ISBN: 978-3-95450-249-3 ISSN: 2673-5350 doi: 10.18429/JACoW-IBIC2024-TUP15

# NEURAL NETWORK TECHNIQUE FOR IMPROVING ACCURACY, RELIABILITY AND ROBUSTNESS OF BEAM POSITION MONITOR SYSTEM

Fang-Qi Huang*[1], Tao-Guang Xu[1], Yan-Feng Sui†[1,2], Jun He[1]
[1]Institute of High Energy Physics, Chinese Academy of Sciences, Beijing, China
[2]University of Chinese Academy of Sciences, Beijing, China

*Abstract*

The beam position monitor (BPM) is a crucial instrumentation system for the commissioning and operation of the accelerator. Its accuracy and robustness are essential for ensuring the stability of the accelerator. Currently, the beam position is calculated by fitting a polynomial to the four voltage signals obtained from the BPM electrodes in BEPCII and HEPS. To improve the system's robustness, a formula is provided that expresses the relationship between the three voltage signals and the position. The average fitting error is 40 $\mu$m, but the error of the three-electrode calculation is not high. Therefore, we propose using neural networks for beam position calculation to improve the system's robustness while guaranteeing its accuracy. This will ensure that the beam position can be provided stably, even in the case of one single electrode error. In our experiments, we use BPM calibration data from HEPS. The trained neural network's performance on the test set meets the accuracy requirements, with an error of less than 15 $\mu$m in both four-electrode and three-electrode predictions, and an average value of fitting error is 1 $\mu$m. Furthermore, we validate the neural network's generalization ability by using data measured by BPM on HEPS.

## INSTRUCTION

Beam Position Measurement (BPM) is a crucial instrumentation system in accelerators, serving as a vital reference for beam commissioning and operation, as well as a fundamental data source for physics research. In HEPS, there are over 700 BPMs used for beam orbit measurement and correction, beam current measurement, beam loss diagnosis and analysis, and other beam state monitoring[1]. The accuracy and robustness of the beam position measurement are essential for ensuring the stable operation of the accelerator. Currently, the beam position is calculated indirectly. Voltage signals measured from four electrodes on BPM, which are then calculated using a formula. The value corrected called beam position and reported to the database. Moreover, there is a set of three-electrode calculations to address the issue of single-electrode signal loss. The existing four-electrode polynomial fitting has an error of approximately 40 $\mu$m[2], but the three-electrode calculation does not meet the required level of accuracy. Therefore, we propose using neural networks for beam position calculation to enhance the robustness and ensure the accuracy. This approach will guarantee that the beam position can be provided reliably, even in the case of single-electrode error.

## TRADITIONAL BEAM POSITION CALCULATION METHOD

The four electrodes of the BPM measure the voltage as the beam passes over the BPM. The normalized horizontal and vertical beam positions are calculated using Eq. (1)[2].

$$\begin{aligned} X_{\text{norm}} &= \frac{V_1 - V_2 - V_3 + V_4}{\sum_{i=1}^{4} V_i} \\ Y_{\text{norm}} &= \frac{V_1 + V_2 - V_3 - V_4}{\sum_{i=1}^{4} V_i} \end{aligned} \tag{1}$$

where Vi denotes the voltage measured by electrode i. The final beam positions X, Y are then obtained by polynomial fitting approximation as in Eq. (2).

$$\begin{aligned} X &= \sum_{i=0}^{n} \sum_{j=0}^{i} A_{i-j,j} X_{\text{norm}}^{i-j} Y_{\text{norm}}^{j} \\ Y &= \sum_{i=0}^{n} \sum_{j=0}^{i} B_{i-j,j} X_{\text{norm}}^{i-j} Y_{\text{norm}}^{j}, \end{aligned} \tag{2}$$

where n is the highest order of the polynomial fit. Based on the calibration data, the values of a, b were obtained by fitting. When n=6, the average value of the fitting error is 40 $\mu$m[2]. When one of the electrodes is out of order, the beam position can be obtained by using deviation[2], but the result is not satisfactory.

## BP NEURAL NETWORK CALCULATION METHOD

The BP neural network is composed of neurons, including input, output, and activation functions. The structure of the neural network is divided into three layers: input, hidden, and output. The activation function allows for non-linear mapping of input data, and the number of layers in the hidden layer determines the complexity of the function that can be modeled. Theoretically, two layers of hidden layers are sufficient to fit any bounded continuous function[3]. When training the network, several factors must be considered, such as the neural network structure (including the choice of activation function, number of hidden layers, and number of neurons per layer), data processing (including the selection

* huangfq@ihep.ac.cn
† syf@ihep.ac.cn

ISBN: 978-3-95450-249-3 ISSN: 2673-5350 doi: 10.18429/JACoW-IBIC2024-TUP15



of a training set and normalization of inputs), and hyper-parameter tuning (including the learning rate, loss function, number of iterations, and optimizer).

### *Neural Network Structure Adjustment*

The structure of a neural network is composed of three main components: the activation function, the number of hidden layers, and the number of neurons in each hidden layer. This can be seen in Fig. 1. The structure of a neural network plays a crucial role in determining the complexity of the function that the network is able to fit. The activation function is responsible for mapping the input data to the output data, and a nonlinear activation function is able to perform nonlinear mappings. Common nonlinear activation functions include sigmoid, leru, tanh, and others. In our experiments, we have used the sigmoid function, as shown in Fig. 2. It is worth noting that two layers of hidden layers are theoretically sufficient to simulate any bounded continuous function, and using three layers of hidden layers requires a much smaller number of parameters compared to using only two layers[3]. The number of neurons in each layer also has a significant impact on the size and performance of the neural network. Increasing the number of neurons in each layer can accelerate the learning speed and improve the accuracy of the network to a certain extent.

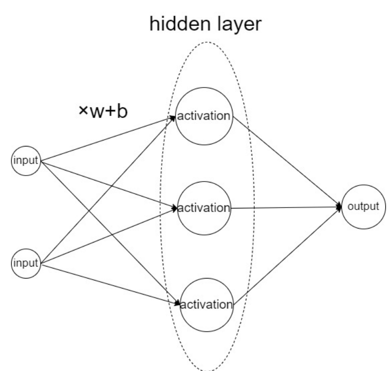


Figure 1: Neural network structure.

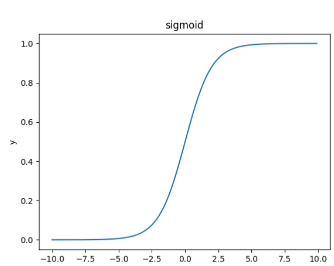


Figure 2: Sigmoid function.

### *Data Processing: Training Set Selection, Normalization*

Once the structure of the neural network has been determined, the next step is to train it. The quality of the data set used for training can greatly impact the learning effectiveness of the neural network. The Institute of High Energy Physics of the Chinese Academy of Sciences has accumulated significant data during the design, manufacturing, and use of the BPM. The calibrated data set is used to train the neural network. The data set was divided into two distinct subsets: one comprising 70% of the data was used for training purposes, while the remaining 30% was designated as the test set. The data set includes the voltages generated by the the BPM, as well as the corresponding beam current positions (x, y). However, these voltage values are quite large, which can hinder the neural network's learning efficiency when used as inputs. Additionally, the measured voltage values of the electrodes are influenced by the strength of the beam current. Therefore, it is necessary to normalize the input values to improve the neural network's training efficiency and enhance its generalization ability. Common normalization methods include Min-Max, Z-Score, Log, Atan, etc. In our experiments, we selected the most suitable normalization method for this problem. Eq. (3) is used when there are four voltages inputs, and Eq. (4) is used when there are three voltages inputs. $V_0$ is one of the three electrodes selected from the four electrodes.

$$V_i = \frac{V_i}{\sum_{i=1}^{4} V_i} \tag{3}$$

$$V_i = -\ln \frac{V_i}{V_0} + 1 \tag{4}$$

### *Hyper-parameters*

The selection of hyper-parameters has a significant impact on both the efficiency of training and the performance improvement of the neural network. These hyper-parameters include the loss function, learning rate, number of iterations, and optimizer, among others. The loss function plays a crucial role in guiding the model training process. According to the theory of maximum likelihood estimation, the model trained using a specific loss function is considered optimal (see Eq. (5)). The neural network computation process can be represented by Eq. (6), and the learning process involves adjusting the values of w and b to reach the desired loss function value $\epsilon$. The number of iterations is another important factor that can affect the training of the neural network. If the number of iterations is too high and the desired loss function value is not reached, it can lead to over-fitting, which means the model performs well on the training set but does not generalize well to test set. Therefore, it is crucial to carefully select the appropriate hyper-parameters to ensure the generalization ability of the neural network.

$$Loss = \frac{\|Y - \bar{y}\|}{2} < \varepsilon \tag{5}$$

$$Y = \sum_{i=0}^{n} Net(w_i \times y_i + b_i) \tag{6}$$

Where n represents the number of hidden layers and Net represents the activation function. The learning rate is a

Proc. 13th International Beam Instrumentation Conference,Beijing JACoW Publishing
ISBN: 978-3-95450-249-3 ISSN: 2673-5350 doi: 10.18429/JACoW-IBIC2024-TUP15

crucial factor in adjusting the parameters w and b, as it directly impacts the efficiency and effectiveness of the neural network's learning process. The optimizer is responsible for determining the direction and value of the learning rate. Commonly used optimizer include SGD, Adam, and LM. In this problem, we utilize the LM algorithm to obtain the optimal values for w and b that will minimize the loss function and reach the desired target[4].

## ANALYSIS OF THE POSSIBILITY OF THREE-ELECTRODE AND TWO-ELECTRODE FOR THE CALCULATION

Using the principle of calculating the beam position using voltages, it is possible to determine the beam position with three electrodes or with two neighbouring electrodes. The theoretical analysis considers the electrode position as the center of a circle, with the distance from the beam position to the electrodes equivalent to the voltage measured by the electrodes and used as the radius. When using three electrodes as the center of the circle and the voltage as the radius, the circle will intersect at a unique point, allowing for the calculation of the beam current position. However, when using two electrodes as the center and the voltage as the radius, the two circles will intersect at two points. In this case, the beam current position should fall within a quadrilateral with four electrodes as vertices. While theoretically one of the two intersection points should fall within the range, experimental results have shown that this method is not applicable to real beam current situations.

## BP NEURAL NETWORK CALCULATION METHOD

### *Effect of Activation Function on Training Results and Efficiency*

The activation function introduces a nonlinear mapping to the network. In our experiments, we found that the choice of the nonlinear activation function does not significantly impact the training efficiency and accuracy of the results, as shown in Fig. 3. We used the sigmoid function in our experiments. It is worth noting that the sigmoid function should not be used in more than three consecutive hidden layers, as it may cause the gradient to disappear.

### *Effect of the Number of Hidden Layers on Efficiency and Results*

In order to conserve resources and guarantee precision, we examined the impact of the number of concealed layers on the outcomes. The experimental results demonstrate that the impact of three and two hidden layers on the precision of the results is not significant when the four voltages input is employed to determine the beam position, as illustrated in Fig. 4. Accordingly, two hidden layers were employed for the four-electrode. For the three voltages input, the three hidden layers yielded superior results compared to the two

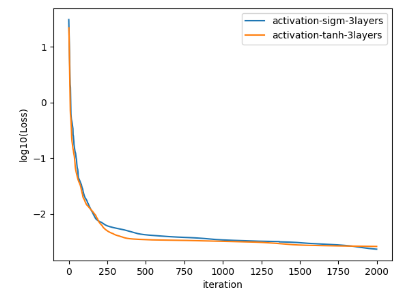


Figure 3: Loss curves for different activation functions.

hidden layers for the same number of iterations, as illustrated in Fig. 5. Consequently, three hidden layers were utilized for the three-electrode.

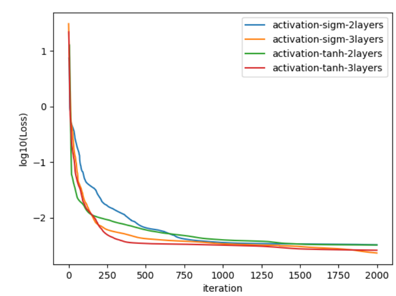


Figure 4: Loss curves of three hidden layers and four hidden layers with four voltage inputs.

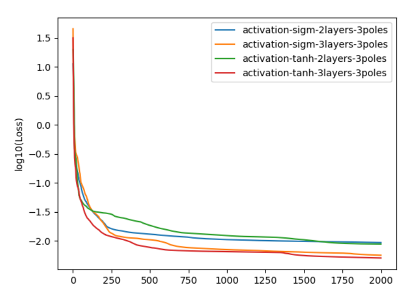


Figure 5: Loss curves of three hidden layers and four hidden layers with three voltage inputs.

### *Normalization*

In the experiments, it was found that different normalization methods greatly influence the results of the three-electrode. When using $V_i = V_i/V_0$ or Eq. (3) as a normalization method, the results are excellent on the learning set, as shown in Fig. 6, but poor on the validation set, as shown in Fig. 7. However, when using Eq. (4) as a normalization method on the validation set, the performance of both the four electrodes and the three electrodes tends to be similar, as shown in Fig. 8, with an average difference of 20 $\mu$m. The large computational difference in both cases are

ISBN: 978-3-95450-249-3 ISSN: 2673-5350 doi: 10.18429/JACoW-IBIC2024-TUP15

caused by beam instability or the dependence on flow intensity. In practical applications, the flow intensity can affect the voltage values, so normalization not only improves training efficiency, but also enhances the model's generalization ability.

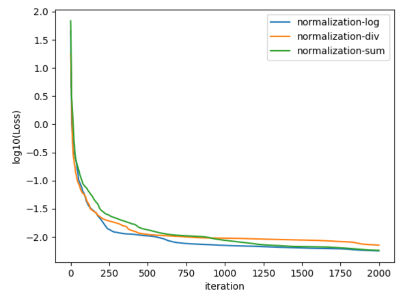


Figure 6: Loss plots for different normalization methods.

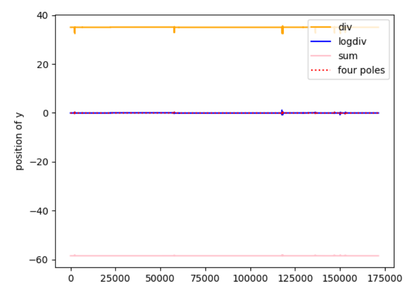


Figure 7: Comparison of validation sets for different normalization methods.

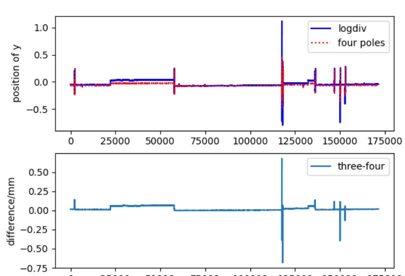


Figure 8: Comparison of four-electrode and three-electrode on the validation set.

## *Neural Network Training Effect*

We utilized the calibration data from BPM to train the neural network. Based on the experimental findings, it is evident that the errors for three and four voltages are mostly below 5 $\mu$m on test set, as shown in Fig. 9, with a few exceeding 5 $\mu$m but still within 10 $\mu$m, and only a small number falling within the 15 $\mu$m range. The average error for four electrodes is 0.849 $\mu$m, while the average error for three electrodes is 1.035 $\mu$m.

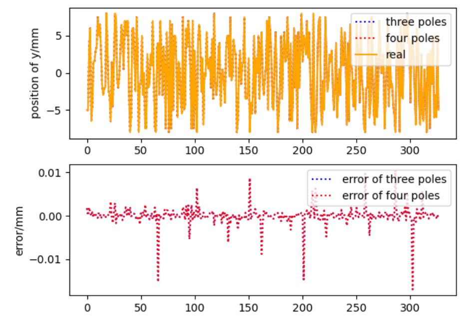


Figure 9: Comparison error of three-electrode and four-electrode on test set.

# CONCLUSION

It has been demonstrated that neural networks are capable of fitting any bounded continuous function. In our experiments, we have selected the appropriate normalization method, activation function and optimizer with the objective of reducing the average error in the calculation of the beam position to 1 $\mu$m. Furthermore, we have conducted experiments which have shown that three electrodes can be used to calculate the beam position, thus improving the robustness of the system.

# REFERENCES

[1] J. He *et al.*, "Beam position monitor design for the high energy photon source", *Meas. Sci. Technol.*, vol. 33, no. 11, p. 115 106, 2022. doi:10.1088/1361-6501/ac8277

[2] J. He *et al.*, "Design and fabrication of button-style beam position monitors for the heps synchrotron light facility", *Nucl. Sci. Tech.*, vol. 33, no. 11, p. 141, 2022. doi:10.1007/s41365-022-01126-7

[3] A. R. Barron, "Universal approximation bounds for superpositions of a sigmoidal function", *IEEE Transactions on Information theory*, vol. 39, no. 3, pp. 930–945, 1993. doi:10.1109/18.256500

[4] M. I. A. Lourakis, "A brief description of the levenberg-marquardt algorithm implemented by levmar", Foundation of Research and Technology, Heraklion, Crete, Greece, Tech. Rep. 4(1), 2005, pp. 1–6.