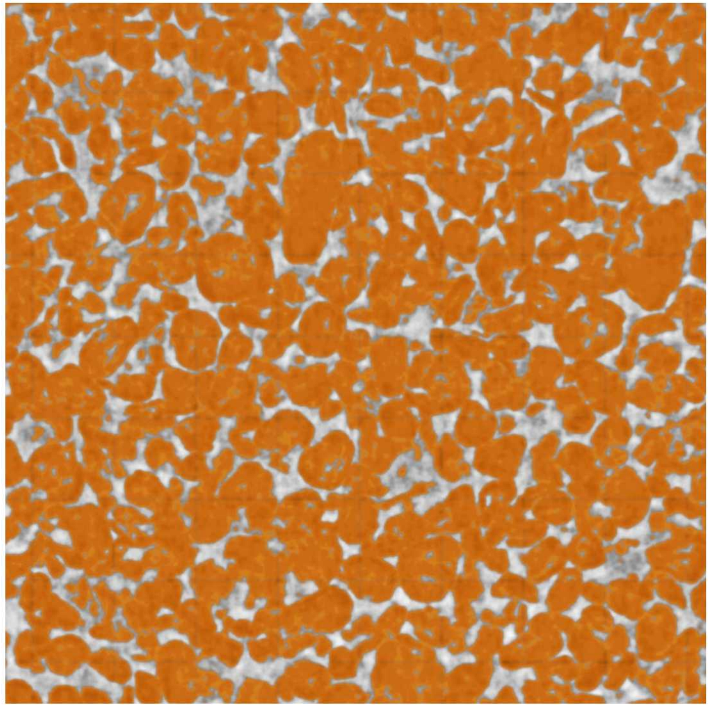

Electrode II_3

Inference results in training domain are as expected

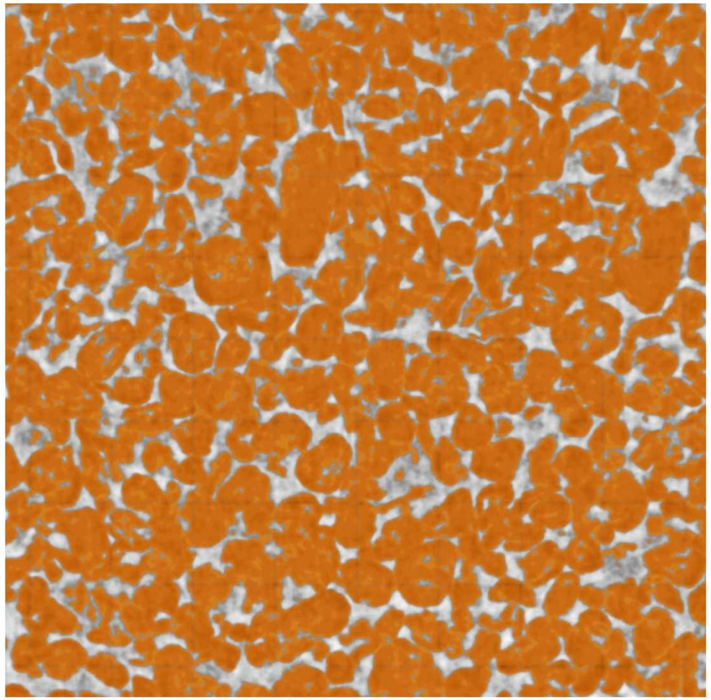
Litarion CT scan slice



Human label



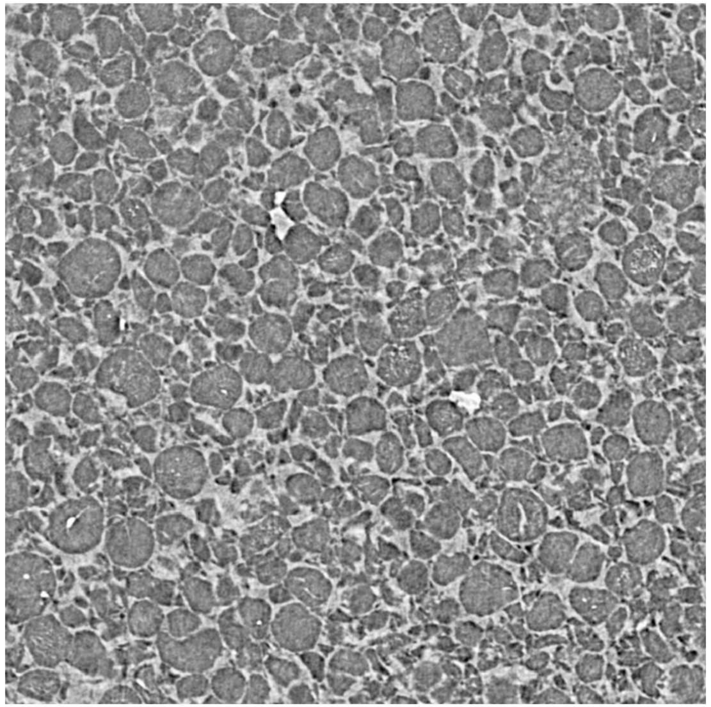
ML prediction



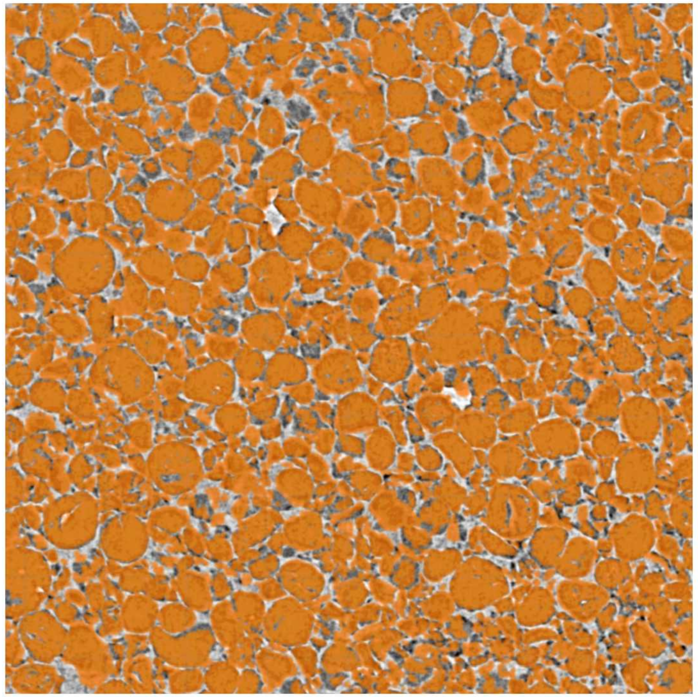
ML segmentation is 96.6% accurate to the human label

Inference results outside the training domain are qualitatively better than accuracy measurements indicate

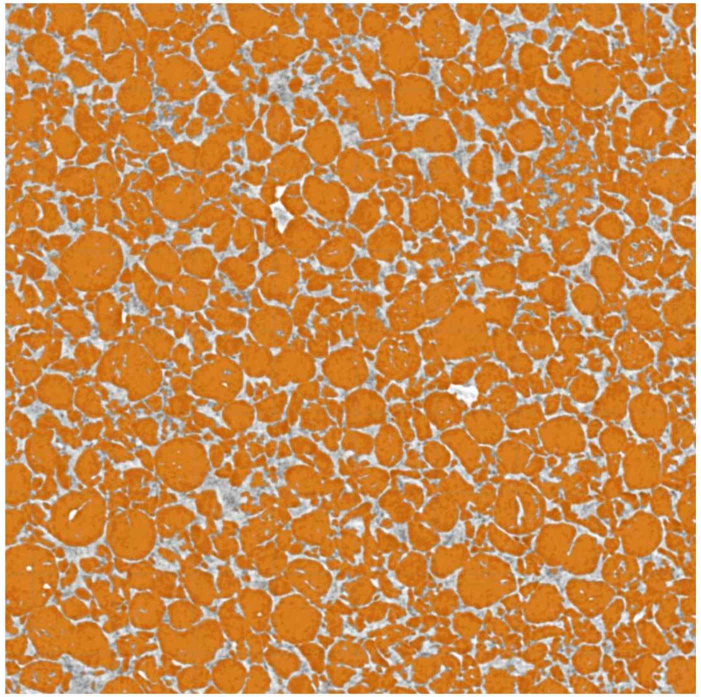
Electrode II_3 CT scan slice



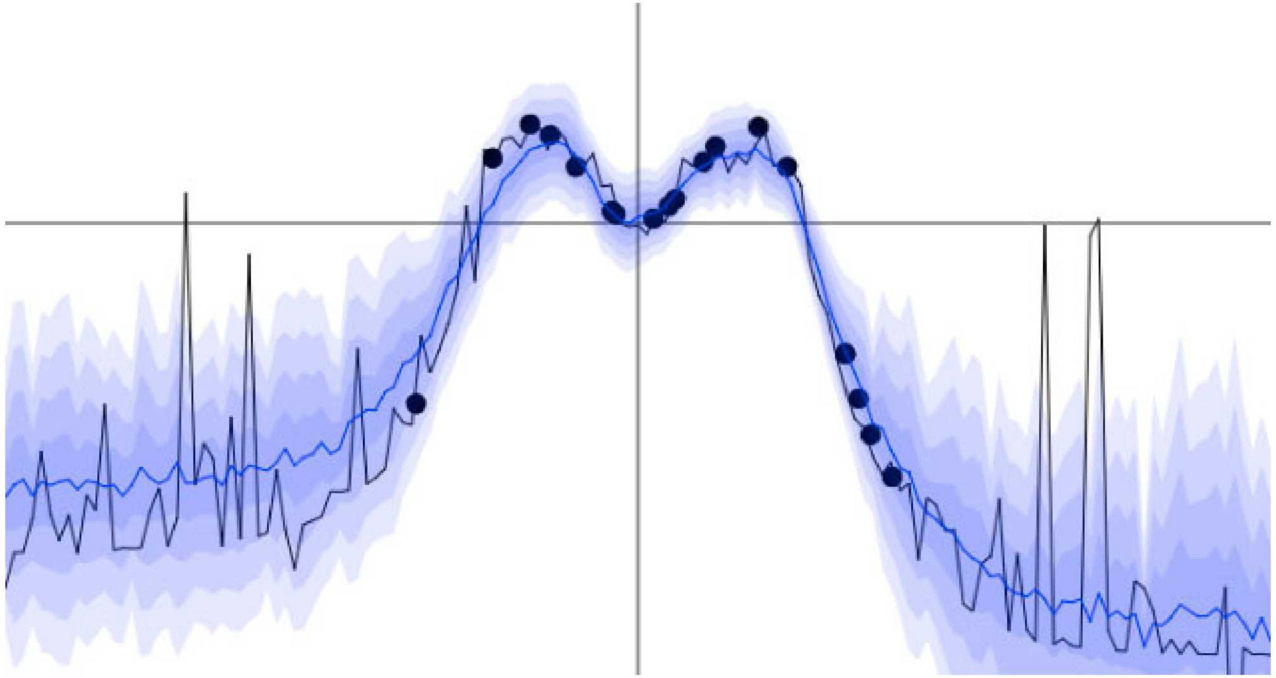
Human label



ML prediction



ML segmentation is 69.9% accurate to the human label…but looks qualitatively better http://www.cs.ox.ac.uk/people/yarin.gal/website/blog_3d801aa532c1ce.html



Using dropout to estimate segmentation confidence

UNCERTAINTY QUANTIFICATION

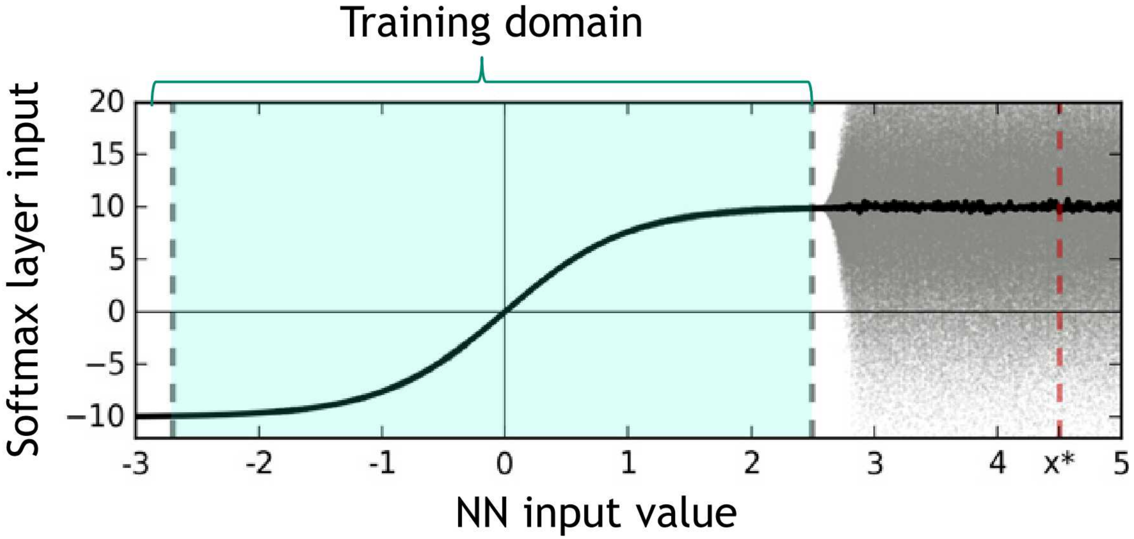
Credible Uncertainty Quantification (UQ) is an open question

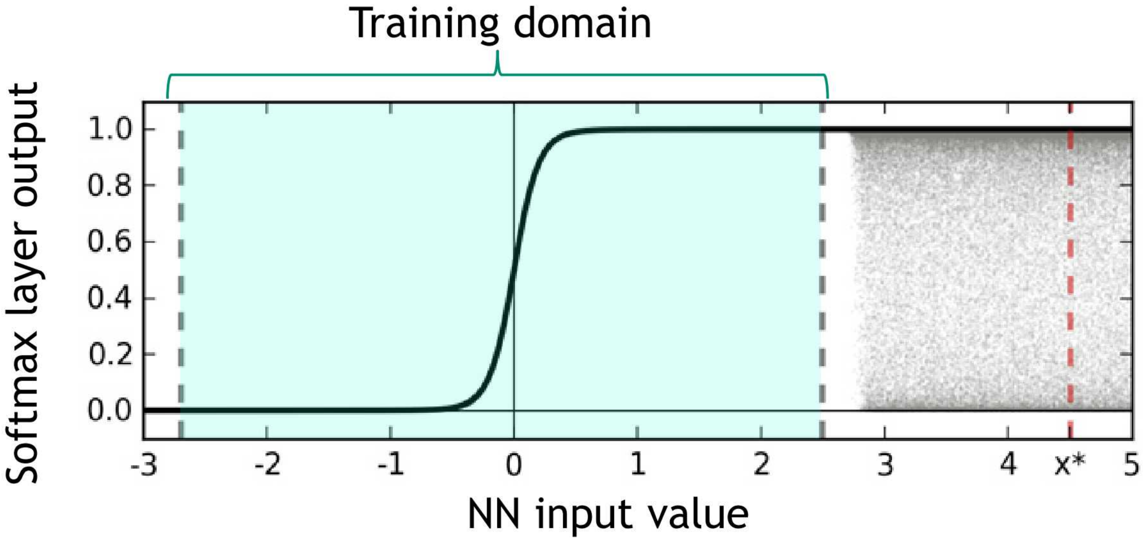
- Uncertainty quantification is critical when segmentations are used in high-consequence scenarios
- But deep learning does not provide UQ estimates – they are not statistical models!
- How can we cast a neural network as a statistical model to obtain theoretically justified UQ?

Monte Carlo Dropout Network (MCDN) (Gal 2016)	Bayesian Neural Network (BNN) (Blundell 2015)
Insert dropout layers into neural network to approximate Deep Gaussian Process	Learn distributions instead of pointwise estimates for each weight in the neural network
Learns uncertainty in the **output space**	Learns uncertainty in the **weight space**
Has been implemented for 3D domains (Liu)	Thought to be infeasible in 3D (Gal 2016)
Statistical soundness questioned (Osband 2016)	Statistically sound (Graves 2011)
Easy to implement, no parameter increase	Very hard to implement, 2x parameter increase

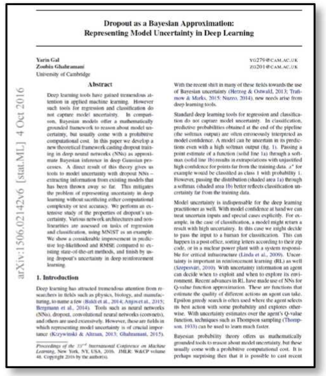
Standard neural network models do not provide error bars

Softmax output is not enough to determine model uncertainty





- Solid black line is the result when dropout is not used
- Grey area shows inputs and outputs when dropout is used during inference
- We can use dropout at inference time to approximate uncertainty


Dropout as a Bayesian Approximation:
Representing Model Uncertainty in Deep Learning

Yarin Gal
Zoubin Ghahramani
University of Cambridge

Abstract

"Dropout as a Bayesian Approximation" – Gal et. al., 2016

https://www.cs.ox.ac.uk/people/yarin.gal/website/blog_3d801aa532c1ce.html

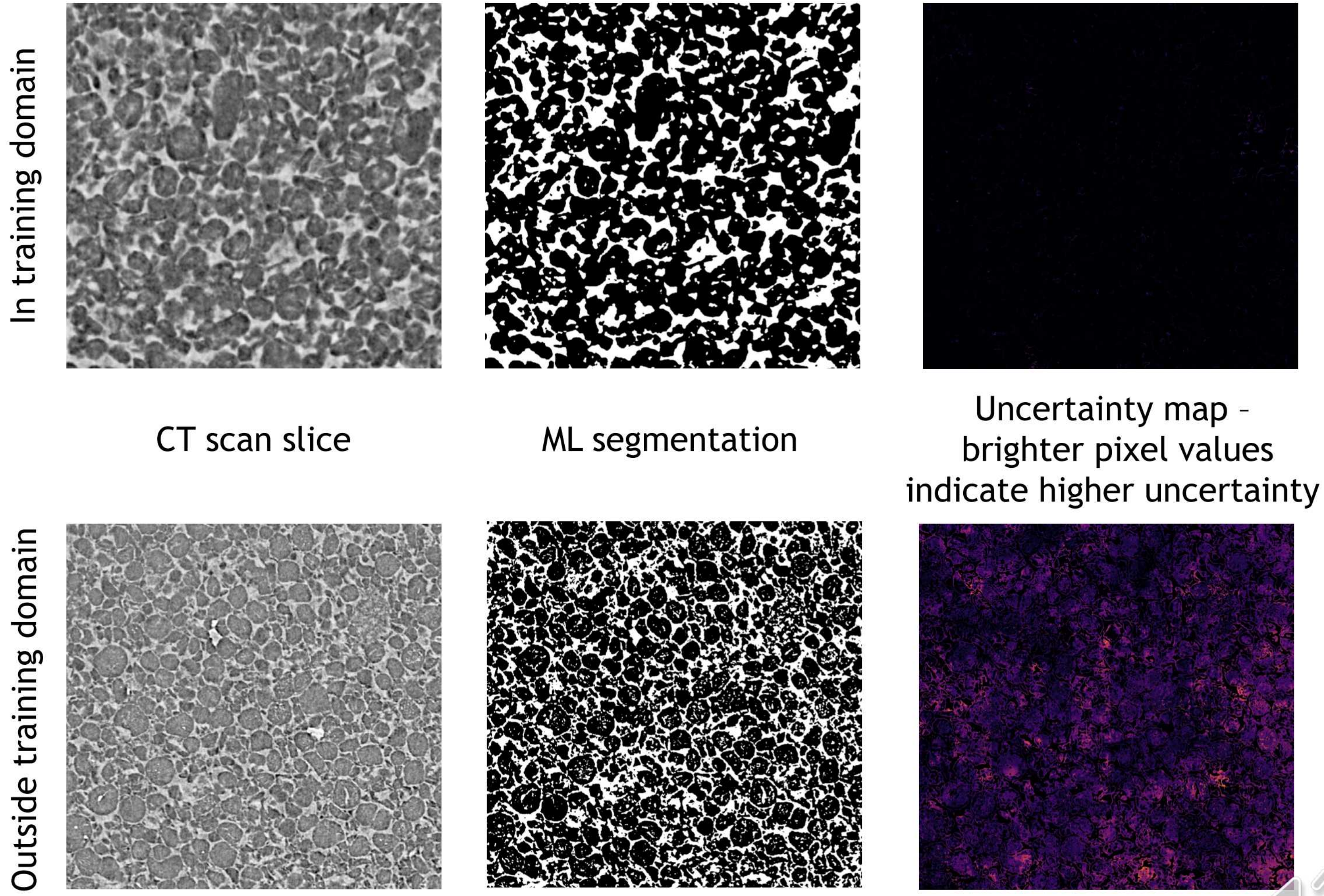
Uncertainty quantification allows us to add error bars to our deep learning models

Using a deep learning technique called dropout, we sample segmentation results from the trained model.

For each pixel, we quantify the level of uncertainty in our model, and we can make decisions about the model's credibility on a particular task.

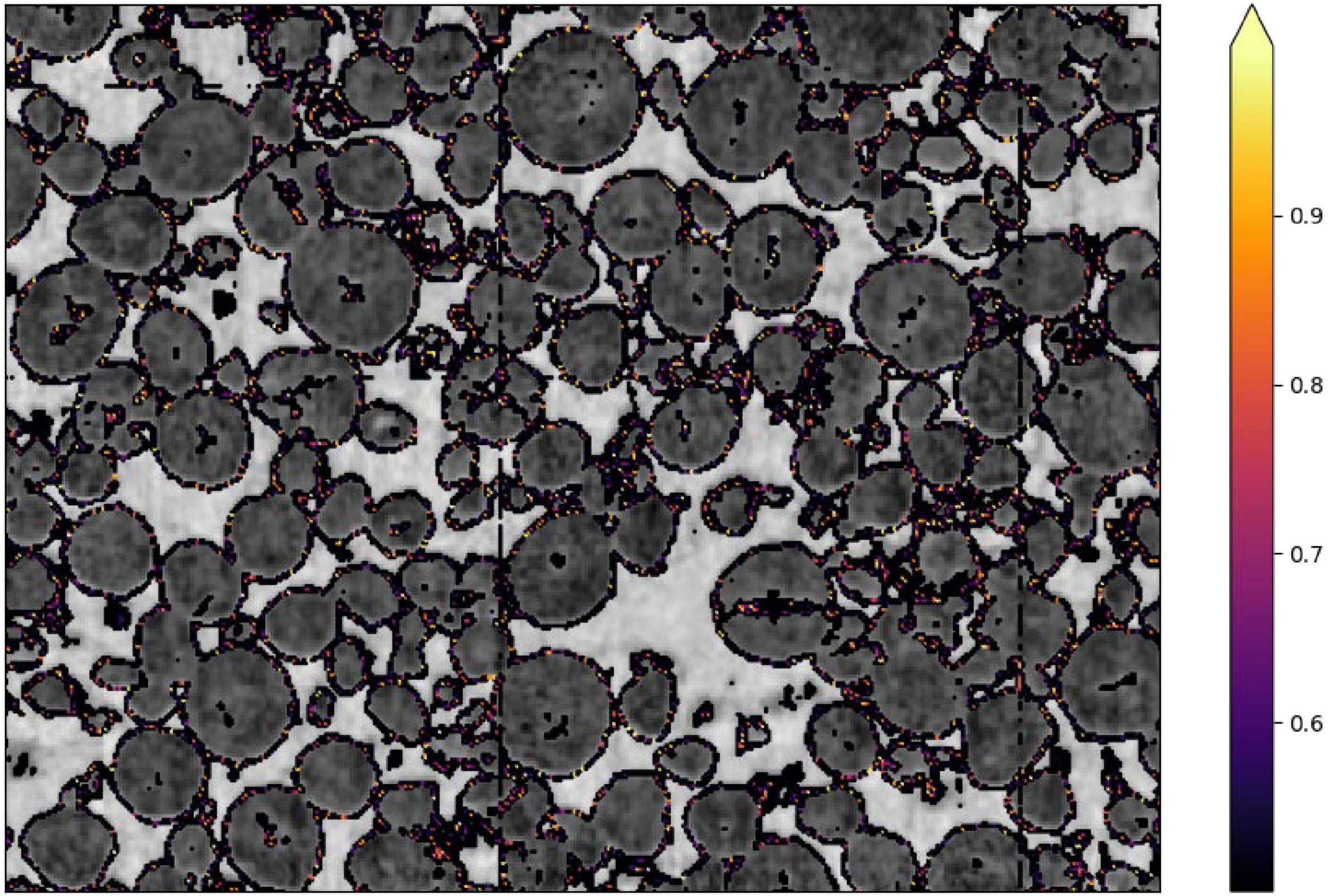
The trained model has less confidence in segmentations of inputs that fall outside of the training distribution.

Work in progress: Use uncertainty maps to bound variance in geometries of as-built parts for use in simulations



Gal, Yarin, and Zoubin Ghahramani. "Dropout as a Bayesian approximation: Representing model uncertainty in deep learning." international conference on machine learning. 2016.

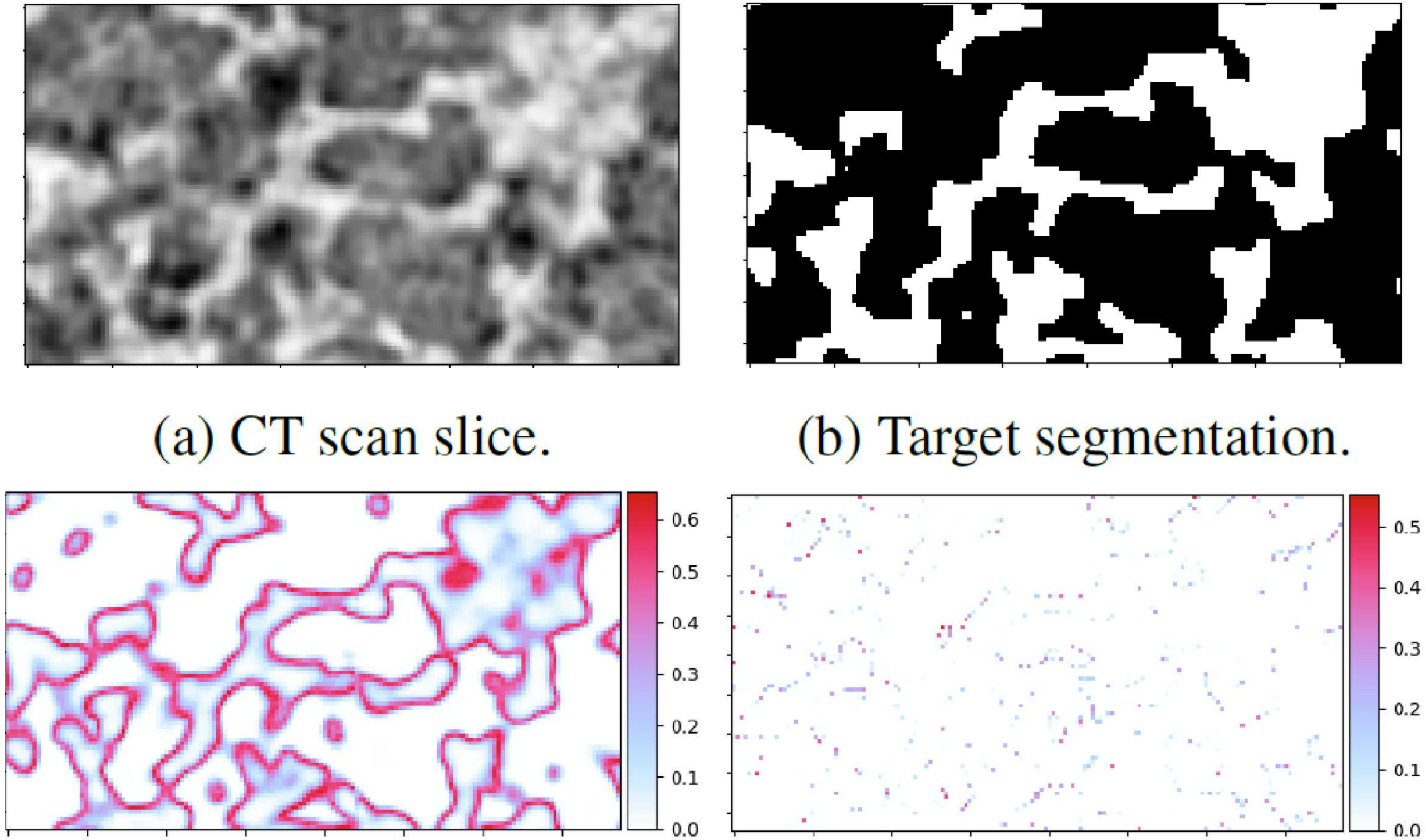
How can we understand geometric uncertainties in deep learning segmentations?



CT scan of battery with output of softmax layer overlaid

We present first-ever 3D Bayesian CNN (BCNN) for UQ

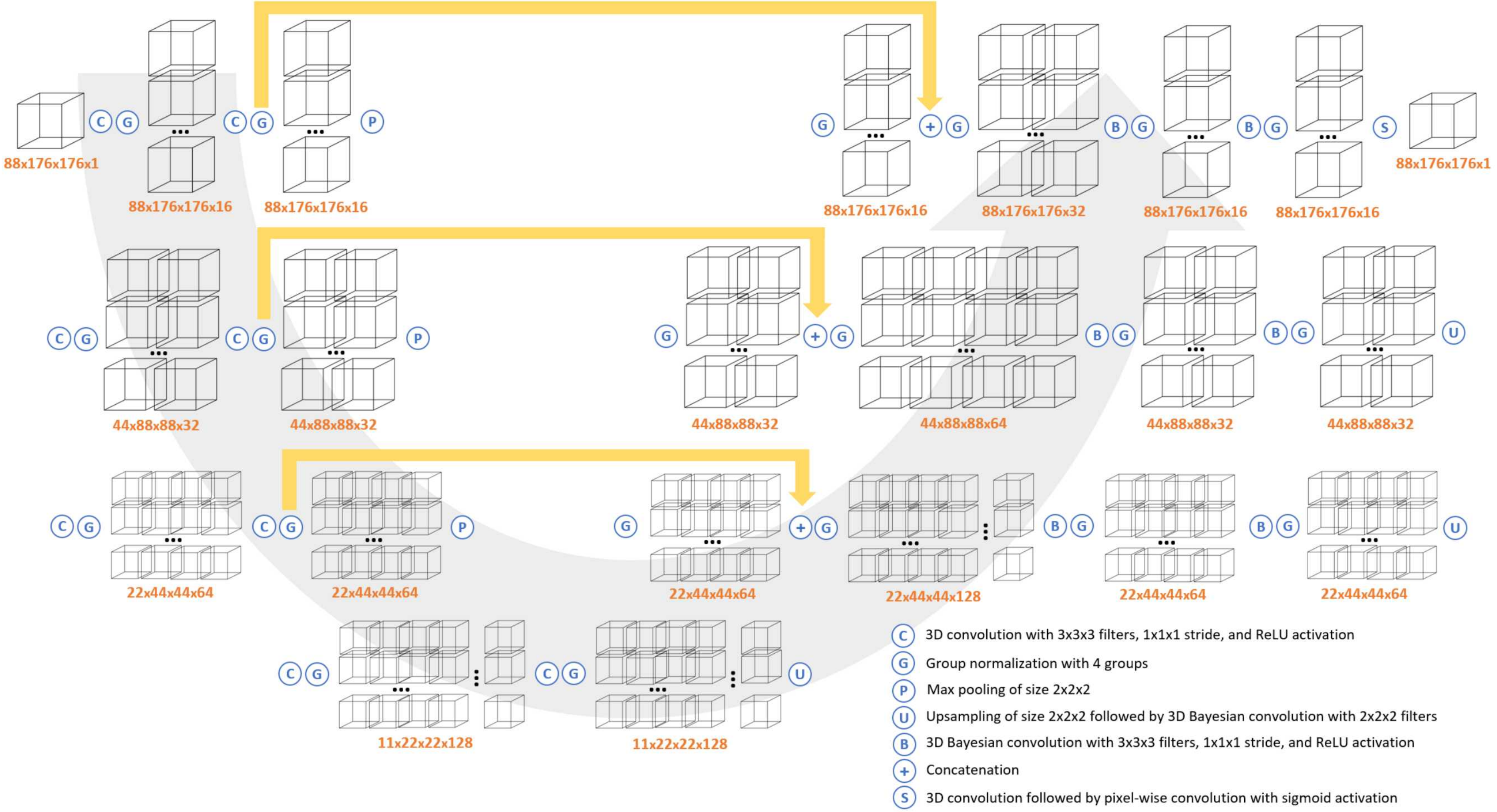
- Refutes theory that BCNNs are infeasible in 3D due to poor convergence
  - Used many recent advances in field to produce a reliable deep learning system
- Tested on CT scan datasets of lithium-ion battery electrodes & laser-welded metals
  - Can extend to multiple datasets of mission-critical materials
- Outperforms MCDN in quality of uncertainty maps
  - Also outperforms MCDN in most recent uncertainty validation metrics



(a) CT scan slice. (b) Target segmentation. (c) BCNN uncertainty (ours). (d) MCDN uncertainty.

3D BCNN Architecture

- Draws from original 3D segmentation architecture called V-Net (Milletari 2016)
- Uses Bayesian layers with standard normal prior in decoder half (right)
- Total of 1.92 million parameters and 49 layers



Variational learning allows distributions on weights

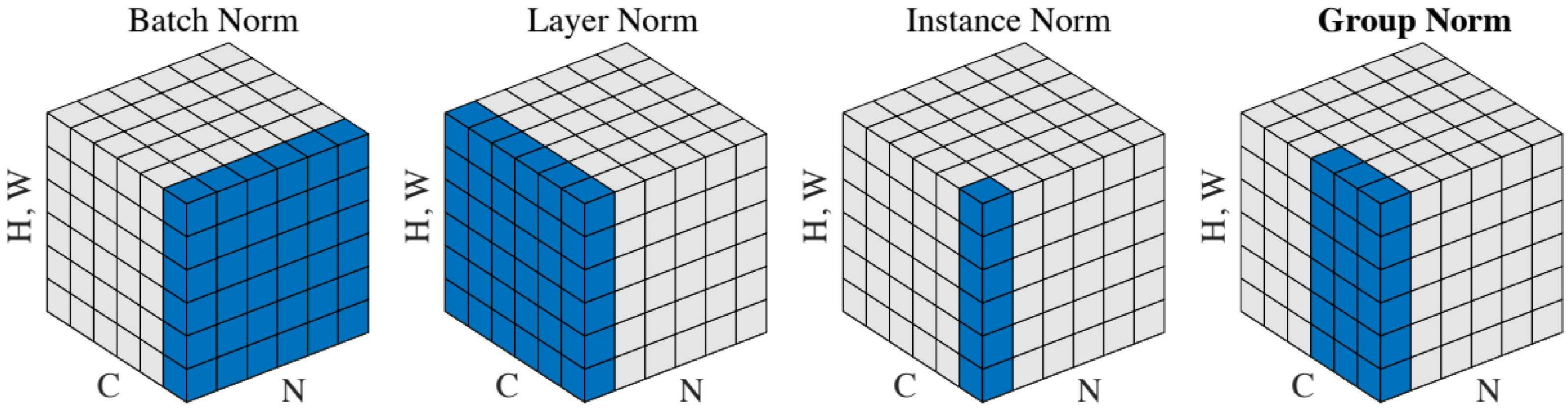
- Problem: intractable to calculate the exact posterior distribution of the weights
- Solution: perform variational learning (Graves 2011) with Bayes by Backprop (Blundell 2015)

$$\mathcal{F}(\mathcal{D}, \theta) = \mathrm{KL}[q(\boldsymbol{w}|\theta) \parallel P(\boldsymbol{w})] - \mathbb{E}_{q(\boldsymbol{w}|\theta)}[\log P(\mathcal{D}|\boldsymbol{w})].$$

- Minimize variational free energy loss above
  - Left side: simplicity of prior (KL divergence)
  - Right side: complexity of dataset (Negative log-likelihood)
  - Annealing of KL term necessary for reliable convergence
- Bayes by Backprop: integrate distribution updates into backpropagation
  - Minimizes extra training computation

Group normalization helps train quickly and reliably

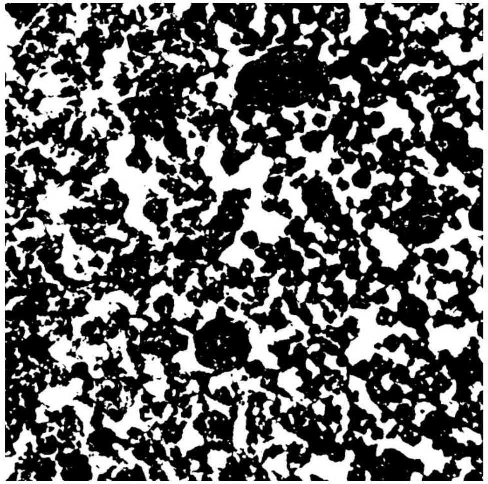
- Problem: large volumes only allow for small mini-batches
  - Makes batch normalization fail, but batch normalization is necessary for reliable convergence
- Solution: use group normalization (Wu and He 2018)
  - Calculates normalization independent of batch size
  - Outperforms batch norm for size 8 or lower



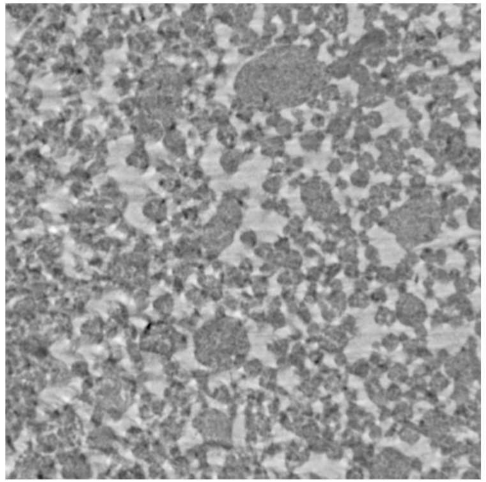
Group normalization represented in tensor format (Wu and He 2018)

Test results on lithium-ion battery electrodes

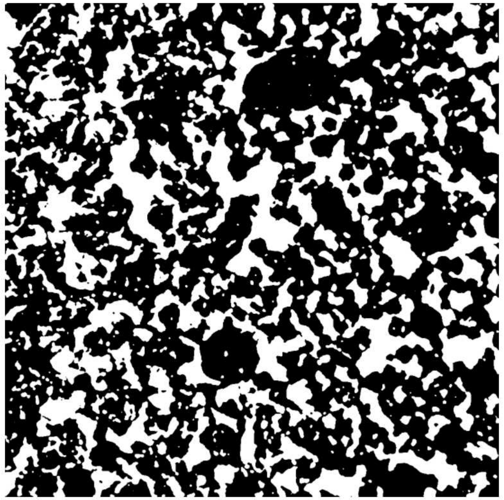
- Accuracy roughly equal to MCDN
- BCNN UQ is usable and interpretable, outperforming MCDN UQ



Target segmentation ←

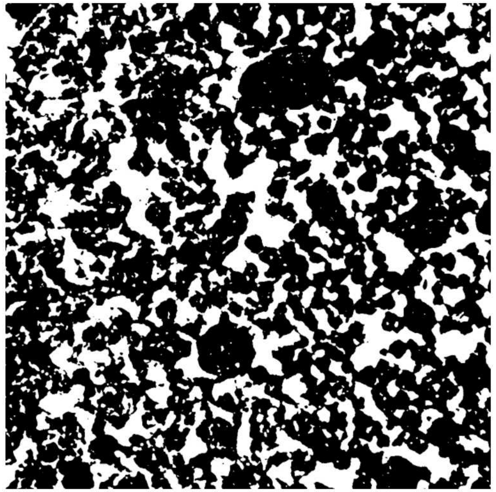


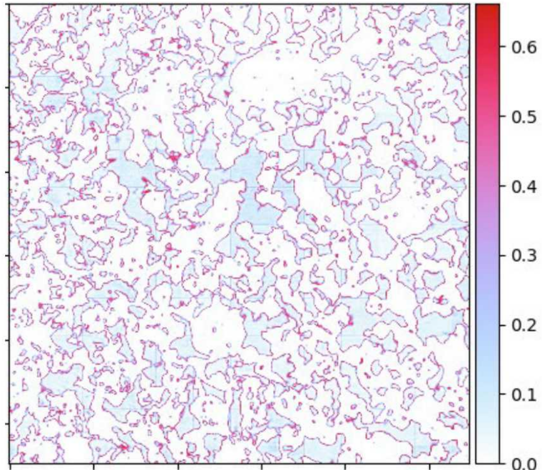
CT scan slice ←



BCNN segmentation ←

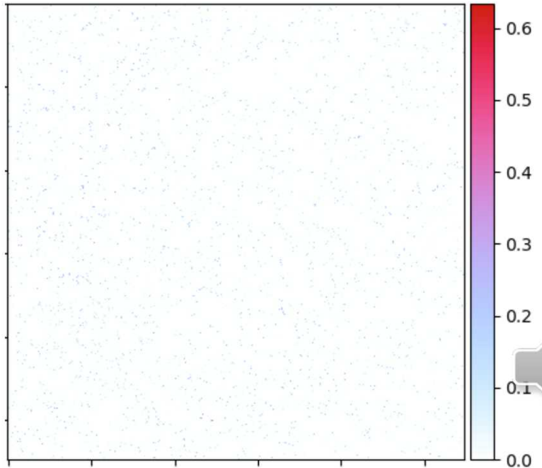
MCDN segmentation →





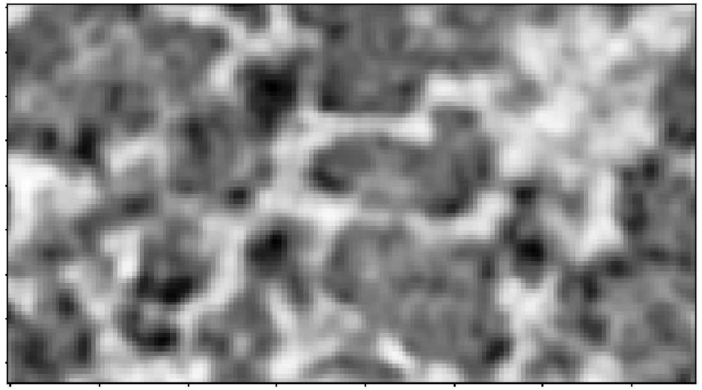
BCNN uncertainty ←

MCDN uncertainty →

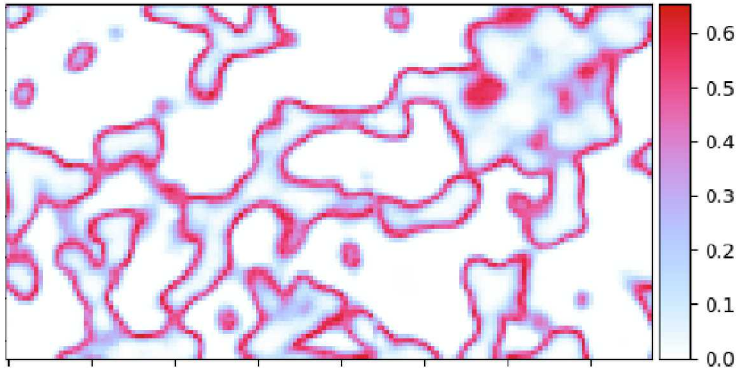


Test results on lithium-ion battery electrodes

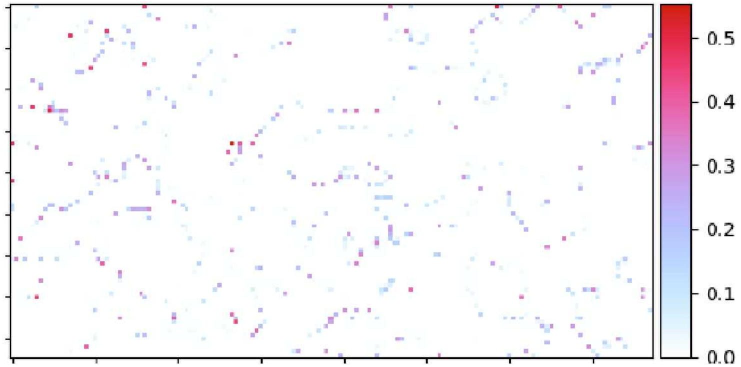
- BCNN uncertainty captures continuity and visual gradients
- BCNN uncertainty more often identifies areas of high visual uncertainty in the CT scan


(a) CT scan slice.


(b) Target segmentation.



(c) BCNN uncertainty (ours).



(d) MCDN uncertainty.

Test results on laser-welded metal joints

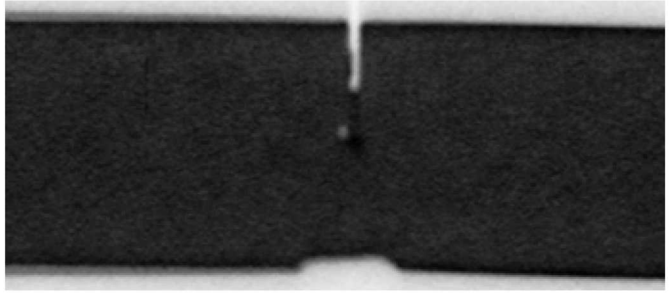
- Accuracy roughly equal to MCDN
- BCNN UQ is usable and interpretable, outperforming MCDN UQ
  - Captures visual gradients and continuity

Target segmentation

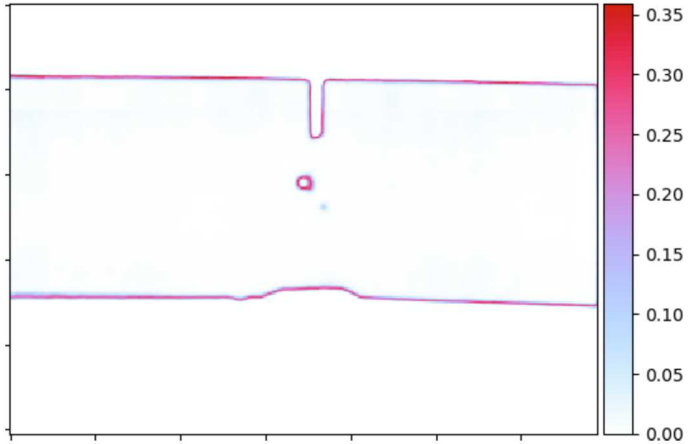
BCNN segmentation

MCDN segmentation

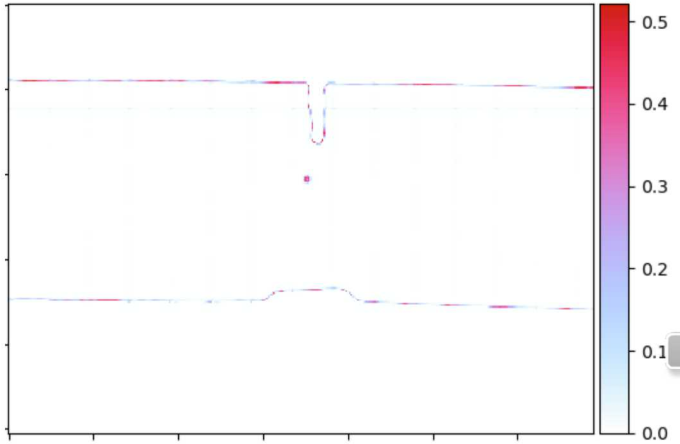
CT scan slice



BCNN uncertainty



MCDN uncertainty



Acknowledgements

- Kevin Potter, Matthew Smith, Emily Donahue, Charlie Snider
- John Korbin
- David Peterson
- Scott Roberts
- Kyle Karlson
- Lincoln Collins
- Chance Norris
- Tyler LaBonte
- CAMI LDRD team

Questions?

cmarti5@sandia.gov