6D BEAM PHASE SPACE MEAASUREMENT BY CONVOLUTED NEURAL NETWORK

S. Mukherjee*, M. Kuriki, Z. Liptak, L. Guo,
Hiroshima University, Hiroshima, Japan

Abstract

The six-dimensional (6D) phase space distribution of beam is an extremely important indicator of beam performance and provides useful information for understanding the actual state of the accelerator. On the other hand, the beam diagnostics for the 6D phase space is generally difficult and only a projection on a 1D or 2D phase space is usually obtained. We developed an algorithm based on Convolutional Neural Network (CNN) to reconstruct the 6D phase space with a limited number of transverse beam images in *x*-*y* plane. The advantage of this method is that it does not require as many computing resources as conventional back projection techniques. In this presentation, using the STF injector as an example, we show through simulation that a six-dimensional phase space can be reconstructed only from 4+4 beam images. The reconstructed 6D beam distribution matched the simulation data with high accuracy. Data for experimental verification of this method has already been obtained using the ATF injector, and analysis is currently underway.

INTRODUCTION

Six-dimensional phase space information is essential for understanding various beam parameters such as beam size, momentum spread, pulse length, energy spread, emittances, and various correlations among those variables. Recent and next-generation particle accelerators are thus dependent on these profiles [1]. C. B. McKee, P. G. O'Shea, and J. M. J. Madey [2], as well as S. Hancock [3], first introduced the technique of reconstructing transverse and longitudinal phase space using back projection. In recent years many studies have focused on improving these techniques by applying Machine Learning algorithms; for example, Roussel *et al.* [4] and Wolski *et al.* [5], among others.

We developed a simple CNN-based algorithm to reconstruct the six-dimensional phase space. It consists of the input layer, convolution layer, pooling layer, Rectified Linear unit (ReLU), fully connected layer, and a loss function [6]. The network starts by taking the raw *x*-*y* distribution data which can be obtained as a beam image on a screen monitor. The beam image is transformed into a 2D histogram. This image is then fed into a convolution layer where filters are applied to extract important features. A ReLU function introduces non-linearity into the process, which enables the network to capture more complex patterns. Subsequently, a pooling layer reduces the spatial dimensions of the image, simplifying the data. Finally, the processed image is flattened into a one-dimensional array and passed through fully connected layers that combine the extracted features and learn complex relationships, ultimately mapping them to the final output.

In this study, we used the STF injector which is composed from a L-band RF Gun (1.3 GHz, 1.5 cell, copper cavity) with a couple of solenoid magnets (main solenoid and bucking solenoid) followed by a chicane as a model. The beam image in the middle of chicane where there is a dispersion, is observed because of the sensitivity to the longitudinal space. We reconstruct the 6D phase space at the cathode by observing the beam images. We developed two separate algorithms: one for learning the non-linear evolution of 2D longitudinal phase space as a function of the RF phases of the RF Gun, and another for generating 4D transverse phase spaces that are rotated using solenoid fields. We will report the results.

METHODOLOGY

In our study, we use the ASTRA [7] to perform the simulation of the KEK-STF. The beam is generated at the cathode and accelerated by the RF Gun up to several MeV. The acceleration is a rotation in the longitudinal phase space and the rotation "angle" depends on the RF phase. Solenoid fields (Main and Bucking) rotate the beam in transverse phase space. In short, the beam is rotated in 6D phase space by varying the RF phase and solenoid fields. Due to the dispersion in the middle of the chicane, the transverse beam image in *x*-*y* plane is a projection not only from the 4D transverse phase space (x-y-x'-y'), but also the 2D longitudinal phase space (z-E). The *x*-*y* beam image at the chicane is a projection on the *x*-*y* space, which is sensitive not only for the transverse phase space, but also the longitudinal phase space.

1.3 GHz L-band normal conducting RF gun [8] employes Cs_2Te cathode driven with a 266 nm laser. We assume an isotropic Fermi Dirac momentum distribution at the photocathode, whose geometrical emittances are given by the following expressions [9]:

$$\epsilon_{x,y} = \sigma_{x,y} \sqrt{\frac{E_{\mathrm{phot}} - \Phi_{\mathrm{eff}}}{3\, m_0\, c^2}}, \tag{1}$$

where $\sigma_{x,y}$ are the r.m.s. beam sizes in x and y, E_{phot} is the photon energy, Φ_{eff}, the work function of the photocathode, and m_0c^2 is the rest mass energy of electron respectively. The photon energy is assumed constant at 4.66 eV, and the work function of Cs_2Te photocathode is kept fixed at 3.3 eV at room temperature.

A multimodal CNN is constructed by training using simulation data. We generate 435 distinct phase space images

* d223836@hiroshima-u.ac.jp, sayantan.96.phy@gmail.com

(ground truths) at the photocathode. For each ground truth, measurements are taken at the chicane under four different conditions—by varying either the RF phase (for the longitudinal case) or the solenoid field (for the transverse case)—resulting in approximately 1700 corresponding real-space images. In the longitudinal configuration, the RF gun phase is varied from 10° to 55° at 15° intervals while the x-y distribution is measured at the chicane's center. The multimodal CNN is then constructed which accepts these four different x-y distribution images, together with an array of RF phases, and maps them to the t-Δp_z distribution at the photocathode (treated as the ground truth). In the transverse configuration, a separate CNN is used where the peak field of the solenoid's main coil is varied from 0.16 T to 0.19 T in 0.01 T increments, and this parameter is similarly arranged in an array. Here, too, the network takes the x-y distribution images measured at the chicane but maps them to two ground truths, x-x' and y-y', at the photocathode region. These procedures are illustrated in Fig. 1, and training is performed by considering various combinations of initial beam parameters such as bunch charge, laser pulse duration, and transverse beam sizes, as detailed in Table 1.

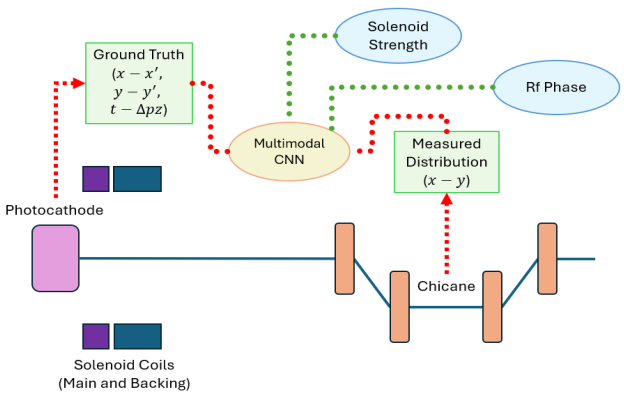

Figure 1: Overview of the simulation workflow for multimodal CNN at the injector region of KEK-STF beamline.

Table 1: Beam Parameters at Cathode for Ground Truth

Parameter	Range	Step Size	Unit
Bunch charge	20–60	20	pC
Laser pulse duration (FWHM)	4.9–19.1	4.73	ps
Beam size in x-direction	0.5–3.0	0.5	mm
Beam size in y-direction	0.5–3.0	0.5	mm

Network Architecture

The CNN model uses the layers described in Ref. [6], with the architectures designed for longitudinal and transverse phase space reconstruction shown graphically in Fig. 2(a) and (b), respectively. The measured x-y distributions at the chicane are represented as 2-D histograms with a resolution of 64×64 bins. These images pass through four convolution

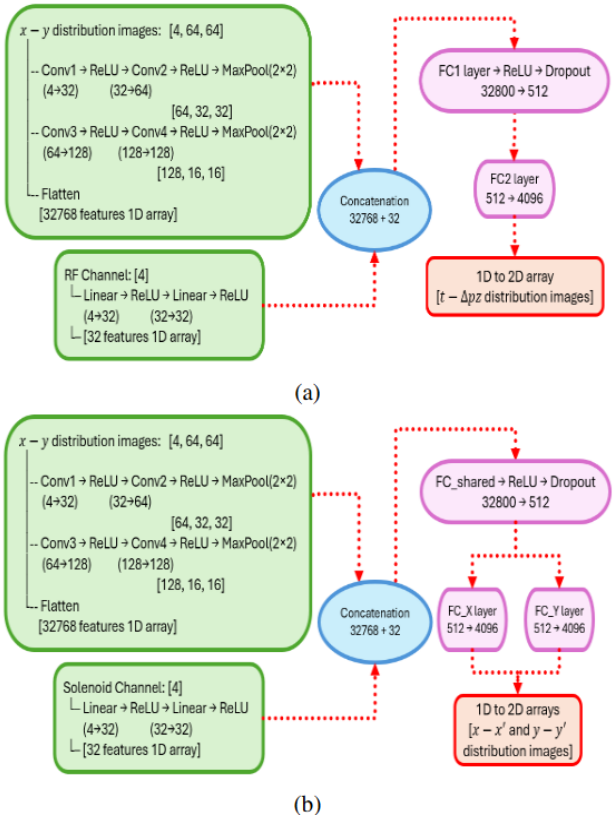

Figure 2: Multimodal CNN architectures for reconstructing (a) 2D longitudinal phase space and (b) 4D transverse phase space.

layers that produce 128 feature maps, which are then downsampled by max pooling to a resolution of 16×16 and finally flattened into 32 768 features. The RF phase or solenoid field, provided as four-scalar values, is nonlinearly transformed into a 32-dimensional feature vector that is concatenated with the image features. A first fully connected (FC) layer reduces this concatenated feature vector to 512 features, and a second FC layer expands it to 4096 features. These are reshaped to form an output image with a 64×64 resolution. The longitudinal CNN uses one such layer, specifically, the second FC layer to yield t-Δpz distribution, while the transverse CNN uses two similar layers to generate x-x' and y-y' distributions.

Loss Function and Hyperparameter Tuning

The CNN algorithm employs the Adam optimizer [10] to minimize the error between the actual and predicted values using Smooth L1 loss function, which combines the strengths of both Mean Square Error (MSE) and Mean Absolute Error (MAE). For small errors, it uses a squared penalty for smooth gradients, and for large errors, it switches to an absolute penalty to reduce outlier effects. The dataset is split into 60% for training and 40% for validation. The losses on training and validation data are optimized by tuning the hyperparameters such as learning rate, epoch, batch size, and dropout. Optimized results of training on the longitudinal dataset are shown in Fig. 3(a). The horizontal axis

represents the full training cycles (epochs) while the vertical axis shows the loss function. The training and validation loss curves, shown by blue and orange lines, converge to 9.0×10^{-4} and 8.3×10^{-2}, implying no significant overfitting or underfitting in the process. Figure 3(b) shows the sum of the losses on the x-x' and y-y' datasets. Here, the training and validation losses have converged to 3.4×10^{-1} and 4.4×10^{-1}, respectively. The transverse model exhibits minimal overfitting due to slightly higher validation loss; however this remains within acceptable limits and does not adversely affect the model's performance.

Table 2: Simulation and CNN Results for t-Δp_z Distribution

Method	ε (keV/c)·ps	α 10^{-3}	β ps/(keV/c)	γ (keV/c)/ps
ASTRA	1.60	3.1	23.13	0.044
CNN	1.61	3.4	22.97	0.043

Table 3: Simulation and CNN Results for x-x' and y-y' Distributions

Space	Method	ε_n mrad.mm	α 10^{-3}	β mm	γ 1/mm
x-x'	ASTRA	1.56	2.21	1.35	0.881
	CNN	1.63	0.46	1.20	0.832
y-y'	ASTRA	2.05	12.4	1.43	0.697
	CNN	2.10	11.9	1.49	0.672

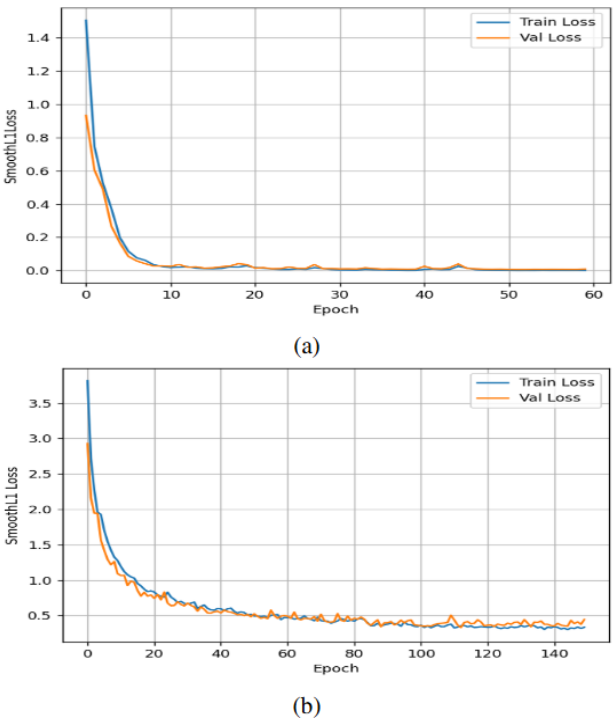

Figure 3: Training and Validation loss curves for (a) 2D longitudinal and (b) 4D transverse datasets.

RESULTS

After the training is over, the optimized CNN model is saved for future predictions. Next, we generate test distributions unknown to the training set by providing hidden initial beam parameters that lie within the range of the training values, but the parameters are not included in the training sample. The longitudinal distribution t-Δpz generated by ASTRA and reconstructed by CNN are shown in Fig. 4, where the vertical axis is the momentum in keV/c and the horizontal axis is the time in ps.

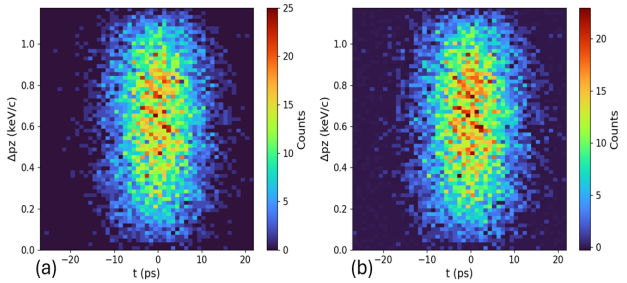

Figure 4: Longitudinal phase space (t-Δpz) for a test sample at photocathode is shown using (a) ASTRA Simulation and (b) CNN Reconstruction.

Similarly, the transverse distributions x-x' and y-y', are shown in Figs. 5 and 6, where the vertical axis is in mrad and the horizontal axis is in mm. Tables 2 and 3 list the Emittances and Twiss parameters for longitudinal and transverse phase space by ASTRA and CNN for comparison. These parameters are reproduced with very high accuracy in the order of 10^{-3} to 10^{-1}. One exception is α in x space.

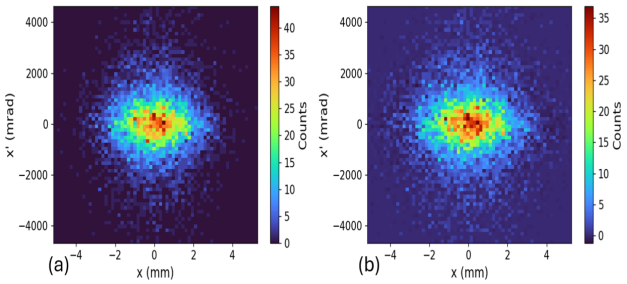

Figure 5: Transverse phase space (x-x') for a test sample at photocathode is shown using (a) ASTRA Simulation and (b) CNN Reconstruction.

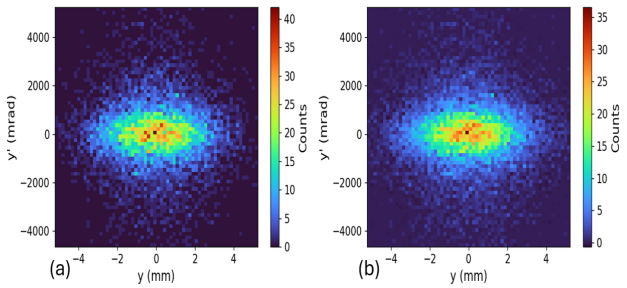

Figure 6: Transverse phase space (y-y') for a test sample at photocathode is shown using (a) ASTRA Simulation and (b) CNN Reconstruction.

CONCLUSION

We developed a multimodal CNN model to reconstruct the 6D phase space by observing *x*-*y* beam images at a section with a dispersion. By assuming the KEK-STF injector , which consists of the L-band RF Gun, solenoid magnets, and a chicane, we demonstrated the 6D phase space reconstruction. With 4+4 *x*-*y* images at the chicane center varying RF phase of the RF Gun and the main solenoid field to obtaine the longitudinal phase space and the transverse phase space, respectively; those phase space were reconstructed with a high accuracy, typically in the order of 10^{-3} to 10^{-1} relative. We conclude that the CNN model demonstrated promising performance in fully reconstructing the phase spaces that show consistency with simulation results. Furthermore, its efficiency in requiring fewer readings minimizes the loss of valuable beam time. In principle, this method enables the reconstruction of any two-dimensional spatial distribution within the 6D phase space such as *x*-δ, *y*-*z*, etc., and we are currently developing such an algorithm. Data has already been collected at KEK-ATF for experimental verification. In the experiment, the *x*-*y* beam images at a chicane were taken with 4×4 combinations of 4 phases and 4 solenoid currents to reconstruct any two variable spaces within the 6D phase space. The development of the new algorithm and analysis are underway.

Until now, measuring 6D phase space required high-performance beam monitors and large amounts of data. With this method, it has become clear that 6D space can be reconstructed with surprising accuracy from a small amount of data by incorporating the latest AI techniques based on CNN algorithms. This method enables a more detailed understanding of accelerator beams. We expect this method to contribute to the advancement of accelerator science.

ACKNOWLEDGEMENT

We would like to express our gratitude to Professors Okuki, Sakai, Terunuma, and Kirk Yamamoto of KEK. The information on the KEK-STF beamline was indispensable for performing simulations. We are also deeply grateful for the opportunity to acquire data at KEK-ATF with their warm support and for providing us with various information.

REFERENCES

[1] J. Blazey *et al.*, "Accelerator and beam physics roadmap: August 2022 workshop," in *Proc. of the DOE HEP Accelerator Beam Physics Roadmap Workshop*, Bethesda, MD, USA, September 6-8, 2022, https://science.osti.gov/hep/-/media/hep/pdf/2022/ABP_Roadmap_2023_final.pdf

[2] C. B. McKee, P. G. O'Shea, and J. M. J. Madey, "Phase space tomography of relativistic electron beams," *Nucl. Instrum. Methods Phys. Res., Sect. A*, vol. 358, pp. 264-267, 1995.

[3] S. Hancock, "A Simple Algorithm for Longitudinal Phase Space Tomography," CERN, Geneva, Switzerland, Rep. PS/RF/Note 97-06, May 1997.

[4] R. Roussel *et al.*, "Efficient six-dimensional phase space reconstructions from experimental measurements using generative machine learning," *Phys. Rev. Accel. Beams*, vol. 27, p. 094601, 2024. doi:10.1103/PhysRevAccelBeams.27.094601

[5] A. Wolski *et al.*, "Accelerator beam phase space tomography using machine learning to account for variations in beamline components," *J. Instrum.*, vol. 19, p. P07013, 2024. doi:10.1088/1748-0221/19/07/P07013

[6] L. Alzubaidi *et al.*, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *J. Big Data*, vol. 8, p. 53, 2021. doi:10.1186/s40537-021-00444-8

[7] K. Floettmann, "ASTRA: A Space-Charge Tracking Algorithm," DESY, Hamburg, Germany, 2017, https://www.desy.de/~mpyflo/

[8] H. Sugiyama *et al.*, "L-band photocathode RF gun at KEK-STF," *J. Phys. Conf. Ser.*, vol. 298, p. 012015, 2011. doi:10.1088/1742-6596/298/1/012015

[9] K. Floettmann, *ASTRA: A Space-Charge Tracking Algorithm*, Version 3.2, DESY, Hamburg, Germany, March 2017, https://www.desy.de/~mpyflo/Astra_manual/Astra-Manual_V3.2.pdf

[10] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *arXiv:1412.6980*, 2015. doi:10.48550/arXiv.1412.6980