Machine Learning for Beam Correction Study of the Injection Beamline at Wuhan Advanced Light Source

Authors: Hui Wang, Hao-Hu Li, Jia-Bao Guan, Yuan Chen, Ji-Ke Wang, Yuan-Cun Nie, Jian-Hua He, Yuan-Cun Nie

Date: 2025-03-10T00:00:00+00:00

Abstract

As a fourth-generation synchrotron radiation light source working at 1.5 GeV, Wuhan Advanced Light Source (WALS) is being designed, which uses a full-energy linear accelerator (LINAC) as its electron beam injector. The injection beamline adopts a three-stage scheme: firstly, the beam from the LINAC that is 6 m under the storage ring is horizontally deflected below the storage ring, then it gradually climbs from underground to the same altitude as the storage ring, and finally the beam is delivered horizontally into the injection straight section inside the storage ring. Meanwhile, the Twiss parameter matching between the LINAC and storage ring is completed. During the construction of the beamline, magnet manufacturing errors, installation errors and beam injection errors from the LINAC will cause beam deviations from predetermined ideal orbits, and even particle losses. As a result, the electron beam correction is required during beam commissioning. Different from the single-direction beam correction of general transfer lines, the horizontal and vertical directions of the beam are coupled in the WALS injection transfer line, which greatly increases the complexity and difficulty of beam correction. Machine learning technology has been developed extensively in recent years, and its powerful algorithm of invertible neural network model is expected to be able to solve the beam commissioning difficulty of the beam injection transfer line at the WALS. Therefore, an invertible neural network model has been designed and trained to simulate the beam transport and beam correction of the WALS injection beamline. By optimizing the number and location of beam profile diagnostics, the accuracy of bidirectional prediction and beam correction effect can be greatly improved. The method is of great practical significance for the commissioning and operation of similar complex beam transport systems.

Full Text

Preamble

Machine Learning for Beam Correction Study of the Injection Beamline at Wuhan Advanced Light Source

Hui Wang,[1,2] Hao-Hu Li,[3] Jia-Bao Guan,[1,2] Yuan Chen,[1,2] Ji-Ke Wang,[1,2] Yuan-Cun Nie,[1,2,†] and Jian-Hua He[1,2]
[1]The Institute for Advanced Studies, Wuhan University, Wuhan 430072, China
[2]Advanced Light Source Research Center, Wuhan University, Wuhan 430072, China
[3]Shanghai APACTRON Particle Equipment Co. Ltd, Shanghai 201800, China

The Wuhan Advanced Light Source (WALS) is being designed as a fourth-generation synchrotron radiation light source operating at 1.5 GeV, utilizing a full-energy linear accelerator (LINAC) as its electron beam injector. The injection beamline employs a three-stage scheme: first, the beam from the LINAC, located 6 m beneath the storage ring, is horizontally deflected below the storage ring; then it gradually climbs from underground to the same altitude as the storage ring; and finally the beam is delivered horizontally into the injection straight section inside the storage ring, completing Twiss parameter matching between the LINAC and storage ring.

During beamline construction, magnet manufacturing errors, installation errors, and beam injection errors from the LINAC will cause beam deviations from predetermined ideal orbits, potentially leading to particle losses. Consequently, electron beam correction is required during beam commissioning. Unlike the single-direction beam correction typical of general transfer lines, the horizontal and vertical directions of the beam are coupled in the WALS injection transfer line, which greatly increases the complexity and difficulty of beam correction. Machine learning technology has developed extensively in recent years, and its powerful invertible neural network model algorithms are expected to solve the beam commissioning challenges of the beam injection transfer line at WALS.

Therefore, an invertible neural network model has been designed and trained to simulate beam transport and correction for the WALS injection beamline. By optimizing the number and location of beam profile diagnostics, the accuracy of bidirectional prediction and beam correction effectiveness can be greatly improved. This method holds significant practical significance for the commissioning and operation of similar complex beam transport systems.

Keywords: Beam correction, Injection transfer beamline, Machine learning, Beam dynamics, Invertible neural network

Introduction

Synchrotron radiation light sources are large-scale facilities for observing and studying the microscopic world using synchrotron radiation phenomena [1], with accelerators being essential components. After decades of development, these facilities have evolved from first-generation part-time light sources that simply utilized synchrotron radiation generated from particle accelerators to fourth-generation diffraction-limited storage ring light sources featuring ultra-low beam emittance and numerous undulator insertions. Several fourth-generation synchrotron radiation light source facilities are currently in operation or under construction worldwide. The MAX IV facility in Sweden employs multi-bend achromat (MBA) lattices to reduce beam emittance below 0.1 nm · rad and increase brightness by two orders of magnitude [2]. Brazil' s Sirius achieves even lower natural beam emittance compared to MAX IV through stronger horizontal focusing [3]. The ESRF-EBS in Europe represents the world' s first 6 GeV storage ring light source upgraded from ESRF, significantly enhancing brightness and coherence [4]. HEPS [5], HALF [6], and several other fourth-generation synchrotron radiation light sources are under construction. Synchrotron radiation light sources have found widespread applications in physics, materials chemistry, biology, life sciences, electronic information engineering, and numerous other fields [7].

The Wuhan Advanced Light Source (WALS), currently under design, is a fourth-generation diffraction-limited synchrotron radiation light source employing a 1.5 GeV full-energy linear accelerator (LINAC) as its injector [8]. Electron beams are transported to the 1.5 GeV storage ring via the beam injection transfer line (BITL) [9–11]. To accommodate future construction of a 4.0 GeV storage ring, the LINAC is positioned 6 m underground and laterally offset relative to the 1.5 GeV storage ring. This configuration necessitates a "three-stage" design scheme for the WALS BITL: first, the beam is horizontally deflected below the storage ring; then it gradually climbs from the underground tunnel to the same altitude as the storage ring, completing a vertical ascent of 6 m; and finally the beam is transferred horizontally into the injection straight section inside the storage ring [12]. This special design scheme introduces horizontal and vertical beam coupling during transport, particularly affecting the dispersion function, which requires simultaneous correction in both transverse directions during beam commissioning. For such complex beam dynamics problems, where beam diagnostics equipment is limited in practice, neither traditional beam commissioning methods like the SVD (Singular Value Decomposition) linear fitting algorithm [13–16] nor the global beam commissioning method proposed for the WALS BITL [12] is optimal.

It is therefore necessary to leverage machine learning algorithms to simulate beam transport and perform horizontal and vertical beam corrections simultaneously for the WALS BITL. In recent years, machine learning algorithms have been widely applied in accelerator physics. The multi-hidden layer structure of neural networks and the algorithmic activity provided by activation functions can accurately realize beam dynamic simulations of particle beams in accelerators. Applications have been reported at many light sources. For example, SLAC National Accelerator Laboratory trained two independent neural network constructs at its Linac Coherent Light Source (LCLS) [17, 18] to achieve accurate prediction of 2D images of electron bunches [19]. The Shanghai Synchrotron Radiation Facility (SSRF) [20-23] used machine learning to extract beam information bunch by bunch and correct beam orbits in the storage ring [24]. Using convolutional neural networks (CNN) [25, 26], the National Synchrotron Radiation Laboratory (NSRL) at the University of Science and Technology of China (USTC) successfully fitted and calculated beam cross-section dimensions through noise suppression of collected beam images [27]. Additionally, online correction using Lasso regression [28, 29] and Beta function correction using neural networks have been carried out [30]. Increasingly, light sources are employing machine learning algorithms to solve problems in accelerator physics [31, 32].

In our study, an invertible neural network (INN) [33, 34] containing 8 affine coupling modules [35] has been successfully designed after extensive testing, employing the ReLU (Rectified Linear Unit) activation function [36, 37] and incorporating the Adam (adaptive moment estimation) algorithm [38] and the Back-propagation algorithm [39, 40]. By learning a dataset consisting of two groups of variables—one group being the magnet K-values of (m−2) quadrupole magnets with adjustable electric current in the WALS BITL (as input x), and another group being the horizontal (σx) and vertical (σy) beam sizes collected at beam profile monitoring systems (as output y)—the model can realize not only forward prediction from magnet K-values to beam size σ, but also inverse prediction from beam size σ to magnet K-values. The AT toolbox in MATLAB has been used to model the WALS BITL, collect beam-related data, and verify the inverse prediction results of the INN model. Optimization of beam correction has been successfully achieved in simulation for the WALS BITL.

The remainder of this paper is organized into five sections. Section II introduces the layout of the WALS BITL and the global beam correction method proposed during initial beamline design. Section III presents the INN model designed for the WALS BITL, described from six aspects: model structure assumption, beam correction process, placement and quantity setting of beam profile monitoring systems, dataset preparation, model structure setup, and training results. Section IV verifies the beam correction method of the INN model. Section V analyzes the beam correction method, introducing the dynamic relationship, loss function definition, and evaluation indicator selection. Summaries and conclusions are presented in Section VI.

II. The WALS BITL and Global Beam Correction

A. Design of the WALS BITL

For synchrotron radiation light sources, the injection transfer line must deliver the electron beam into the storage ring with high quality and high transmission efficiency. Stable operation of a synchrotron radiation light source highly relies on precise design and construction of the injection transfer line. When designing the transfer line, one must ensure accurate matching of Twiss parameters and maintain control of dispersion functions. Simultaneously, considering installation errors, injection errors from the LINAC, magnet manufacturing errors, and other factors during beamline construction, beam correction must be performed based on theoretical design. Typically, beam diagnostic devices such as BPMs [41] and beam profile monitoring systems are placed along the beamline to collect beam information, with horizontal/vertical correction (HVC) magnets arranged for beam orbit corrections.

At WALS, the BITL consists of three sections as mentioned earlier. The first section is designed with two 15° dipole magnets and three quadrupole magnets to form a horizontal achromatic unit, supplemented by ten quadrupole magnets for beam envelope manipulation. In the second section, the beam is deflected 30° vertically, climbing over the 6 m slope, then deflected back by 30° to reach the same horizontal plane as the storage ring. Both ends of the second section feature two 15° dipole magnets and three quadrupole magnets forming vertical achromatic units, which adopt the same magnet parameters as the horizontal achromatic unit in the first section to reduce beam commissioning complexity. In the third section, four quadrupole magnets deliver the beam from the transfer line to the storage ring entrance. A horizontal achromatic unit consisting of two 10.5° dipole magnets, an injection septum, and three quadrupole magnets matches the Twiss parameters. [Figure 1: see original paper] plots the 3D model of the WALS BITL. [Figure 2: see original paper] shows the evolution of Twiss parameters along the WALS BITL, with the beta function not exceeding 25 m and the dispersion function not exceeding 0.6 m. [Figure 3: see original paper] presents the beam size in the BITL, where beam size refers to the standard deviation of the electron beam bunch following a Gaussian distribution.

B. Global Correction Method

During beamline construction, installation errors, manufacturing errors, and other factors cause the particle beam to deviate from the predetermined orbit, leading to Twiss parameter mismatch and even beam losses, necessitating beamline correction optimization. The global correction method is a beam correction approach applied in the WALS BITL design [12]. Considering the horizontal-vertical coupling problem of the WALS BITL, particularly for dispersion function correction, simultaneous correction in both horizontal and vertical directions is essential, prompting the initial proposal of the global correction method.

The global correction process includes two parts: orbit correction and optical correction. Orbit correction is achieved with six sets of HVC magnets and BPMs placed in pairs, with specific locations observable in [Figure 2: see original paper] and [Figure 3: see original paper]. Optical correction is realized using the AT toolbox in MATLAB, which involves randomly generating 20-50 sets of quadrupole magnet seeds to test parameters, retaining seeds with better results (such as lower maximum beta function and lower dispersion function at the beamline end), and iterating until ideal quadrupole settings are found. This computational approach enables simultaneous bidirectional beam correction in simulation. Since the quadrupole magnets of the WALS BITL are powered individually, the 18 quadrupole magnets in the non-dispersion region and the 12 quadrupole magnets in the dispersion region (30 total) can be adjusted separately. Twenty error seeds were randomly generated, and beam optics parameters and beam sizes under different error seeds were analyzed.

In global correction simulations, orbit correction was performed using six sets of HVC magnets and BPMs, while optical correction was achieved through MATLAB calculations using beam information collected along the beamline. [Figure 4: see original paper] shows the comparison before and after beam orbit correction, where x represents the horizontal orbit and y represents the vertical orbit, with curves corresponding to the 20 error seeds mentioned earlier. Comparing (a) and (b), the maximum horizontal orbit error before correction is 22 mm, with end values ranging from -18 mm to 12 mm. After correction, the maximum orbit error is reduced to 2.7 mm, with end values nearly 0 mm (ideal orbit). Similarly, comparing (c) and (d), the maximum vertical orbit error before correction is 27 mm, with end values ranging from -11 mm to 16 mm. After correction, the maximum error becomes 2.7 mm, with end values also close to 0 mm, indicating that orbit positions and directions at the injection point have been essentially corrected ideally.

For beam optics parameters, [Figure 5: see original paper] plots results before correction, while [Figure 6: see original paper] presents results after correction, with curves also corresponding to the 20 error seeds. Comparing these figures reveals that after correction, the maximum βx value at the end decreases from 17.8 m to 12.6 m. For βy, the maximum end value decreases from 12.3 m to 10.4 m. The maximum x value at the end decreases from 0.5 m to 0.01 m, and the maximum y value at the end decreases from 0.4 m to 0 m. Overall, the beta function at the end is close to 12 m, and the dispersion function is close to 0 m, meeting correction requirements of 15 m and 0.2 m, respectively.

To compare with the INN method, 20 new seeds were selected for correction using the global method. These 20 seeds differ from the previous 100 seeds [12] and have a slightly smaller error range. The global correction achieved similar results to the previous study, with particular focus on dispersion function correction, especially its deviation in the beamline middle section, resulting in better overall correction. However, correction results at the injection point of the storage ring injection straight section were very similar. While the global correction method achieves beam correction for the WALS BITL by capturing information from all points along the transport line in simulations, in reality, limited locations for beam diagnostic elements provide far less information than simulations. Consequently, the global correction method is unavailable for actual beam correction, necessitating a new commissioning method that can simultaneously solve horizontal and vertical beam corrections while completing beam correction with limited information. Therefore, we have investigated beam correction for the WALS BITL using machine learning algorithms, proposing and implementing an INN model for this purpose.

III. Invertible Neural Network Model

With continuous progress in artificial intelligence, machine learning systems have developed rapidly [42], with increasing applications in cutting-edge research [43-46], particularly in physics [47-50]. In the field of particle accelerators, machine learning technology applications in beam commissioning are also under rapid development.

A. Network Structure for the WALS BITL

Currently, machine learning algorithms are often used to design storage rings or study beam instabilities, with neural networks typically featuring one to three hidden layers and node counts ranging from tens to thousands. Multi-objective genetic algorithms [51, 52] are usually applied to find optimal working points for accelerator components. In our case, beam transport at the WALS BITL is unidirectional, with few adjustable components and a very specific three-stage beamline design, necessitating a unique neural network model for beam correction. With 30 adjustable quadrupole magnets and limited locations for beam profile monitoring systems along the BITL, the neural network must determine K-values of quadrupoles to achieve desired horizontal (σx) and vertical (σy) beam sizes at the monitoring systems.

In beam transfer lines, the forward process from magnet K-value to beam size σ is accurately defined, but the inverse process from beam size σ to magnet K-value is ambiguous, with potential for "many-to-one" solutions. Therefore, an INN model has been designed for WALS BITL beam correction. Through forward learning, the network can predict beam size from quadrupole magnet K-values, while through invertible mapping, it can also predict K-values corresponding to expected beam sizes. The network thus realizes both forward prediction from magnet K-value to beam size and inverse prediction from beam size to magnet K-value, with the latter based on the former. The forward learning process aligns with the physical nature of beam dynamics along the WALS BITL, and the inverse prediction results are highly accurate, making this method efficient for actual beam commissioning.

Beam size is selected as a learning feature for several reasons: (1) it can be easily collected by beam profiles; (2) ensuring efficient electron transport without loss makes beam size the primary commissioning goal; (3) since the storage ring has strict dispersion requirements when injecting electron beams, the dispersion function is also an important correction target. Beta function and other optical parameters must be considered, as beam size is determined by both beta and dispersion functions, as shown in Equation 1: $\sigma\mathrm{i}^2 = \beta\mathrm{i}\ \mathrm{i} + (\ \mathrm{i}\sigma\mathrm{e})^2$, where σi is beam size, βi is the beta function, i is beam emittance, i is the dispersion function, and σe is energy spread (i refers to horizontal or vertical). This equation shows that beam size changes with beta and dispersion functions when beam emittance and energy spread are constant.

The Adam optimization algorithm is introduced as an optimizer, differing from common stochastic gradient descent by combining adaptive learning rates with momentum to efficiently and automatically adjust learning rates and parameter update speeds during training. The Back-propagation algorithm is chosen to compute gradient functions. Unlike forward propagation, which transmits input data to the output layer, back-propagation applies the chain rule to find partial derivatives of all parameters, passing them backward and continuously adjusting network weights and biases to better fit data and improve model generalization [39].

While machine learning techniques typically employ Artificial Neural Networks (ANN) in accelerator physics, the unique characteristics of the WALS BITL warrant the INN model. Since electron beam motion in the transfer line is unidirectional, multiple quadrupole magnet combinations may correspond to the same beam size. Therefore, the primary goal is to quickly obtain a feasible K-value combination based on desired beam size targets. The INN model is highly suitable, as its forward learning process can simulate beam transport accurately, while its inverse transmission process can rapidly obtain feasible K-value combinations.

B. Beam Correction Process

The overall INN model learning process is shown in [Figure 7: see original paper], comprising five main steps from (a) to (c). In Figure 7: see original paper: Step 1 uses the AT toolbox of MATLAB to model the transfer line, randomly generating injection parameter errors from the LINAC and magnet installation errors within specified ranges to create error seeds. Step 2 corrects beam orbit with HVC magnets in simulation (in practice, orbit correction can be completed without machine learning). Step 3 randomly generates multiple sets of quadrupole magnet K-values for each error seed, collecting horizontal (σx) and vertical (σy) beam sizes simulated by AT (or measured during actual commissioning) at each profile. K-values and beam sizes are treated as input x and output y, respectively, forming a learning dataset (training:validation:test ratio of 4:1:1.25).

In Figure 7: see original paper, Step 4 preliminarily trains the model with error seed No.1, continuously adjusting network structure and hyperparameters to design the most reasonable model. Step 5, shown in Figure 7: see original paper, verifies model usability: the dataset from error seeds is learned and predicted, and if results are reasonable, the error seed is replaced by the next one. If unreasonable, hyperparameters are reset until ten error seeds can be reasonably predicted. The INN model is then considered to have reasonable structure and sufficient usability, enabling beam transfer line correction in simulation. R^2 (coefficient of determination) is selected as the criterion for judging prediction rationality, with predictions considered accurate enough when the average forward prediction R^2 exceeds 0.90 and the average inverse prediction R^2 exceeds 0.85. R^2 is described in Section V C.

C. Optimization of Beam Profile Monitoring System Settings

Approximately ten locations are available for beam profile monitoring system installation in the WALS BITL, designated No.1 to No.10 from beamline start to end. The specific distribution is indicated in [Figure 1: see original paper] using arrows. To save space and cost, the number of beam profiles should be minimized while ensuring beam correction accuracy. Optimization has been performed using the ideal error-free WALS BITL, starting from Step 3 in Figure 7: see original paper. illustrates R^2 values for different profile numbers, beginning with ten systems and continuously reducing the number for comparison.

Since the model uses the Adam optimizer, some fluctuations in R^2 values occur. As shown in , when the number of profiles is reduced to five, the inverse prediction R^2 value drops sharply. Fewer profiles provide fewer observable points and less data for model learning. Considering practical model requirements and device costs, the final profile number is determined to be six. Since the final model application involves predicting quadrupole magnet K-values based on desired beam size, a beam profile monitoring system must be placed at position No.10 behind the last quadrupole magnet. Because positions No.1, No.2, No.3, and No.4 are close to each other with few nearby quadrupole magnets, one or two representatives should suffice. Based on this, comparisons have been conducted among various combinations of six profile locations, as listed in . Finally, positions No.1, No.4, No.5, No.6, No.8, and No.10 have been selected to achieve satisfactory R^2 values.

D. Preparation of Datasets

The main considered errors for the WALS BITL include injection errors at the BITL start point (e.g., Twiss parameter errors and position errors), magnet installation errors, and magnetic field errors of dipoles and quadrupoles. Using the AT toolbox of MATLAB, ten corresponding error seeds are randomly generated. For each error seed, a dataset is produced for the neural network model, where input x consists of quadrupole magnet K-values and output y consists of horizontal and vertical beam sizes at the six beam profile monitoring systems. Each dataset contains 10,000 groups of x and y.

Error ranges are listed in Tables 3-6. Specifically, Tables 3 and 4 show positioning and Twiss parameter error ranges of the injected beam from the LINAC. shows installation error ranges for dipole and quadrupole magnets. lists magnetic field error ranges for dipoles and quadrupoles. The ten random error seeds are generated within these ranges, and for each seed, the dataset containing quadrupole K-values is produced based on values. Ten random error seeds have been selected to verify whether the INN model structure and hyperparameter settings are appropriate. If the model can successfully correct the beam under ten different error seeds, it demonstrates good versatility.

E. Design of the Network Structure

After comprehensive consideration, an INN model with eight affine coupling modules has been designed, applying the ReLU activation function and incorporating the Adam algorithm and Back-propagation algorithm. [Figure 8: see original paper] shows the schematic diagram of the neural network structure.

F. Training Results

Prediction results for the test set are visually analyzed by comparing predicted values with true beam size values for both forward and inverse processes. Box plots [53] are applied for visual analysis, clearly displaying data quartiles and realistically reflecting data distribution. As shown in [Figure 9: see original paper], for beam size, the median error between forward, inverse, and true values does not exceed 0.015 mm, while 75% and 25% of the data, along with maximum and minimum values, are almost symmetrically distributed around the median.

For forward prediction, $ $3_y results are about 0.015 mm larger than inverse and true values overall. However, since the model is eventually used with inverse prediction, the difference between the inverse process and true value is of greater concern. As evident, the difference between inverse predicted value and true value distributions is negligible at each beam profile monitoring system.

[Figure 10: see original paper] compares predicted and true beam size values at the first and last beam profile monitoring locations, with each point representing an individual in the test set. This figure transparently and informatively reveals minor deviations between true and predicted values.

IV. Verification

Since the WALS BITL has not yet been built, MATLAB has been used to model it for virtual verification. [Figure 11: see original paper] compares beam parameters before and after correction using the INN model with an error seed and $±$1% quadrupole field error. Continuous lines represent dispersion function and beam size under these errors, while dark blue dotted lines marked with circles show corrected beam parameters using the INN model. Both dispersion function and beam size are well corrected, with horizontal and vertical dispersion at the transfer line exit being 0.03 m, horizontal beam size 0.14 mm, and vertical beam size 0.12 mm.

compares beam sizes between expected and corrected values using INN. Input values refer to expected beam sizes imported into the INN model, which directly predicts K-values. Corrected beam sizes are calculated by MATLAB using the WALS BITL with INN-predicted K-values, showing results very close to expected values. [Figure 11: see original paper] and further verify INN model accuracy.

[Figure 12: see original paper] shows beam dispersion function and beam size after INN model correction for ten error seeds. Segmented blue curves represent theoretical ideal beam parameters along the transfer line without errors, while continuous color curves show corrected results under ten error seeds. The corrected dispersion function and beam size are close to ideal conditions, with dispersion at the transfer line exit within 0.15 m and beam size corrected to within 0.3 mm.

Comparing [Figure 12: see original paper] and [Figure 6: see original paper] reveals that in the x-direction, the global correction method corrects the dispersion function at the endpoint within 0.01 m, while the INN model corrects it to within 0.1 m. In the y-direction, the global correction method corrects the dispersion function at the endpoint close to 0 m, while the INN model corrects it to within 0.15 m. Both methods meet the 0.2 m correction requirement. Although the global correction method achieves more precise results in simulation environments, it is limited to simulations because it requires more beam information than actual diagnostic systems can provide. In contrast, the INN model is specifically designed for practical beam correction applications and can effectively utilize limited diagnostic information. This practical capability makes the INN model more advantageous in actual operations, especially when beam information is limited. Therefore, despite the global correction method' s excellent simulation performance, the INN model' s feasibility and flexibility in practical applications make it a more practical choice.

V. Method Analysis

The INN model code programming is based on previous work [54]. This section analyzes the method applied to the WALS BITL. Section V A introduces dynamic relationships in the WALS BITL and analyzes mathematical relationships of the INN model application. Section V B details loss function considerations, with summarizing hyperparameters used in the final model. Section V C introduces R^2 and model evaluation steps.

A. Dynamic Relationships

The dynamic equations for particle beam and quadrupole magnet K-values at the WALS injection transfer line are given by $\hat{f}$: D_K $\times$ D $\rightarrow$ D_σ, where $\hat{f}$(k, d) = $\hat{\sigma}$(d). Here K represents quadrupole magnet values, d D represents beamline location, function $\hat{f}$ represents the transfer line model, and $\hat{\sigma}$ represents electron beam size at location d.

INN models generally require equal input and output data dimensions. However, for this problem, input data dimension is 30 while output data dimension is 12, necessitating dimension compensation. We define total data dimension as A. Since inverse problem solutions are not unique and may involve many-to-one mappings, the solution maps not only to pre-designed values but also to a prospective spatial distribution (I distribution) learned during training. By randomly sampling in I, the model continuously seeks the most suitable distribution value to complete the inverse problem' s optimal solution [54]. The total dimension A can be described as A = dim(K) + 1 + dim(xpad) = dim(σ) + dim(ypad) + I, where xpad is the complementary dimension of input data and ypad is the complementary dimension of output data. I is the potential spatial distribution dimension. For this paper, only output data dimension is supplemented, so the forward prediction mathematical expression for the transfer line physical model is f̂(k, d) = ($\hat{\sigma}$(d), dim(ypad), I). The inverse prediction mathematical expression is $\hat{f}^{-1}$: D_σ × D_{dim}(ypad) × D_I → D_K × D, where $\hat{f}^{-1}$($\hat{\sigma}$(d), dim(ypad), I) = (k, d).

B. Loss Function

The loss function definition during training consists of five components: mean square error L_{MSE} between predicted output and true values, mean square error L_α of output data complementary dimensions, reconstruction error L_r to reduce influence of small perturbations (δ) on reverse prediction, L_k to ensure sampled magnet K-values follow the dataset distribution, and L_I to ensure potential spatial distribution D_I follows the desired distribution. The total loss function is L_{all} = ω{*MSE*} *L*{MSE} + ω*r L_r +* $\omega\alpha$ L_α + ω_k L_k + ω*I L_I, where* ω{MSE}, ω*r,* $\omega\alpha$, ω_k, and ω_I are the corresponding error weights.

The mean square error and reconstruction error equations are: MSE = (1/n) Σ(y_i - ŷ_i)^2 L_r = ||$\hat{f}^{-1}$(f̂(k_i, d_i) + δ) - (k_i, d_i)||^2

In these equations, n is the total number of quadrupole magnets per sample (i.e., 30). shows other detailed data configurations for the INN model.

C. Evaluation Indicators

At the transfer line end, beam size is determined by all quadrupole magnets in the beamline (ignoring dipole influences), so in inverse prediction, simply comparing predicted and true values cannot truly reflect model accuracy. Therefore, R^2 has been selected as the evaluation metric. R^2 ranges between 0 and 1, with values closer to 1 indicating higher prediction accuracy. R^2 is expressed as: R^2 = 1 - [Σ(ŷ_i - y_i)^2] / [Σ(ȳ_i - y_i)^2] where y_i is the true value, ȳ_i is the true value average, and ŷ_i is the predicted value.

Model evaluation steps are: (1) The model makes forward predictions on the validation set, obtaining forward prediction σ values corresponding to quadrupole magnet K-values (true K-values). (2) True σ values and forward predicted σ values calculate forward prediction R^2. (3) The model makes inverse predictions on the validation set, obtaining quadrupole magnet K-values predicted from true σ values. (4) The model makes forward predictions on these predicted K-values, obtaining inverse prediction σ values. (5) True σ values and inverse prediction σ values calculate inverse prediction R^2. [Figure 13: see original paper] illustrates these specific steps, where "Inverse prediction-σ obtained by the Pending K-value after the forward prediction" represents the predicted σ value for the entire inverse prediction process.

After adequate testing, R^2 can reach approximately 0.96 maximum for forward prediction and about 0.93 for inverse prediction. From both corrected Twiss parameters and beam sizes, and from R^2 values, the feasibility of this INN model for WALS BITL beam correction study is confirmed. lists R^2 data for an error seed, showing average forward prediction R^2 reaching 0.9483 and inverse prediction R^2 reaching 0.9228.

VI. Conclusions

An INN model has been designed using machine learning neural network algorithms to optimize beam correction for the WALS BITL. The method has proven feasible with remarkable beam correction effects in simulations. Given the unique "three-stage"design of the WALS BITL, the neural network algorithm can simultaneously solve horizontal and vertical beam correction problems while addressing beam correction difficulties caused by limited beam information at restricted measurement points.

The final INN model design realizes both forward prediction from magnet K-value to beam size σ and inverse prediction from beam size σ to quadrupole magnet K-value. The AT toolbox of MATLAB has been used to model the transfer line, collect beam-related data, verify INN model inverse prediction results, and successfully achieve commissioning optimization for the WALS BITL in simulation. This method can provide visual guidance for actual beam correction processes. In the future, we will further explore designing a proper INN model with fewer beam profile monitoring systems to reduce construction costs. During WALS BITL construction and commissioning, this method will be applied in practice. This study of machine learning algorithms for WALS BITL beam correction may be applied to similar accelerator physics problems at other facilities.

Acknowledgments: This work was supported by the Major Science and Technology Project of Hubei Province (2021AFB001).

Author Contributions: All authors contributed to study conception and design. Material preparation, data collection, and analysis were performed by Hui Wang, Hao-Hu Li, Jiao-Bao Guan, and Yuan-Cun Nie. The first draft was written by Hui Wang, Hao-Hu Li, and Yuan-Cun Nie reviewed and edited the manuscript. All authors commented on previous versions and approved the final manuscript.

Data Availability: Data supporting this study' s findings are available from the corresponding author upon reasonable request.

Declarations: The authors declare no conflict of interest.

References

[1] H.N. Chapman, Fourth-generation light sources. IUCrJ 10, 246 (2003). doi: 10.1107/S2052252523003585
[2] P.F. Tavares, E. Al-Dmour, A. Andersson, et al., Commissioning and first-year operational results of the MAX IV 3 GeV ring. J. Synchrotron Rad. 25, 1291 (2018). doi: 10.1107/S1600577518008111
[3] Y. Jiao, Z.H. Bai, X. Li, Accelerator physics and technology of the fourth-generation synchrotron radiation light source. PHYSICS 53, 71 (2024). doi:10.7693/wl20240201 (in Chinese)
[4] L. Valle, T. Brochard, N. Carmignani, et al., Off-energy operation for the Extremely Brilliant Source at the European Synchrotron Radiation Facility. Phys. Rev. Accel. Beams 27, 051601 (2024). doi:10.1103/PhysRevAccelBeams.27.051601
[5] J. Yi, X. Gang, X.H. Cui, The HEPS project. doi:10.1107/S1600577518012110
[6] Z.H. Bai, G.W. Liu, T.L. He, et al., A Modified Hybrid 6BA Lattice for the HALF Storage Ring. Paper Presented at the IPAC 2021, Campinas, SP, Brazil, August 2021, pp. 407-409. doi:10.18429/JACOW-IPAC2021-MOPAB112
[7] F.C. Li, R.Z. Liu, W.J. Li, et al., Synchrotron radiation: A key tool for drug discovery. Bioorg. Med. Chem. Lett. 144, 129990 (2024). doi: 10.1016/j.bmcl.2024.129990
[8] Z.Y. Dai, Y.C. Nie, Z. Hui, et al., Design of S-band photoinjector with high bunch charge and low emittance based on multi-objective genetic algorithm. Nucl. Sci. Tech. 34, 41 (2023). doi:10.1007/s41365-023-01183-6
[9] H.H. Li, J. Wang, L. Tang, et al., Project of Wuhan photon source. Paper Presented at the IPAC 2021, Campinas, SP, Brazil, August 2021, pp. 346-349. doi:10.18429/JACOW-IPAC2021-MOPAB092
[10] Y. Zou, H.H. Li, Y. Chen, et al., Studies on beam injection system for Wuhan advanced light source storage ring. Paper Presented at the IPAC 2023, Venice, Italy, September 2023, pp. 1176-1179. doi:10.18429/JACoW-IPAC2023-MOPM086
[11] Z.Y. Dai, Y.C. Nie, J.H. Zhong, et al., Beam dynamics study of the photoinjector at Wuhan advanced light source. Radiation Detection Technology and Methods 8, 1319 (2024). doi:10.1007/s41605-024-00455-y
[12] H.H. Li, J. Li, G. Wei, et al., Beam transfer line of Wuhan advanced light source. Paper Presented at the IPAC 2023, Venice, Italy, May 2023, pp. 1082-1084. doi:10.18429/JACoW-IPAC2023-MOPM040
[13] A.G. Akritas, G.I. Malaschonok, Applications of singular-value decomposition (SVD). Math. Comput. Simulat. 67, 15 (2004). doi:10.1016/j.matcom.2004.05.005
[14] S.H. Mirza, R. Singh, P. Forck, et al., Closed orbit correction at synchrotrons for symmetric and near-symmetric lattices. Phys. Rev. Accel. Beams 22, 072804 (2019). doi:10.1103/PhysRevAccelBeams.22.072804
[15] A.M. Coxe, J.F. Benesch, R.M. Bodenstein, et al., Beam correction for multi-pass arcs in FFA @ CEBAF: status update. Paper Presented at the

IPAC 2024, Nashville, TN, July 2024, pp.1054-1056. doi:10.18429/JACoW-IPAC2024-TUPC23
[16] H. Jin, D.O. Jeon, J.H. Jang, Results of on-line orbit correction at the medium energy beam transport section of the RAON accelerator. J. Korean. Phys. Soc. 83, 854 (2023). doi:10.1007/s40042-023-00919-2
[17] J. Arthur, G. Materlik, R. Tatchyn, et al., The LCLS: A fourth generation light source using the SLAC linac. Rev. Sci. Instrum. 66, 1987 (1995). doi:10.1063/1.1145778
[18] J.N. Galayda, The LCLS-II: A high power upgrade to the LCLS. Paper Presented at the IPAC 2018, Vancouver, BC, Canada, June 2018, pp. 18-23. doi:10.18429/JACoW-IPAC2018-MOYGB2
[19] C. Emma, A. Edelen, M.J. Hogan, et al., Machine learning-based longitudinal phase space prediction of particle accelerators. Phys. Rev. ST Accel. Beams. 22, 112802 (2018). doi:10.1103/PhysRevAccelBeams.21.112802
[20] H.J. Xu, Z.T. Zhao, Current status and progresses of SSRF project. Nucl. Sci. Tech. 19, 1 (2018). doi:10.1016/S1001-8042(08)60013-5
[21] T.Y. Yang, W. Wen, G.Z. Yin, et al., Introduction of the X-ray diffraction beamline of SSRF. Nucl. Sci. Tech. 26, 020101 (2015). doi:10.13538/j.1001-8042/nst.26.020101
[22] Z.L. Li, Y.C. Fan, L. Xue, et al., The design of the test beamline at SSRF. AIP Conf. Proc. 2054, 060040 (2019). doi:10.1063/1.5084671
[23] J.H. He, X.Y. Gao, Status of the crystallography beamlines at SSRF. Eur. Phys. J. Plus. 130, 1 (2015). doi:10.1140/epjp/i2015-15032-6
[24] R.C. Li, Q.L. Zhang, Q.R. Mi, et al., Application of machine learning in orbital correction of storage ring. High Power Laser and Particle Beams 33, 034007 (2021). doi:10.11884/HPLPB202133.200318 (in Chinese)
[25] T. J. Kattenborn, J. Leitloff, F. Schiefer, et al., Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. ISPRS. J. Photogramm. 173, 24 (2021). doi:10.1016/j.isprsjprs.2020.12.010
[26] L. Alzubaidi, J.L. Zhang, A.J. Humaidi, et al., Review of deep learning: concepts, CNN architectures, applications, challenges, future directions. J. Big Data. 8, 1 (2021). doi:10.1186/s40537-021-00444-8
[27] W.C. Zhu, Z.Y. Wei, C.J. Xie, et al., Development of the NFTHz accelerator beam profile measurement system. High Power Laser and Particle Beams. 36, 034004 (2024). doi:10.11884/HPLPB202436.230361 (in Chinese)
[28] H. Zou, The adaptive lasso and its oracle properties. J. Am. Stat. Assoc. 101:1418 (2006). doi:10.1198/016214506000000735
[29] J. Ranstam, J.A. Cook, LASSO regression. Brit. J. Surg. 105, 1348 (2018). doi:10.1002/bjs.10895
[30] Y.B. Yu, W.B. Ni, G.F. Liu, et al., Initial Application of Machine Learning for Beam Parameter Optimization at the Hefei Light Source II. J. Phys. Conf. Ser. 2687, 072002 (2024). doi:10.1088/1742-6596/2687/7/072002
[31] Z.M. Chu, D.J. Xiao, Y.S. Qiao, et al., Machine Learning Applications for Particle Accelerators. Front. Data. Comput. 1, 110 (2019). doi:10.11871/jfdc.issn.2096-742X.2019.02.010
[32] Y.B. Yu, G.F. Liu, W. Xu, et al., Research on tune feedback of the Hefei

Light Source II based on machine learning. Nucl. Sci. Tech. 33, 28 (2022). doi:10.1007/s41365-022-01018-w
[33] L. Ardizzone, J. Kruse, S.J. Wirkert, et al., Analyzing inverse problems with invertible neural networks. Paper Presented at the ICLR 2019, Ernest N. Morial Convention Center, February 2019. doi:10.48550/arXiv.1808.04730
[34] O. Moran, P. Caramazza, D. Faccio, et al., Deep, complex, invertible networks for inversion of transmission effects in multimode optical fibres. Paper Presented at NeurIPS 2018, Montreal, Canada, December 2018. doi:10.5555/3327144.3327248
[35] D.P. Kingma, P. Dhariwal, Glow: Generative flow with invertible 1x1 convolutions. Paper Presented at NeurIPS 2018, Montreal, Canada, December 2018, pp. 10236-10245. doi:10.5555/3327546.3327685
[36] D. Yarotsky, Error bounds for approximations with deep ReLU networks. Neural networks. 94, 103 (2017). doi:10.1016/j.neunet.2017.07.002
[37] J.C. He, L. Lin, J.C. Xu, et al., ReLU deep neural networks and linear finite elements. J. Comput. Math. 38, 502 (2020). doi:10.4208/jcm.1901-m2018-0160
[38] D.P. Kingma, J. Ba, Adam: A method for stochastic optimization. Paper Presented at ICLR 2015, San Diego, May 2015. doi:10.48550/arXiv.1412.6980
[39] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by back-propagating errors. Nature. 323, 533 (1986). doi:10.1038/323533a0
[40] J. Li, J.H. Cheng, J.Y. Shi, et al., Brief introduction of back propagation (BP) neural network algorithm and its improvement. Advances in Computer Science and Information Engineering: Volume 2. (Springer, Berlin Heidelberg, 2012), pp. 553-558
[41] S. Walston, S. Boogert, C. Chung, et al., Performance of a high resolution cavity beam position monitor system. Nucl. Instrum. Meth. A. 578, 1 (2007). doi:10.1016/j.nima.2007.04.162
[42] Y.C. Wu, J.W. Feng, Development and application of artificial neural network. Wirel. Pers. Commun. 102, 1645 (2018). doi:10.1007/s11277-017-5224-x
[43] B.C. Wang, C.X. Tang, M.T. Qiu, et al., A machine learning approach to TCAD model calibration for MOSFET. Nucl. Sci. Tech. 34, 192 (2023). doi:10.1007/s41365-023-01340-x
[44] Y.J. Ma, Y. Ren, P. Feng, et al., Sinogram denoising via attention residual dense convolutional neural network for low-dose computed tomography. Nucl. Sci. Tech. 32, 41 (2021). doi:10.1007/s41365-021-00874-2
[45] X.D. Guo, P. He, X.J. Lv, et al., Material decomposition of spectral CT images via attention-based global convolutional generative adversarial network. Nucl. Sci. Tech. 34, 45 (2023). doi:10.1007/s41365-023-01184-5
[46] G.X. Wei, S.X. Zhang, Z. Li, et al., Multi-modality measurement and comprehensive analysis of hepatocellular carcinoma using synchrotron-based microscopy and spectroscopy. Nucl. Sci. Tech. 32, 102 (2021). doi:10.1007/s41365-021-00927-6
[47] H.Y. Yang, D. Wang, D.D. Wang, An overview of the application of neural network algorithm in the nuclear field of China. Nucl. Sci. Tech. 8, 19 (2020). doi:10.12677/NST.2020.81003 (in Chinese)
[48] W.B. He, Y.G. Ma, L.G. Pang, et al., High-energy nuclear physics meets

machine learning. Nucl. Sci. Tech. 34, 88 (2023). doi:10.1007/s41365-023-01233-z
[49] Z.P. Gao, Y.J. Wang, H.L. Lü, et al., Machine learning the nuclear mass. Nucl. Sci. Tech. 32, 109 (2021). doi:10.1007/s41365-021-00956-1
[50] Y. Zou, Q.Z. Xing, B.C. Wang, et al., Application of the asynchronous advantage actor-critic machine learning algorithm to real-time accelerator tuning. Nucl. Sci. Tech. 30, 158 (2019). doi:10.1007/s41365-019-0668-1
[51] K. Deb, A. Pratap, S. Agarwal, et al., A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans. Evol. Comput. 6, 182 (2002). doi:10.1109/4235.996017
[52] H.X. Yin, J.B. Guan, S.Q. Tian, et al., Design and optimization of diffraction-limited storage ring lattices based on many-objective evolutionary algorithms. Nucl. Sci. Tech. 34, 147 (2023). doi:10.1007/s41365-023-01284-2
[53] R. McGill, J.W. Tukey, W.A. Larsen. Variations of box plots. Am. Stat. 32, 12 (1978). doi:10.1080/00031305.1978.10479236
[54] R. Bellotti, R. Boiger, A. Adelmann, Fast, efficient and flexible particle accelerator optimisation using densely connected and invertible neural networks. Information 12, 351 (2021). doi:10.3390/info12090351