ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-WEPAB308

MEASUREMENT-BASED SURROGATE MODEL OF THE SLAC LCLS-II INJECTOR

L. Gupta*, Y.-K. Kim
University of Chicago, Chicago, IL, USA
A. Edelen, N. Neveu, A. Mishra, C. Mayes
SLAC National Accelerator Laboratory, Menlo Park, CA, USA

Abstract

Machine learning models of accelerator systems ("surrogate models") are able to provide fast, accurate predictions of accelerator physics phenomena. However, approaches to date typically do not include measured input diagnostics, such as the initial beam distributions, which are critical for accurately representing the beam evolution through the system. Surrogate models that can leverage both simulation and measured data successfully are needed. We introduce an approach based on encoder-decoder style convolutional neural networks that uses the drive laser distribution and scalar settings as inputs for a photoinjector system model (here, the Linac Coherent Light Source II, or LCLS-II, injector frontend). The model is able to predict scalar beam parameters and the transverse beam distribution downstream, taking into account the impact of time-varying non-uniformities in the initial transverse laser distribution. These approaches for improving ML-based online modeling of injector systems could be easily adapted to other accelerator facilities.

INTRODUCTION

Physics simulations of particle accelerators are essential tools for predicting optimal settings for different running configurations (e.g. changing the bunch charge, bunch length). Injector systems are particularly difficult to model accurately *a priori* because of nonlinear forces such as space charge at low beam energies. These simulations can also be computationally expensive, which can be prohibitive during the design stage as well as for online use in accelerators. Thus, there is a general need for fast and reliable models which can be used for online prediction, offline experiment planning, and design of new setups. Fast models would also enable more thorough investigation of differences between physics simulations and the real machine. There is also significant effort [1–8] towards using model-based control methods in real-time machine operation and tuning, with the goal to achieve faster, higher-quality tuning. Machine learning (ML) methods may help to automate tasks such as switching between standard operating schemes, or correcting small deviations that result in poor beam quality. Fast-executing, accurate machine models can aid the development and deployment of these control methods.

Machine learning (ML) based surrogate models are one avenue toward developing fast, reliable, and realistic models of accelerators. For injector systems, data generation and model training requires significant computational resources, but once trained, ML models offer orders of magnitude faster execution speed over classical simulation methods. Amongst the many ML algorithms available, neural network (NN) based surrogate models are being widely applied for addressing the issue of execution speed and obtaining fast, non-invasive predictions of beam parameters. Several studies have verified that ML-based models can be used to support fast optimization, particularly when trained using data that spans the operational range of the physical inputs [2, 4, 9–13, 20–25].

While surrogate models trained on simulation are fast enough for use in online operation, the issue of how accurate these models are with respect to the real accelerator system also needs to be addressed. One way to address this issue is to train surrogate models on measured data, but in many cases there is not enough data to do so. Further, injector surrogate models to date do not include the full transverse laser distribution measurements as inputs, resulting in a loss of important time-varying information for accurately predicting the beam behavior.

Here, we introduce an ML-based approach to address these issues by accounting for variation in the VCC image and conducting domain transfer between the simulation and measured data. We demonstrate that including the VCC image as an input to the model improves the predictions. To achieve this, we combined scalar setting inputs for the injector with a Convolutional Neural Network (CNN) to do the image processing. A similar approach was taken in simulation in [9], and here we take the next step of including measured VCC images as inputs.

THE LCLS-II PHOTOCATHODE INJECTOR

The new Linac Coherent Light Source (LCLS) superconducting linac at the Stanford Linear Accelerator Center (SLAC), or LCLS-II, is one such facility that could benefit from having fast, accurate models for use in experiment planning, online prediction of beam distributions, and model-based control.

The injector, shown in Fig. 1, will be used for the LCLS-II project. The expected operating parameters for the injector are shown in Table 1. At the time of experimentation, the injector was in early commissioning.

With such a high repetition rate, the cathode field gradient is lower than low-repetition rate photocathode injectors [14]. Thus, the kinetic energy of the electrons as they are emitted

* lipigupta@uchicago.edu

ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-WEPAB308

and injected into the buncher is relatively low; up to 750 keV. At this energy, the dynamics are space-charged dominated. To study the dynamics in this regime and optimize the parameters for operation, particle-in-cell simulations are necessary. As such, these calculations can take several minutes to complete.

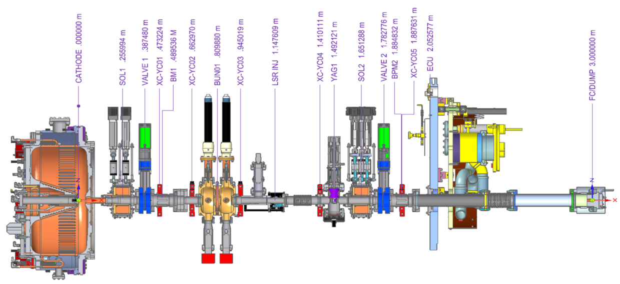

Figure 1: A schematic of the LCLS-II injector, showing each component and its position along the beamline [14].

Table 1: Designed Operating Parameters for the LCLS-II Injector

Parameter	Value	Unit
Charges	100	pC
Laser FWHM	20	ps
Laser Radius	1	mm
Field on Cathode	20	MV/m
Repetition Rate	1	MHz

SURROGATE MODEL DATA CREATION AND ARCHITECTURE

For this study, all simulated data was generated using Astra [15], and particle generation was done using distgen [16]. The SLAC-developed Python wrapper LUME-Astra [17] was used to create, set-up, and process simulated data. Measurements of laser input distributions were done by the following apparatus: the laser beam is passed through an optical splitter, such that approximately 5% of the beam intensity is directed towards a camera. The distance between the splitter and the camera is analogous to the distance between the splitter and the cathode, i.e. the transverse size of the laser at the camera location should be equal to the size at the cathode. The intensity at the camera is recorded, providing an image of how the laser intensity at the cathode appears. The measurements were recorded over many days to capture any variation.

For each laser profile, a particle bunch with 10 k particles was generated and tracked through the injector lattice in Astra to calculate various resulting beam outputs. It was determined that the bulk parameters in the simulation can be recovered with sufficient fidelity and speed using 10 k particles.

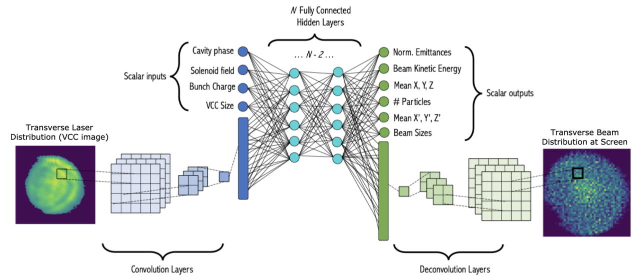

Figure 2: Encoder-Decoder CNN architecture used for prediction of beam transverse distributions and scalar beam parameters, with the VCC laser distribution as a variable input. To process the VCC images (binned into 50 × 50 pixels), the encoder consists of 3 convolutional layers with 10 filters each, alternating with max pooling layers for 2 × downsampling. The scalar input settings are concatenated into the first of 4 fully-connected layers in between the encoder and decoder. The scalar outputs are obtained from the last of these layers. Finally the decoder CNN consists of 3 convolutional layers alternating with 2 × upsampling layers, resulting in an output transverse beam prediction image with 50 × 50 bins.

Architecture and Model Training

All neural network model development and training was done using the TensorFlow and Keras libraries [18]. Each model was trained by minimizing a mean squared error loss function, using the Adam optimization algorithm [19].

The model took scalar settings for the solenoid value and charge as inputs, along with the 2-dimensional histogram representation of the laser intensity on the cathode. The laser distributions were 50 × 50 bins. The size of the laser distribution were given as the horizontal and vertical extents of the histogram, relative to the center of the histogram.These six scalar values and the 50 × 50 bin laser distribution are considered inputs to the models. The binned images were input into convolutional layers. Three convolutional layers with 10 4 × 4 filters each are applied to the image inputs. The resulting nodes are then fed to densely-connected layers. The densely-connected part of the network consists of 6 hidden layers with 1024, 512, 256, 64, 32, 16. The convolution procedure is then mirrored to deconvolve the neurons back into the scalar output nodes, and the electron beam distribution. This is shown in Fig. 2. The training was done by using 1 GPU on Cori at NERSC , and took several hours to complete.

RESULTS AND CONCLUSIONS

The results of the training on selected samples are shown in Fig. 3. Similarly, to determine if the model can successfully interpolate between the laser distributions that were seen during training, to similar distributions which have not been seen, several unique laser distributions were withheld. The predicted projections of the electron distributions are

ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-WEPAB308

shown in Fig. 4. It is clear that the macroscopic shapes and non-uniformity are captured by the surrogate model.

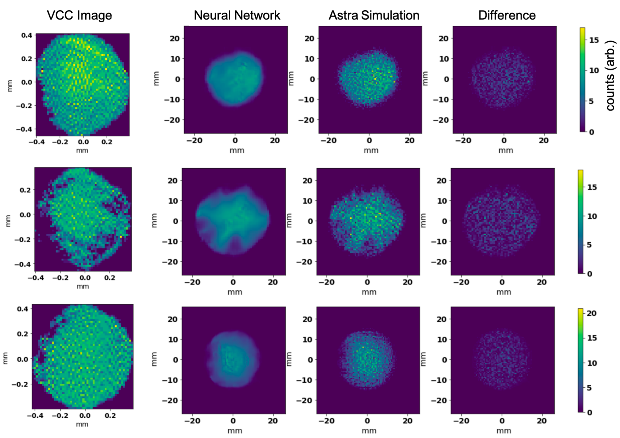

Figure 3: Examples of the neural network predictions and Astra simulation results for the transverse beam distributions. The corresponding VCC inputs used in each case are shown at left. The agreement is good, even for cases with irregular beam distributions. This demonstrates that the model can interpolate between measured input laser distributions (as seen on the VCC) and provide realistic predictions of the expected transverse beam distribution from simulation. This is important for using this model online in the accelerator, as the initial beam distribution will vary with time.

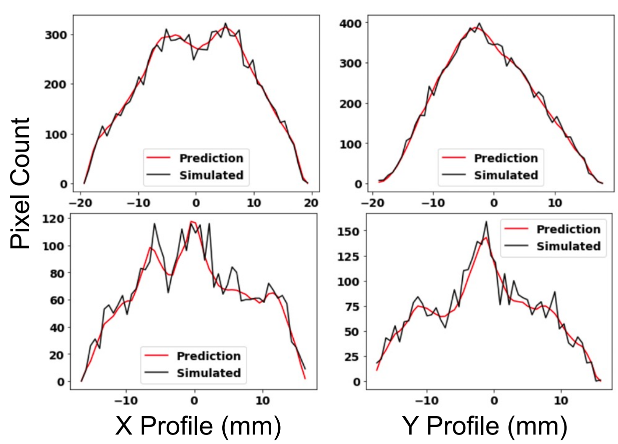

Figure 4: Predicted and simulated profiles for laser profiles withheld from training.

These results suggest that this surrogate model is able to interpolate between real laser distributions successfully, an important step toward reliable online and offline use of the model. More details can be found in [20].

ACKNOWLEDGEMENTS

The authors would like to acknowledge Jane Shtalenkova and Alex Saad for their invaluable help and expertise, and the METSD department at SLAC for CAD drawings. Funding for this work was provided in part by the U.S. Department of Energy, Office of Science, under Contracts No. DE-AC02-76SF00515 and No. DE-AC02-06CH11357, FWP 100494 through BES on B&R code KC0406020. Funding for L. Gupta provided by the U.S. Department of Energy, Office of Science Graduate Student Research (SCGSR) Program, the U.S. National Science Foundation under Award No. PHY-1549132, the Center for Bright Beams, and the University of Chicago Department of Physics.

REFERENCES

[1] J. Duris *et al.*, "Bayesian Optimization of a Free-Electron Laser," *Physical Review Letters*, vol. 124, no. 12, Mar. 2020. doi:10.1103/physrevlett.124.124801

[2] A. Edelen et al., "Recent Applications of Neural Networks to Modeling and Control of Particle Accelerators," presented at IPAC'18, Vancouver, BC, Canada, April-May 2018, paper THYGBE2, unpublished.

[3] A. Edelen *et al.*, "Opportunities in Machine Learning for Particle Accelerators," in *Proc. 1st ICFA Machine Learning Workshop for Particle Accelerators*, 2018, to be published. arXiv:1811.03172

[4] A. Edelen, N. Neveu, M. Frey, Y. Huber, C. Mayes, and A. Adelmann, "Machine learning for orders of magnitude speedup in multiobjective optimization of particle accelerator systems," *Physical Review Accelerators and Beams*, vol. 23, no. 4, p. 044601, Apr. 2020. doi:10.1103/physrevaccelbeams.23.044601

[5] A. L. Edelen, S. G. Biedron, B. E. Chase, D. Edstrom, S. V. Milton, and P. Stabile, "Neural Networks for Modeling and Control of Particle Accelerators," *IEEE Transactions on Nuclear Science*, vol. 63, no. 2, pp. 878–897, Apr. 2016. doi:10.1109/tns.2016.2543203

[6] A. Hanuka *et al.*, "Online tuning and light source control using a physics-informed gaussian process," 2019, to be published. arXiv:1911.01538

[7] A. Scheinker, A. Edelen, D. Bohler, C. Emma, and A. Lutman, "Demonstration of Model-Independent Control of the Longitudinal Phase Space of Electron Beams in the Linac-Coherent Light Source with Femtosecond Resolution," *Physical Review Letters*, vol. 121, no. 4, p. 044801, Jul. 2018. doi:10.1103/physrevlett.121.044801

[8] A. Scheinker, C. Emma, A. Edelen, and S. Gessner, "Advanced Control Methods for Particle Accelerators (ACM4PA) 2019," in *Proc. ACM4PA 2019*, Los Alamos National Laboratory, Santa Fe, NM, USA, Aug. 2019, workshop report paper, pp. 1-24, to be published.

[9] A. Edelen, J. Edelen, S. Biedron, and S. Milton. , "First steps toward incorporating image based diagnostics into particle accelerator control systems using convolutional neural networks," in *Proc. NAPAC'16*, Chicago, IL, USA, Oct. 2016, paper TUPOA51, pp. 390–393.

[10] A. Edelen, N. Neveu, C. Emma, D. Ratner, and C. Mayes, "Machine learning models for optimization and control of xray free electron lasers," in *Proc. NeurIPS Machine Learning for the Physical Sciences Workshop*, Vancouver, Canada, Dec. 2019, paper 209334724, pp. 1–5.

ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-WEPAB308

[11] C. Emma *et al.*, "Machine learning based longitudinal phase space prediction of twobunch operation at facet-ii," in *Proc. IBIC'19*, Malmö, Sweden, Sep. 2019, pp. 679–683. doi:10.18429/JACoW-IBIC2019-THBO01

[12] C. Emma, A. Edelen, M. J. Hogan, B. O'Shea, G. White, and V. Yakimenko, "Machine learning-based longitudinal phase space prediction of particle accelerators," *Physical Review Accelerators and Beams*, vol. 21, no. 11, p. 112802, Nov. 2018. doi:10.1103/physrevaccelbeams.21.112802

[13] H. Komkov, L. Dovlatyan, A. Perevalov, and D. Lathrop, "Reservoir Computing for Prediction of BeamEvolution in Particle Accelerators," in *Proc. NeurIPS Machine Learning for the Physical Sciences Workshop*, Vancouver, Canada, Dec. 2019, paper ML4PS, pp. 1–5.

[14] F. Sannibale *et al.*, "Advanced photoinjector experiment photogun commissioning results," *Physical Review Special Topics - Accelerators and Beams*, vol. 15, no. 10, p. 103501, Oct. 2012. doi:10.1103/physrevstab.15.103501

[15] *ASTRA: A Space Charge Tracking Algorithm*, DESY, Hamburg, Germany, 1997, pp. 1-117, https://www.desy.de/~mpyflo/Astra_manual/Astra-Manual_V3.2.pdf

[16] C. Gulliford, DistGen: Particle distribution generator, 2019, https://github.com/ColwynGulliford/distgen

[17] LUME-astra, https://github.com/ChristopherMayes/lume-astra

[18] Keras, https://keras.io/

[19] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization. In ICLR (Poster)," 2015, to be published. arXiv:1412.6980

[20] L. Gupta, A. Edelen, N. Neveu, A. Mishra, C. Mayes, and Y.-K. Kim, "Improving Surrogate Model Accuracy for the LCLS-II Injector Frontend Using Convolutional Neural Networks and Transfer Learning", Mar. 2021, to be published. arXiv:2103.07540

[21] A. Bartnik, C. Gulliford, I. Bazarov, L. Cultera, and B. Dunham, "Operational experience with nanocoulomb bunch charges in the Cornell photoinjector," *Physical Review Special Topics - Accelerators and Beams*, vol. 18, no. 8, p. 083401, Aug. 2015. doi:10.1103/physrevstab.18.083401

[22] C. Gulliford, A. Bartnik, I. Bazarov, B. Dunham, and L. Cultrera, "Demonstration of cathode emittance dominated high bunch charge beams in a DC gun-based photoinjector," *Applied Physics Letters*, vol. 106, no. 9, p. 094101, Mar. 2015. doi:10.1063/1.4913678

[23] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, Jun. 2002. doi:10.1613/jair.953

[24] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *Journal of Machine Learning Research*, vol. 15, no. 56, pp. 1929-–1958, 2014.

[25] S. Thrun and L. Pratt, *Learning to learn*. New York, USA: Springer Science and Business Media, 2012.