SAND2022-16023C

What do we need subject matter experts for with plots like these?

Cari Martinez

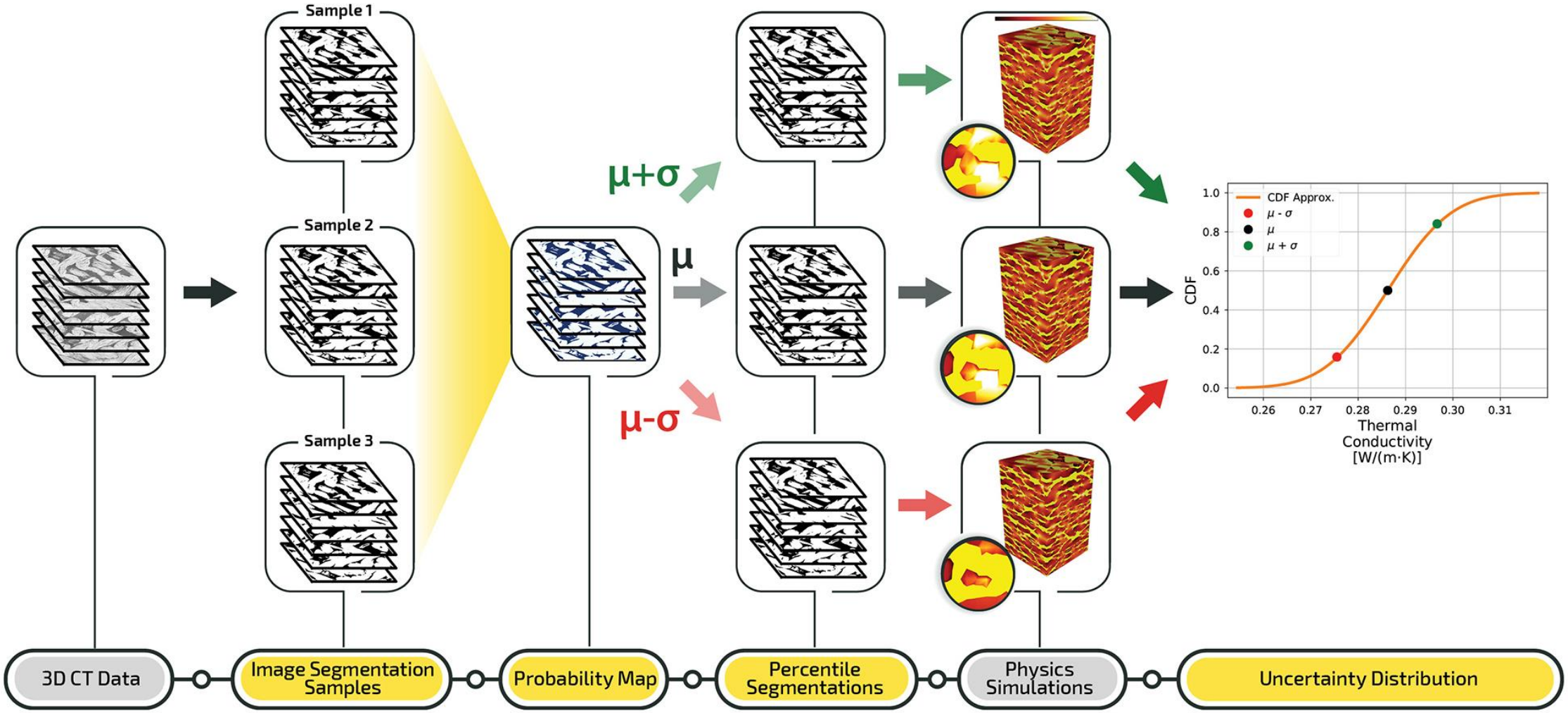
November 16, 2022

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.

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.

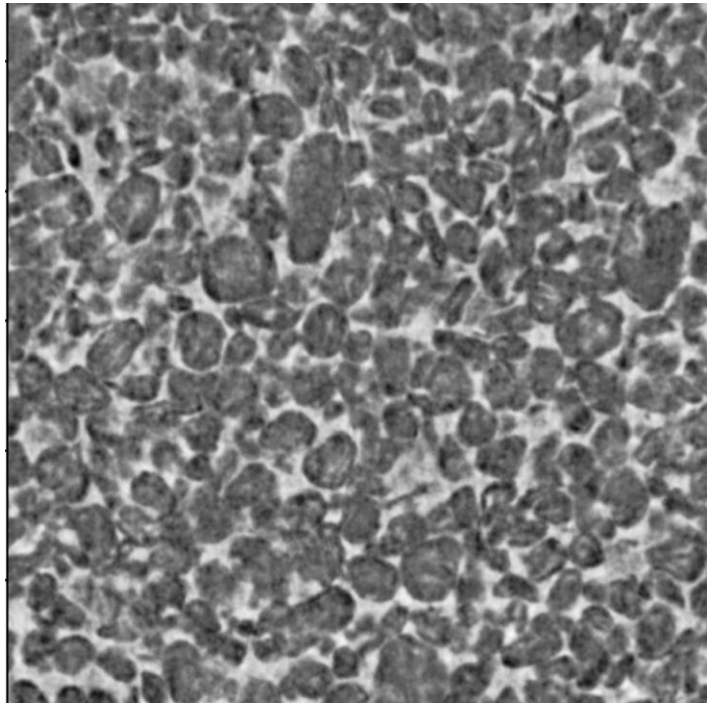
Digital twins require the processing of massive amounts of data.



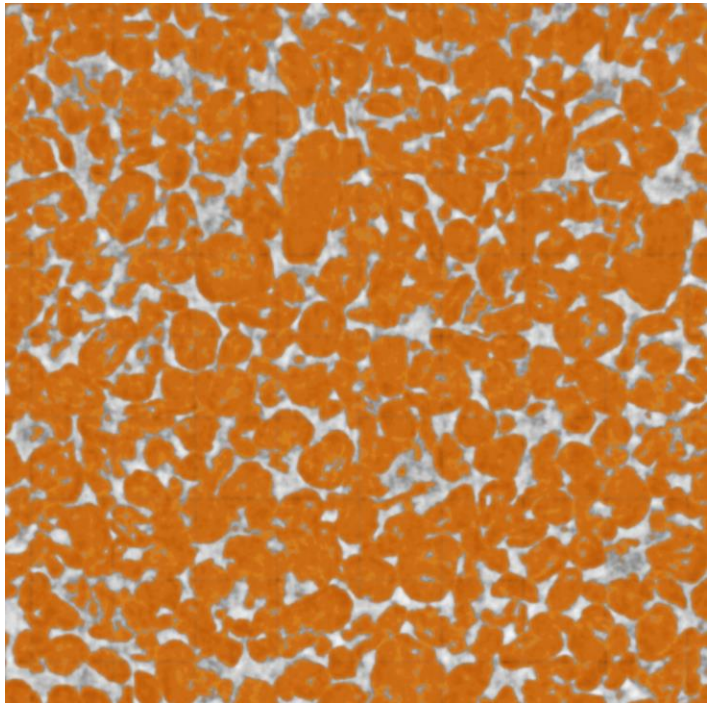
Krygier, M. C., LaBonte, T., Martinez, C., Norris, C., Sharma, K., Collins, L. N., ... & Roberts, S. A. (2021). Quantifying the unknow n impact of segmentation uncertainty on image-based simulations. *Nature communications*, *12*(1), 1-11.

Deep learning models can learn to process data, guided by expert labels.

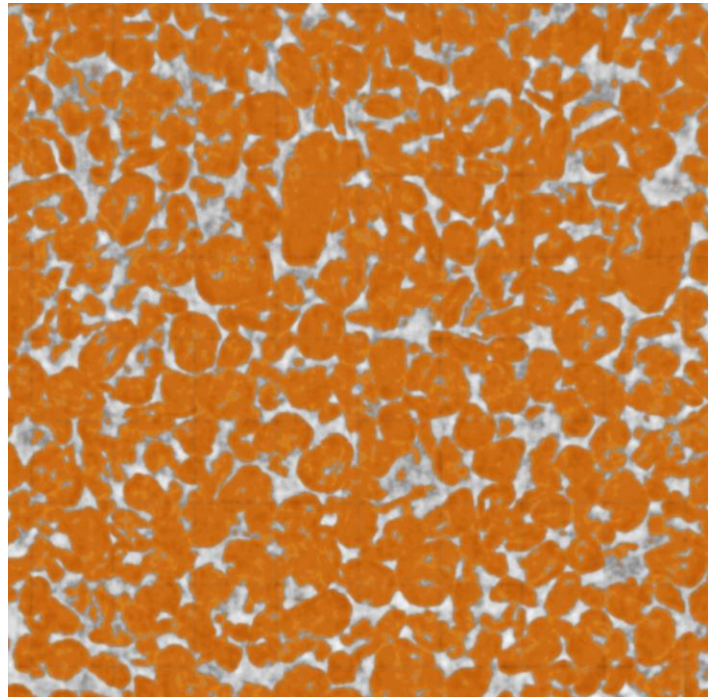
Litarion CT scan slice



Human label



ML prediction

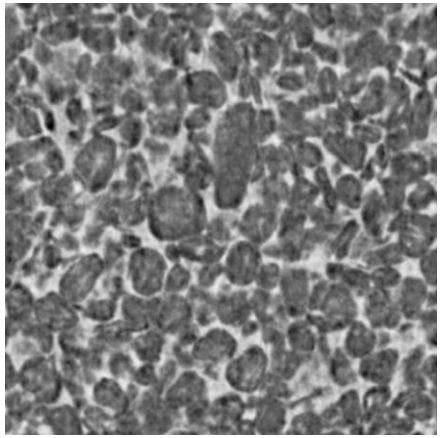


ML segmentation is 96.6% accurate to the human label

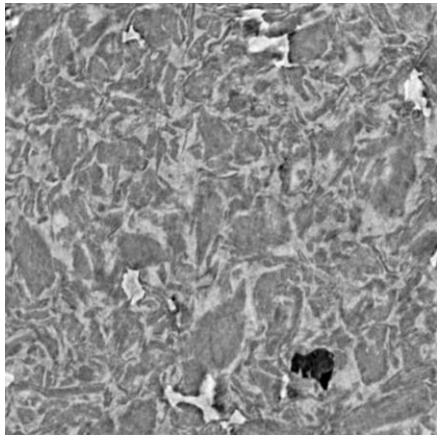
What happens when we ask deep learning models to “extrapolate”?

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

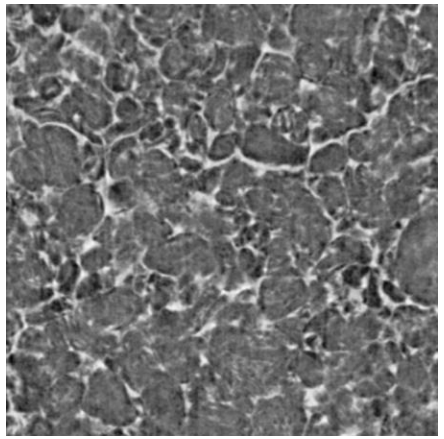


Litarion

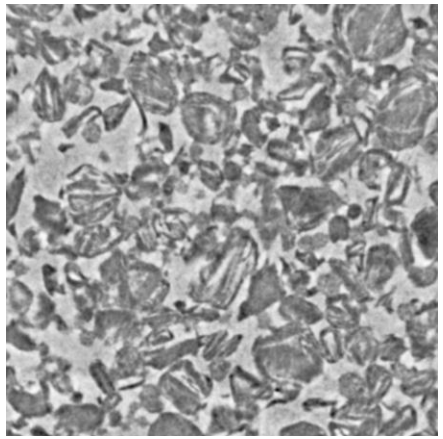


Electrode IV_1

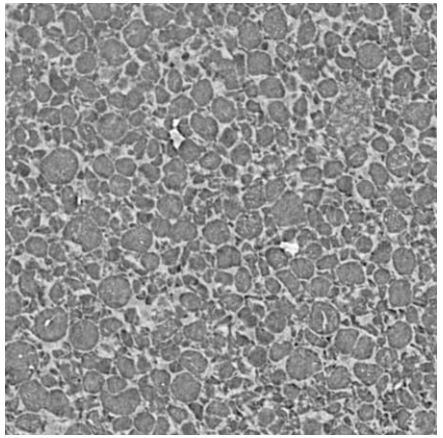
TEST SET



E35



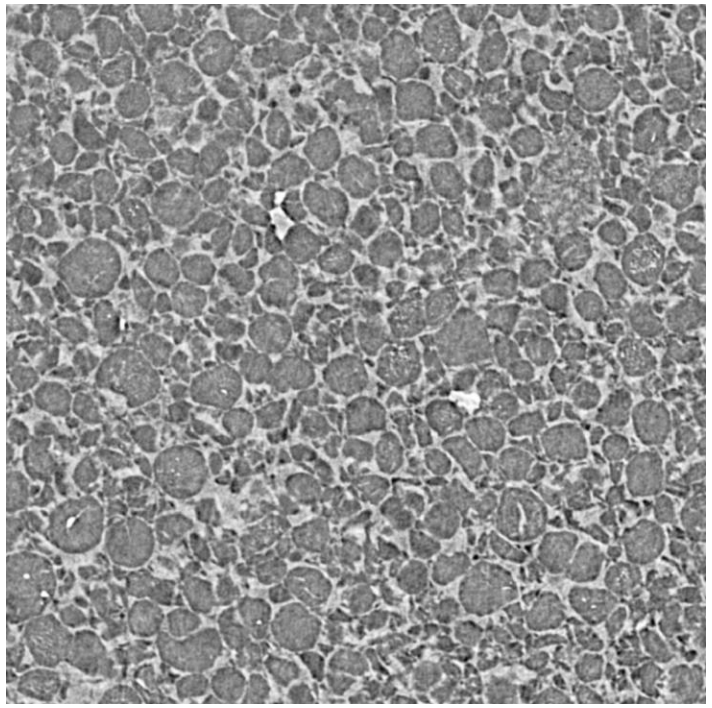
GCA400



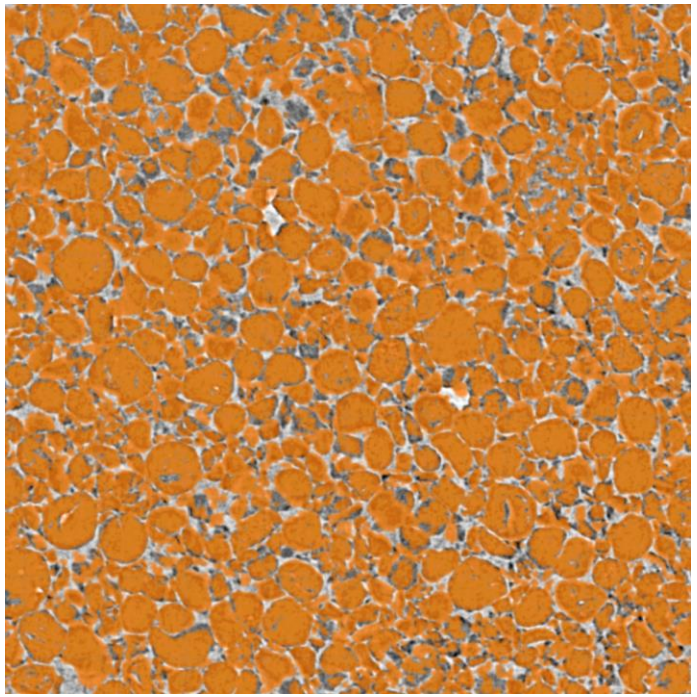
Electrode II_3

Inference results outside the training domain are better than human labels in this case.

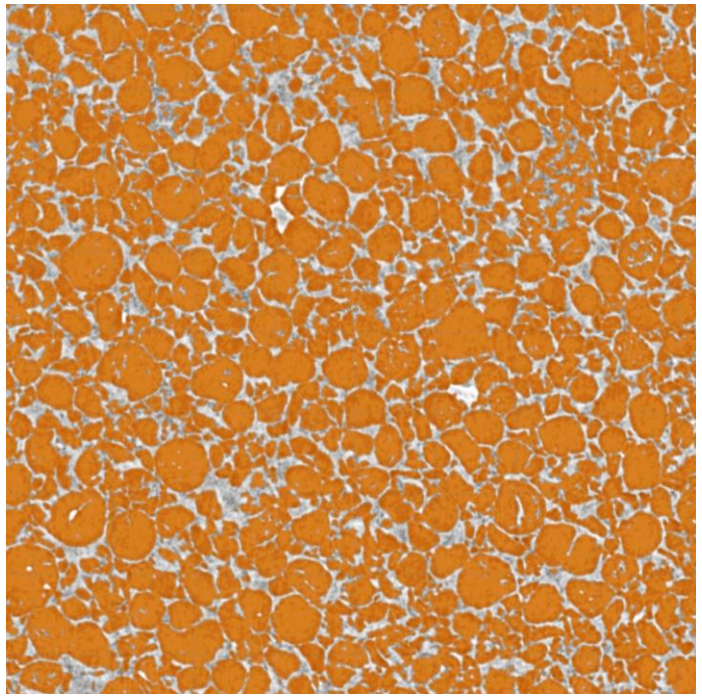
Electrode II_3 CT scan slice



Human label



ML prediction

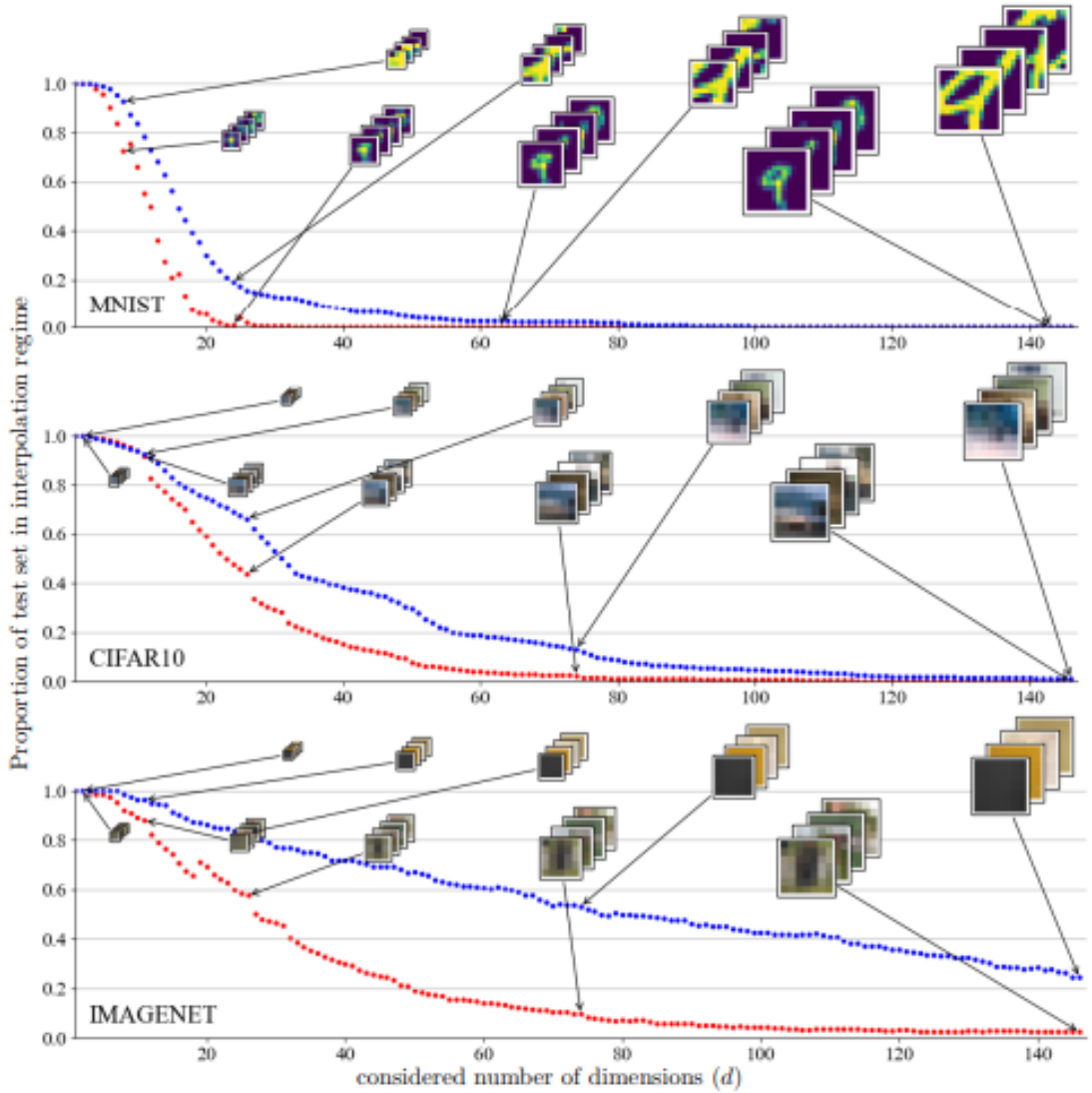


ML segmentation is 69.9% accurate to the human label…but looks qualitatively better

In high-dimensional spaces, do we care about “extrapolation” or generalization?

Compelling preprint from FAIR, NYU [1] observes that in common computer vision datasets, held-out test examples are almost always outside the convex hull of the training data; **we are almost always extrapolating, yet deep learning models are successful.**

Is there a better definition of extrapolation that is more relevant to data-driven scientific modeling in high dimensional spaces?



[1] Balestriero, Randall, Jerome Pesenti, and Yann LeCun. "Learning in high dimension always amounts to extrapolation." *arXiv preprint arXiv:2110.09485* (2021).

Questions?

cmarti5@sandia.gov