Bayesian Neural Networks For Uncertainty Estimation In Particle Accelerator Applications

SLAC NATIONAL ACCELERATOR LABORATORY

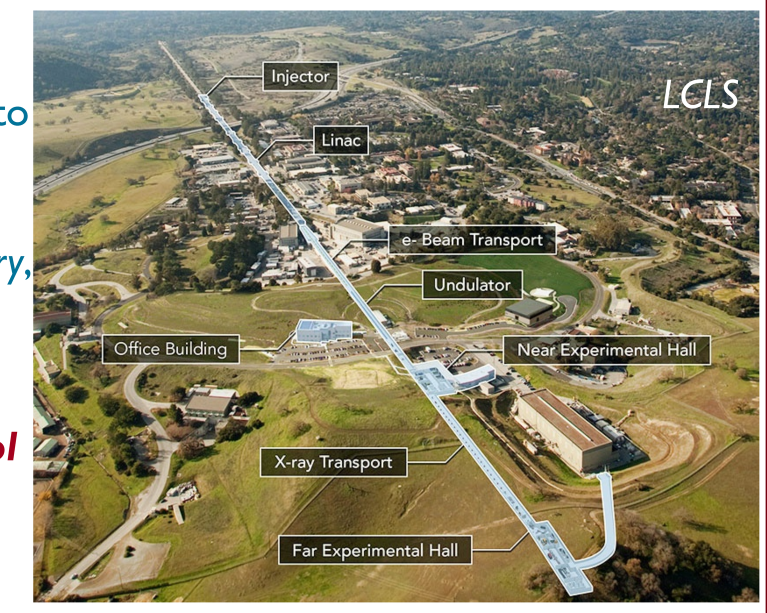
Aashwin Ananda Mishra, Auralee Edelen, Adi Hanuka, Christopher Mayes

SLAC National Accelerator Laboratory

Introduction / Motivation

X-ray Free Electron Laser facilities must provide high-quality custom photon beams to scientific users on-demand. Applications in *cancer therapy, nuclear non-proliferation treaty verification, material discovery,* chemistry and biological insights.

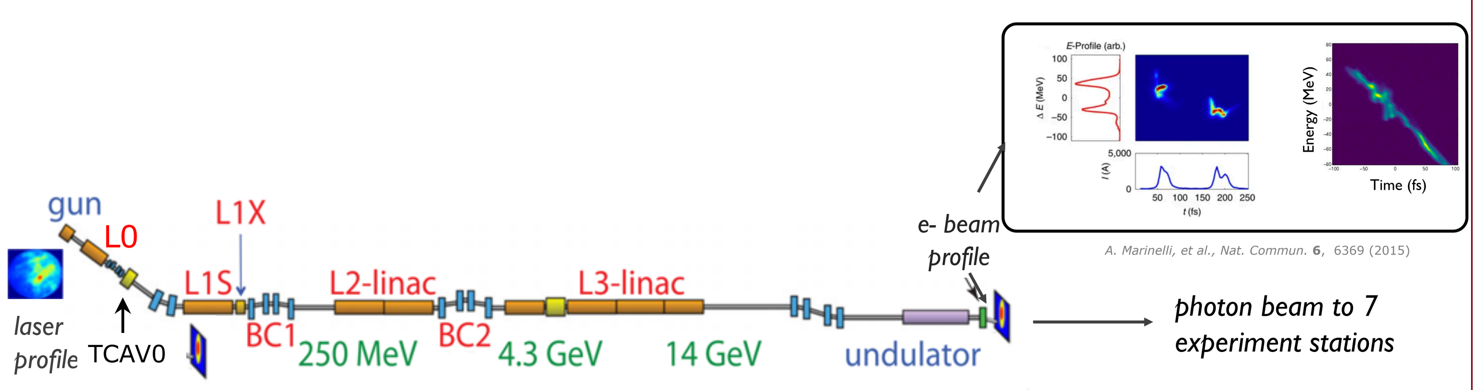
This is a challenging optimization/control problem

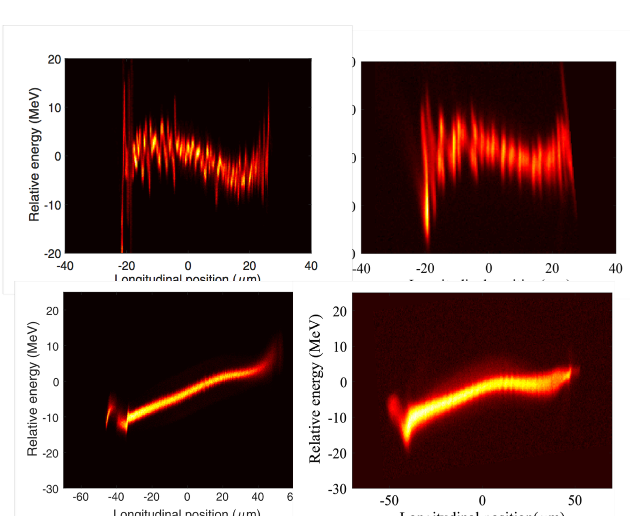


About $30,000 per experiment hour to run the LCLS → efficient tuning matters

nonlinear, non-stationary behavior
high-dimensional parameter space
limited diagnostics / incomplete knowledge
intermittent anomalies / failures

Common to most particle accelerator systems





Start-to-end LCLS simulation
***10 hours** on thousands of cores at NERSC*

Accelerator simulations can be very slow to execute and often don't match the machine well

Fast, accurate system models would be useful for:

online modeling (e.g. extra diagnostic info / model-based control)
prototyping new control algorithms (e.g. reinforcement learning)
experiment planning + start-to-end optimization of new setups

→ How can we use ML to improve the speed and accuracy of our accelerator models, and use these to aid optimization / control ?

- *Particle Accelerators represent **high-regret and safety-critical** systems. Overconfident predictions from deep neural networks, compounded by their lack of intepretability, can have undesirable consequences.*

The first fatality in automated driving vehicles occurred when the AI-agent couldn't differentiate the sky from a white trailer.

An individual was charged with traffic violations when the Chinese AI-detection system mistook her image on a hoarding

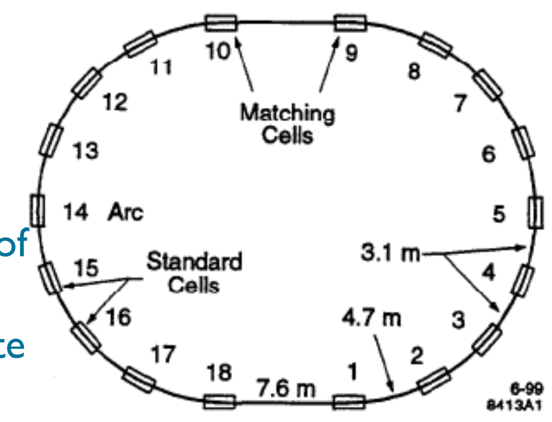
- *In addition to point predictions, we need **reliable measures of predictive uncertainty** from the data driven model.*

***Epistemic**: inherent complexity of problems & lack of all import parameters measured.*
***Aleatoric**: limited instrumentation, sensitivity and noise, compounding errors and trips.*
***Out of Distribution**: Models often used to explore new parameter ranges.*

- *At present, bootstrapped ensembles of neural networks are used to estimate uncertainty in predictions. However, this may underpredict the uncertainty intervals.*
- *In this investigation, we utilize Bayesian Neural Networks as a potential alternative, providing an amalgam of the the predictive capability of neural networks with the uncertainty quantification inherent to the Bayesian formalism.*
- *We utilize the Bayes By Backprop algorithm for inference, and exhibit results for scalar and image inputs and outputs.*
- *The Bayesian Neural Network predictions are compared and contrasted against deterministic neural networks and bootstrapped ensembles, for predictive accuracy (mean absolute error), Prediction Interval Coverage Probability (PICP), and Mean Prediction Interval Width (MPIW at 90% PICP).*

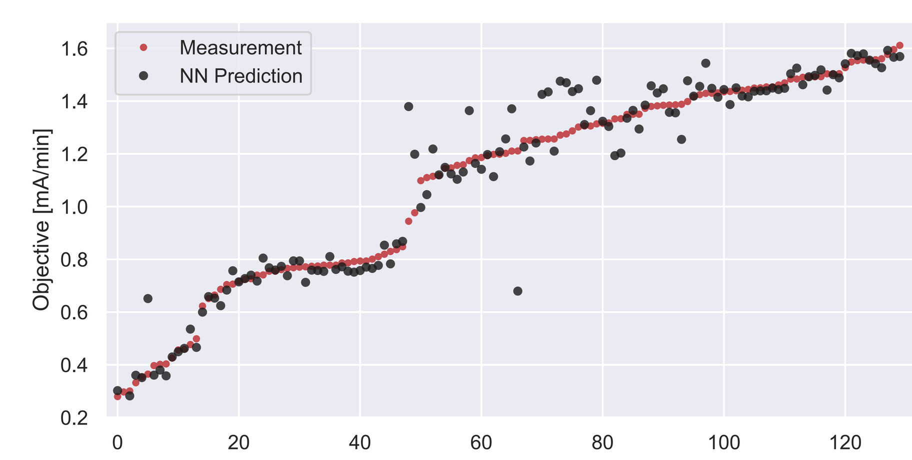
Case I: Spear3 Electron Storage Ring

Spear3 is a 3-GeV, high brightness electron storage ring. 13 skew quadrupoles are adjusted to minimize beam loss and maximize emittance. Owing to the sensitivity of parameters, we require reliable uncertainty estimates from surrogate models.

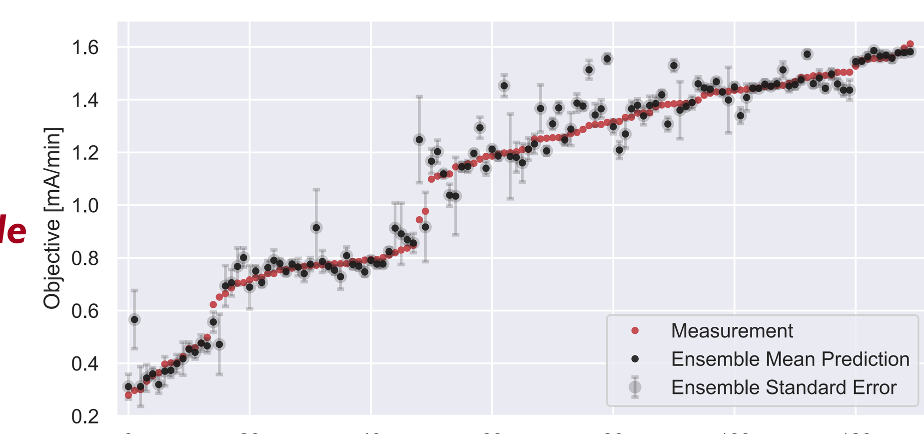


*A. **Experimental Measurements**: We commence with a small dataset of 650 experimental samples, and test on 120 samples.*

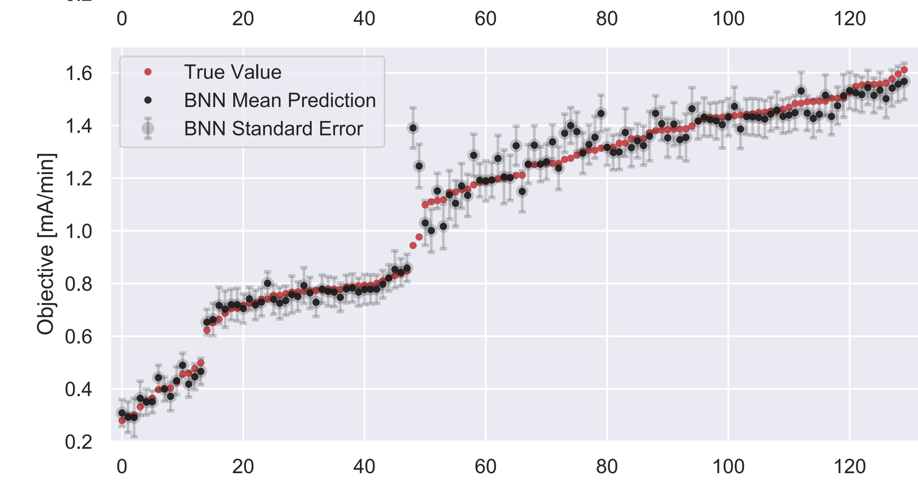
DNN



Ensemble



BNN

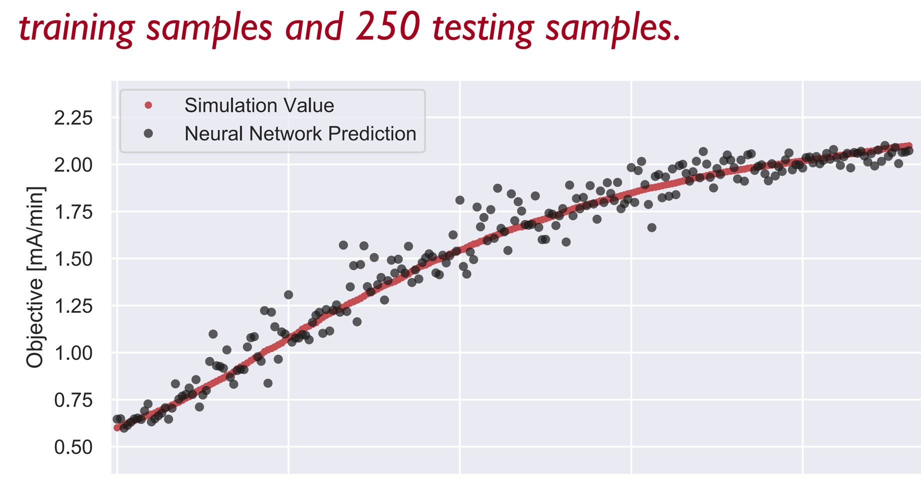


BNNs provide better accuracy, more reliable prediction interval coverage that are more efficient, as opposed to deterministic methods.

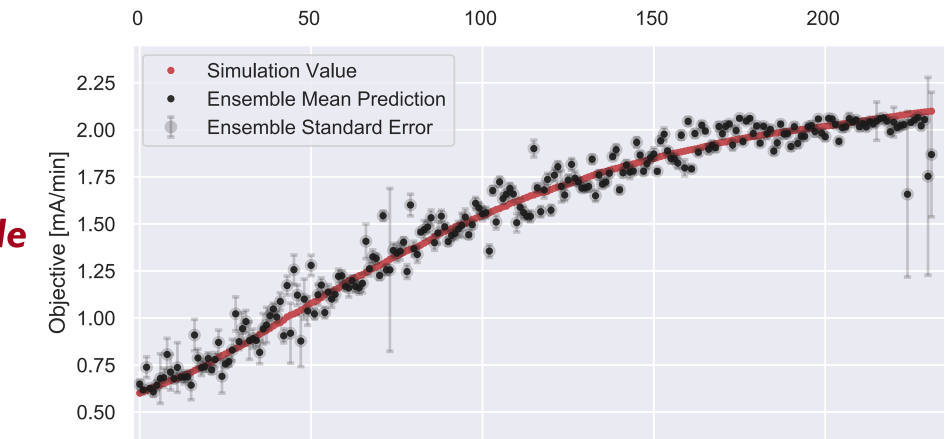
	MAE	PICP	MPIW (@ 90% PICP)
DNN	0.06	--	--
Ensemble	0.043	0.52	0.24
BNN	*0.032*	*0.89*	*0.11*

*B. **Simulation Data**: We repeat for a simulation dataset with 4000 training samples and 250 testing samples.*

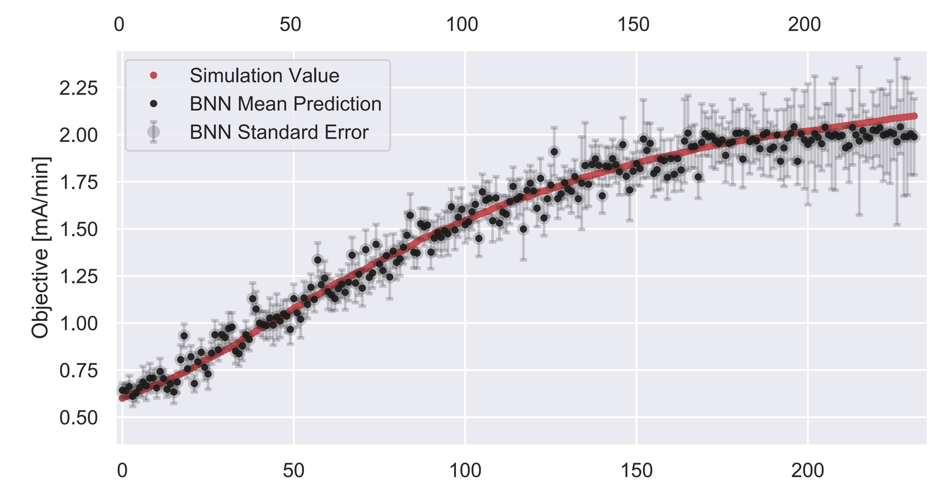
DNN



Ensemble



BNN

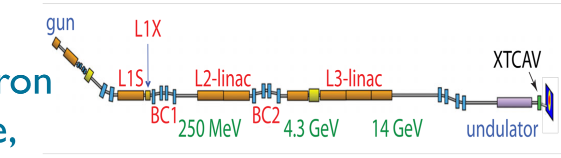


Intervals from bootstrapped ensembles are misleading. BNNs provide better accuracy, more reliable prediction interval coverage that are more efficient.

	MAE	PICP	MPIW (@ 90% PICP)
DNN	0.07	--	--
Ensemble	0.065	0.46	0.36
BNN	*0.055*	*0.87*	*0.19*

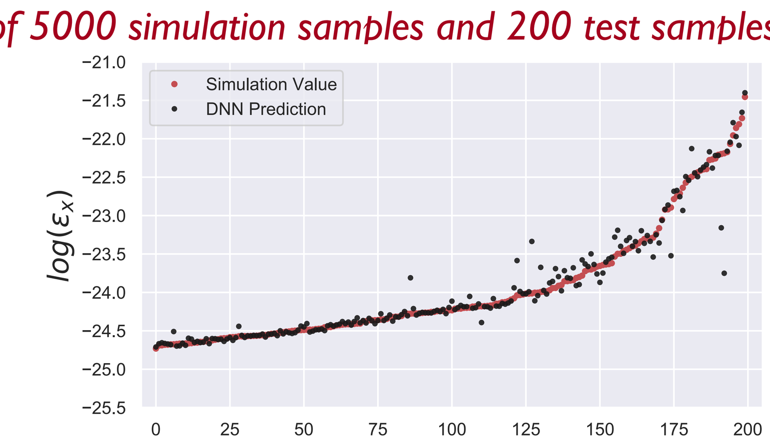
Case II: LCLS Linac

The Linac Coherent Light Source (LCLS) is a Free Electron Laser (FEL) based light source, Providing customized photon beams for scientific experiments. The FEL process is extremely sensitive to small changes in accelerator settings, and the impact of external sources like coherent synchotron radiation. ML models are hamstrung as sampling uniformly from feature space to generate datasets does not sample the target space uniformly, leading to out of distribution effects.

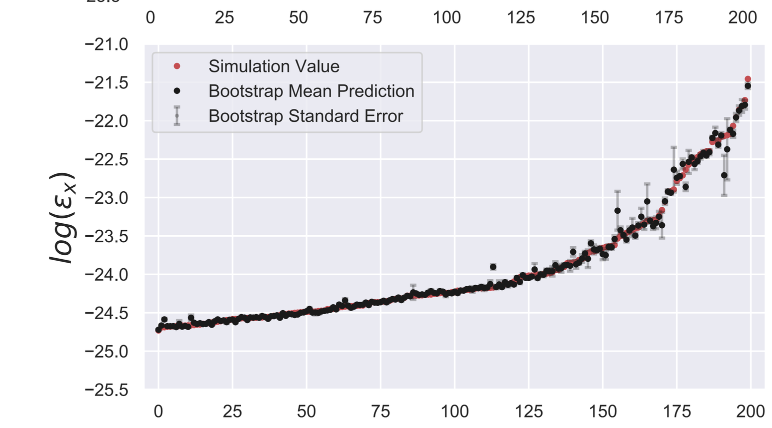


training dataset of 5000 simulation samples and 200 test samples.

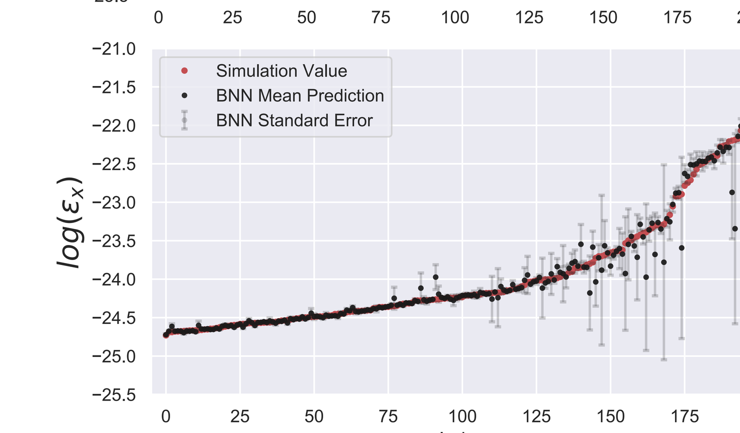
DNN



Ensemble



BNN

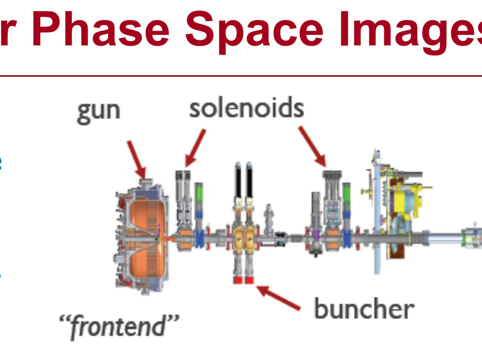


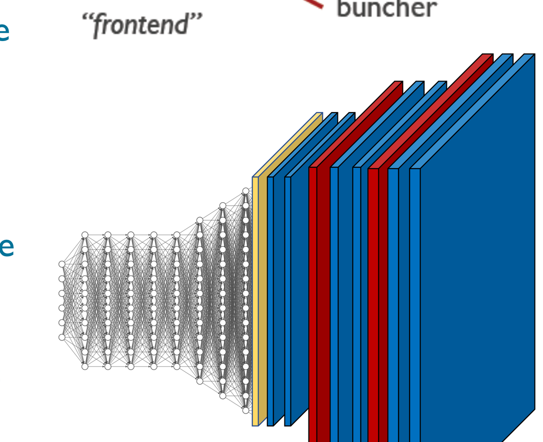
The bootstrapped ensembles offer higher increase in accuracy, but still underpredict the uncertainty. BNNs provide better coverage.

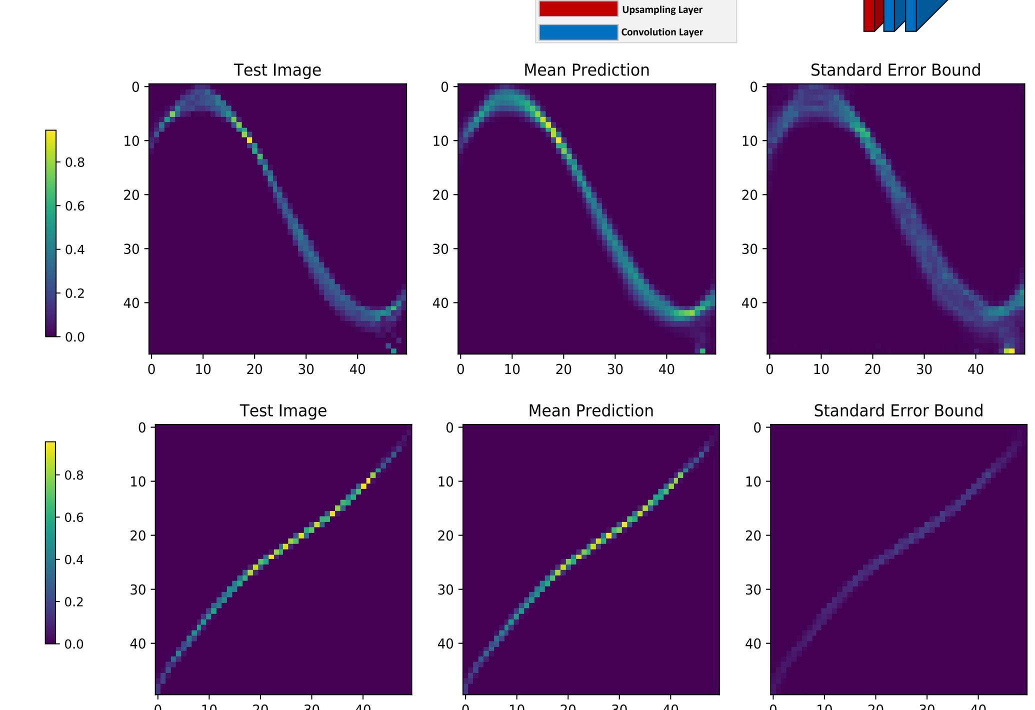
	MAE	PICP	MPIW (@ 90% PICP)
DNN	0.02	--	--
Ensemble	0.009	0.46	0.28
BNN	*0.018*	*0.91*	*0.32*

Case III: LCLS-II Injector Phase Space Images

The Beam is a collection of 6D information, 2D projections of these positions and momenta is required to be predicted by models, as a way to provide non-invasive estimates of beam phase space. We focus on the longitudinal phase space images of the LCLS-II injector. The data Is from simulations using the IMPACT code. The inputs were sampled uniformly from a 6D phase space. The neural network used an encoder-decoder architecture to generate beam phase space images from scalar inputs.







Outlook and Challenges

- Selection of better inference algorithm for inference: deterministic, hybrid, approximate Bayesian.
- Utilize Bayesian Neural Networks for Bayesian optimization for complex black box functions.
- Attribute intepretability measures for uncertainty estimates.
- Integrating physics based constraints in Bayesian Neural Network priors or optimization objectives.