ISBN: 978-3-95450-220-2 ISSN: 2673-5474 doi: 10.18429/JACoW-FEL2022-WEP12

APPLICATION OF MACHINE LEARNING IN LONGITUDINAL PHASE SPACE PREDICTION AT THE EUROPEAN XFEL

Z. H. Zhu[1,2*]
Shanghai Institute of Applied Physics, Chinese Academy of Sciences, Shanghai, China
[1]also at University of Chinese Academy of Sciences, Beijing, China
[2]also at Deutsches Elektronen-Synchrotron DESY, Hamburg, Germany
S. Tomin, J. Kaiser
Deutsches Elektronen-Synchrotron DESY, Hamburg, Germany

Abstract

Beam longitudinal phase space (LPS) distribution is the crucial property to driving the high-brightness free-electron laser. However, the beam LPS diagnostics is often destructive and the relevant physical simulation is too time-consuming to be involved in the control room. Therefore, we explored applying the machine learning models to facilitate the virtual diagnostic of the LPS distribution at European XFEL. Two different model designs are proposed and the performance demonstrates its feasibility based on the simulations. This work lays the further investigation of the real-time virtual diagnostics in the operational machine.

INTRODUCTION

In the past decade, The X-ray free-electron laser (XFEL) facilities around the world provide coherent and ultra-short X-ray radiation with tunable wavelength and high-brightness [1], facilitating the ultra-fast scientific research and discovery with atomic spatial resolution [2–4]. In the daily operation of the facility, the electron beam with high quality is indispensable to the desirable lasing performance in the undulator sections, especially the beam longitudinal properties, such as the charge distribution and slice energy spread distribution along the beam. The acquirement of these essential beam properties requires the measurement of the entire beam longitudinal phase space (LPS) with a diagnostic instrument such as transverse deflecting structure (TDS). However, this measurement is conducted in an interceptive manner, which makes it impossible to be taken during the photon delivery to the experimental stations unless it is implemented downstream of the undulator section. In addition to that, the LPS measurement is subject to the resolution limitations from the TDS. One potential approach is beam physical simulation in which the collective effects are modeled and the beam dynamics results are well matched with experimental measurement [5]. Unfortunately, it is too computationally expensive to be applicable in the control room.

To overcome these difficulties, machine learning (ML), especially the deep neural network, has the potential to cope with system modeling in the accelerator community as a result of the rapid development of computer science. Based on it, we are enabled to construct the surrogate model with the fast-execution speed with ML, which has demonstrated its data processing capabilities in many industrial fields. With this powerful tool introduced, some ML applications in accelerators have been explored and studied in the past few years. Based on the built surrogate model with fast execution and reliable accuracy, some beam dynamics optimization projects can benefit from the orders of magnitude increase in speed [6, 7].In addition to that, it can also facilitate the automatic online tuning and beam phase space control in FEL facility [8, 9].

Because of the powerful capability of interpreting images, there has been some recent research about neural network-based virtual diagnostic of 2D beam LPS distribution in the accelerator community [10–13]. In this paper, we propose an alternative manner to predict the beam LPS distribution. This surrogate model is built to provide the prediction of the slice beam properties distributions and its corresponding LPS is reconstructed based on them. Furthermore, Convolutional neural networks (CNN) have been implemented as the second method. We compare the results from these two approaches and the preliminary study paves the path for further investigation on the online virtual diagnostic in the operational machine.

METHODOLOGY AND RESULTS

Here we demonstrate the feasibility of the machine learning-based beam LPS reconstruction with beam dynamics simulations from European X-ray Free-Electron Laser [14]. The schematic layout of the European XFEL accelerator is shown in Fig. 1. In the main linac section, the beam experiences the longitudinal density modulations at three bunch compressor chicanes with the nominated beam energy of 130 MeV, 700 MeV, and 2.4 GeV respectively. The initial beam distribution at the end of the gun cavity is simulated with ASTRA [15]. The remaining physical tracking, which simulates the beam dynamics from the gun cavity exit to the collimator section at the entrance of the undulator, is conducted with OCELOT code [16].

In the simulations, the momentum compaction factor in each dispersive section is kept as their initial values. We tune the upstream RF parameters to adjust the energy chirp at the entrance of each dispersive section to change the com-

* zihan.zhu@desy.de

ISBN: 978-3-95450-220-2 ISSN: 2673-5474 doi: 10.18429/JACoW-FEL2022-WEP12

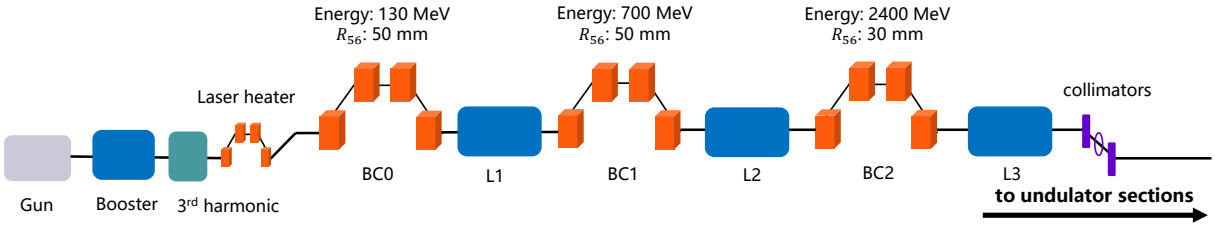

Figure 1: Schematic layout of EuXFEL linear accelerator from the gun cavity exit to the entrance of the undulator.

pression scenarios. The first three parameters are the first-, second-, and third-order coefficient in energy distribution at the injector (I1) exit, which define the detailed compression scenario in the first bunch compressor. The other two parameters control the compression strengths in the second and third chicane sections. In this preliminary study, we adjust the RF settings with relatively small variations based on the nominal design of beam dynamics. The scanning range of each parameter can be found in Table 1.

Table 1: The Five Input Features and these Values Ranges for Simulation

Parameters	Range
I1 chirp (1/m)	[-8.2, -8]
I1 curvaure ($1/m^2$)	[260, 280]
I1 skewness ($1/m^3$)	[45000, 47000]
L1 chirp (1/m)	[-13.7, -12.6]
L2 chirp (1/m)	[0, 2]

The total sample number is 40,000, 70% of which is used for model training and the rest 30% for validation. The open-source machine learning library Pytorch [17] is applied here to build the architecture for all the neural networks presented in this paper.

Slice Parameters-based LPS Prediction

The first method of the LPS distribution prediction is based on the three slice beam property distributions: current, mean energy, and slice energy spread.For each simulation in the sample set, these three distributions are generated with longitudinal beam dynamics analysis with 300 slices. The beam current profile indicates the density distribution along the beam, and the beam energy distribution can be predicted based on the Gaussian distribution assumption in each longitudinal slice.

Additionally, the bin width in the longitudinal position for each sample is collected as the scaling parameter, which describes the bunch longitudinal extent. The constructed fully-connected neural network contains the architecture of four hidden layers. In each of them, the neuron number is 256, and Rectified Linear Unit (ReLU) [18] is selected as the activation function.

After the model construction, we test its performance using examples of different compression scenarios, from

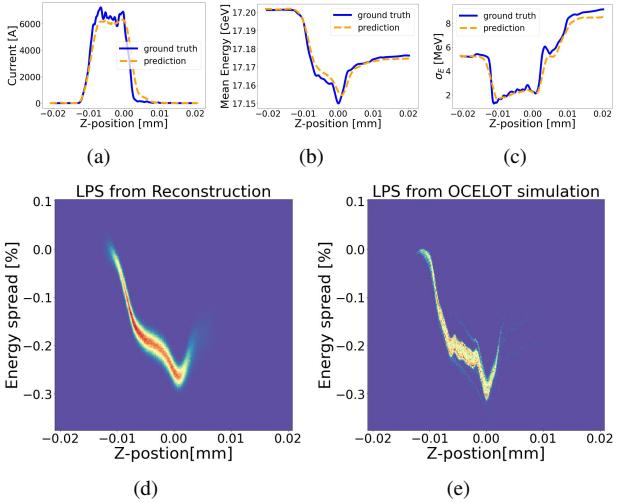

Figure 2: One example to illustrate the reconstruction according to the three property distribution (a-c). (d) The reconstructed LPS distribution based on the NN prediction. (e) The groud truth from OCELOT simulation. The head of the bunch is on the left of the coordinate.

which a good agreement can be found between the model prediction and the corresponding simulation result. Fig. 2 shows the beam property distributions from one validation case of normal compression intensity. In Fig. 2(d), the reconstructed LPS distribution matches well with the one from the simulation that is shown in Fig. 2(e).

Based on this model, the immediate beam LPS distribution can be obtained and this tool is eligible for being introduced in the control room to provide real-time online virtual diagnostic. Furthermore, it is noteworthy that the training process during the model construction is fast, which means it is feasible to introduce the model in the control room and retrain it based on the existing one with updated measurement samples. Consequently, its accuracy will be enhanced as well as its robustness. One potential issue for this method is some information in the LPS distribution is lost during the reconstruction, which is illustrated in Fig. 3. This problem results from the outliers in the LPS distribution, which can be observed in Fig. 3(e). This distribution leads to a larger slice energy spread, resulting in inaccuracy in the prediction based on the Gaussian distribution assumption.

ISBN: 978-3-95450-220-2 ISSN: 2673-5474 doi: 10.18429/JACoW-FEL2022-WEP12

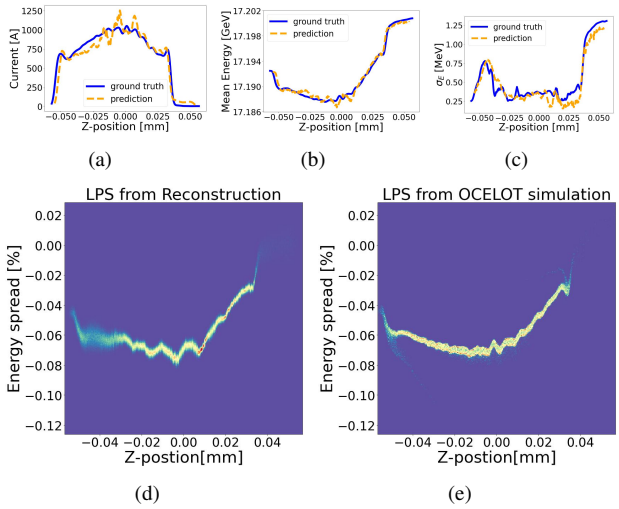

Figure 3: Another example from the validation set in which the distribution in the bunch head and tail is spoiled.

Image-based LPS Prediction

Other than the method presented above, we also give an attempt on treating the entire LPS distribution as an image and conduct image processing. The power of convolutional encoder-decoder network, which is introduced to this image regression project, has been demonstrated with applications from numerous industrial fields including facial recognition, image classification, speech recognition, etc. For each sample, its 2D LPS image is generated after binning the beam distribution in phase space with a 2D histogram of the same shape. Here we construct the convolutional encoder-decoder network to power the virtual diagnostics of LPS distribution after multi-stage longitudinal density modulation and compare its performance of it with the previous method.

The architecture of the network is shown in Fig. 4. Its output is an image of 300×300 pixels that displays the 2D distribution in the phase space. The other fully connected neural network is deployed to provide the resolution prediction based on the same input features. Due to the different bunch lengths between the samples, the pixel resolutions in the two coordinates vary between the samples as well. For a nearly full compression case, the time resolution for the LPS image can be as high as 0.3 fs/pixel.

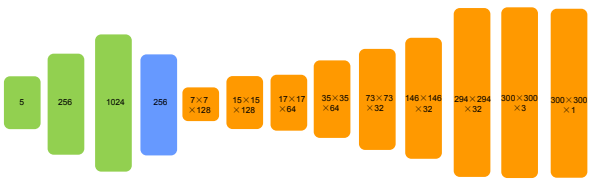

Figure 4: The contains one input layer, which is followed by three fully connected layers (green). The latent layer is highlighted in blue. Downstream it, the eight transposed convolutional layers (orange) are built as the decoder.

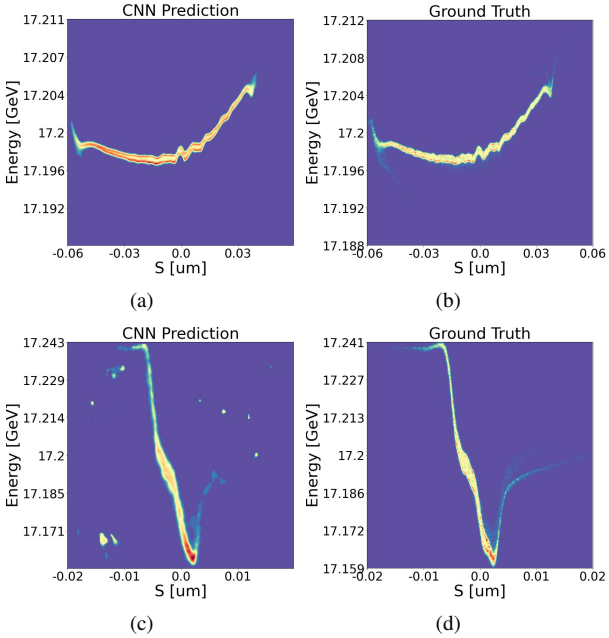

Figure 5: Two examples as the performance of the CNN model. The top row shows the LPS from a long bunch case (under compression), and the bottom one shows the short bunch case (full compression).

As evidenced in Fig. 5, there is a good agreement between the prediction from the model and the ground truth. Compared with the first method whose neural network is designed to predict the slice beam property distributions, the most obvious advantage of this method is the predicted 2D image can reproduce the detail in the beam distribution in LPS, such as the microbunching structure and the outliers around the bunch head. Concerning the training duration, it takes nearly 20h on one Graphics Processing Unit node, which is much more computationally-intense than the previous method in which the model training is easily conducted in a single CPU within 5 minutes.

CONCLUSION

In this paper, the machine learning-based prediction of the beam longitudinal phase space distribution is presented. Two different strategies are validated with simulation samples and the pros and cons of each are discussed. According to performance in the validation set, both of them achieve good and fast prediction of the beam LPS distribution after multi-stage longitudinal charge modulations. With its fast execution and reliable prediction, the online virtual diagnostic and optimization of beam longitudinal properties can be achievable. The performance test of the constructed model on the operational machine is in the following schedule and it lays the foundation for further exploration of machine learning applications in the accelerator operation.

ISBN: 978-3-95450-220-2 ISSN: 2673-5474 doi: 10.18429/JACoW-FEL2022-WEP12

ACKNOWLEDGEMENTS

This work was supported by the Youth Innovation Promotion Association CAS (2021282) and CAS-DAAD Program for Promotion of Outstanding Young Scholars, 2021 (57575641).

REFERENCES

[1] C. Pellegrini, A. Marinelli, and S. Reiche, "The physics of x-ray free-electron lasers," *Rev. Mod. Phys.*, vol. 88, p. 015 006, 1 2016. doi:10.1103/RevModPhys.88.015006

[2] H. Öström *et al.*, "Probing the transition state region in catalytic co oxidation on ru," *Science*, vol. 347, no. 6225, pp. 978–982, 2015. doi:10.1126/science.1261747

[3] C. Dejoie *et al.*, "Serial snapshot crystallography for materials science with swissfel," *IUCrJ*, vol. 2, no. 3, pp. 361–370, 2015. doi:10.1107/S2052252515006740

[4] J. M. Martin-Garcia, C. E. Conrad, J. Coe, S. Roy-Chowdhury, and P. Fromme, "Serial femtosecond crystallography: A revolution in structural biology," *Arch. Biochem. Biophys.*, vol. 602, pp. 32–47, 2016, Protein Crystallography. doi:10.1016/j.abb.2016.03.036

[5] I. Zagorodnov, S. Tomin, Y. Chen, and F. Brinker, "Experimental validation of collective effects modeling at injector section of x-ray free-electron laser," *Nucl. Instrum. Methods Phys. Res., Sect. A*, vol. 995, p. 165 111, 2021. doi:https://doi.org/10.1016/j.nima.2021.165111

[6] J. Wan, P. Chu, and Y. Jiao, "Neural network-based multiobjective optimization algorithm for nonlinear beam dynamics," *Phys. Rev. Accel. Beams*, vol. 23, no. 8, p. 081 601, 2020. doi:10.1103/PhysRevAccelBeams.23.081601

[7] 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," *Phys. Rev. Accel. Beams*, vol. 23, no. 4, p. 044 601, 2020. doi:10.1103/PhysRevAccelBeams.23.044601

[8] Z. Zhu, Y. Chen, W. Qin, M. Scholz, and S. Tomin, "Machine Learning Based Surrogate Model Construction for Optics Matching at the European XFEL," in *Proc. IPAC'22*, Bangkok, Thailand, 2022, paper MOPOTK013, pp. 461–464. doi:10.18429/JACoW-IPAC2022-MOPOTK013

[9] 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," *Phys. Rev. Lett.*, vol. 121, p. 044 801, 4 2018. doi:10.1103/PhysRevLett.121.044801

[10] 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," *Phys. Rev. Accel. Beams*, vol. 21, p. 112 802, 2018. doi:10.1103/PhysRevAccelBeams.21.112802

[11] A. Hanuka *et al.*, "Accurate and confident prediction of electron beam longitudinal properties using spectral virtual diagnostics," *Sci. Rep.*, vol. 11, no. 1, pp. 1–10, 2021. doi:10.1038/s41598-021-82473-0

[12] J. Zhu, Y. Chen, F. Brinker, W. Decking, S. Tomin, and H. Schlarb, "High-fidelity prediction of megapixel longitudinal phase-space images of electron beams using encoder-decoder neural networks," *Phys. Rev. Appl.*, vol. 16, p. 024 005, 2021. doi:10.1103/PhysRevApplied.16.024005

[13] A. E. Pollard, D. J. Dunning, and M. Maheshwari, "Learning to Lase: Machine Learning Prediction of FEL Beam Properties," Shanghai, China, Oct. 2021, presented at ICALEPCS'21, Shanghai, China, Oct. 2021, paper WEPV020, unpublished.

[14] W. Decking *et al.*, "A MHz-repetition-rate hard x-ray free-electron laser driven by a superconducting linear accelerator," *Nat. Photonics*, vol. 14, no. 6, pp. 391–397, 2020. doi:10.1038/s41566-020-0607-z

[15] K. Flöttmann *et al.*, *ASTRA: A space charge tracking algorithm*, 2011.

[16] S. I. Tomin, I. V. Agapov, M. Dohlus, and I. Zagorodnov, "OCELOT as a Framework for Beam Dynamics Simulations of X-Ray Sources," in *Proc. IPAC'17*, Copenhagen, Denmark, May 2017, pp. 2642–2645. doi:10.18429/JACoW-IPAC2017-WEPAB031

[17] A. Paszke *et al.*, "Pytorch: An imperative style, high-performance deep learning library," *Adv. Neural Inf. Process Syst.*, vol. 32, 2019.

[18] A. F. Agarap, "Deep learning using rectified linear units (relu)," *CoRR*, vol. abs/1803.08375, 2018. http://arxiv.org/abs/1803.08375