ISBN: 978-3-95450-204-2 ISSN: 2673-5350 doi:10.18429/JACoW-IBIC2019-WEPP021

MACHINE LEARNING IMAGE PROCESSING TECHNOLOGY APPLICATION IN BUNCH LONGITUDINAL PHASE DATA INFORMATION EXTRACTION*

X.Y.Xu †, Y.M.Zhou, Shanghai Institute of Applied Physics, CAS, Shanghai, China
also at University of Chinese Academy of Sciences, Beijing, China
Y.B.Leng#, Shanghai Advanced Research Institute, CAS, Shanghai, China

Abstract

To achieve the bunch-by-bunch longitudinal phase measurement, Shanghai Synchrotron Radiation Facility (SSRF) has developed a high-resolution measurement system. We used this measurement system to study the injection transient process, and obtained the longitudinal phase of the refilled bunch and the longitudinal phase of the original stored bunch. A large number of parameters of the synchronous damping oscillation are included in this large amount of longitudinal phase data, which are important for the evaluation of machine state and bunch stability. The multi-turn phase data of a multi-bunch is a large two-dimensional array that can be converted into an image. The convolutional neural network (CNN) is a machine learning model with strong capabilities in image processing. We hope to use the convolutional neural network to process the longitudinal phase two-dimensional array data and extract important parameters such as the oscillation amplitude and the synchrotron damping time.

INTRODUCTION

With the enhancement of computer computing performance, machine learning technology has been greatly developed in recent years. In particular, the concept of deep learning allows us to use algorithms to deal with the complex problem with a large amount of data. Accelerator scholars have also begun to gradually apply machine learning technology to beam measurement and diagnosis, have achieved some remarkable results [1].

The development of machine learning in image processing is obvious to all, such as face recognition, automatic driving and so on. Using image processing techniques such as convolutional neural networks, we can give machines the ability to process image data. In our field of beam measurement, we collect huge amounts of data every day. These data are often stored in two-dimensional or even multi-dimensional arrays. If the two-dimensional array is restored according to the storage method of the grey value image, it can be drawn into a greyscale image. So can we use cutting-edge deep learning techniques such as multi-layer convolutional neural networks to process such data in the form of multi-dimensional arrays and extract the information we need from it? In this regard, we made a preliminary attempt. The object we are dealing with is the electrical signal data of the BPM electrode at the instant of the accelerator injection. We hope to use this data to extract the relevant parameters of synchrotron damping oscillation when injecting transients.

The study of the injection transient is helpful for optimizing the state parameters of the injector and understanding the physical process of the fusion of the stored charge and the refilled charge during the injection [2]. Therefore, starting from 2012, the BI group of SSRF (Shanghai Synchrotron Radiation Facility) performs bunch-by-bunch phase measurement analysis, and uses phase multi-channel delay sampling technology and table look-up method to achieve phase acquisition. These phase data are fitted by gradient descent method or the fitting function library in commercial data software such as Matlab to obtain the synchronous damping oscillation parameters. The most primitive data is the BPM electrical signal obtained using a digital acquisition board. These electrical signals are stored in the form of multi-dimensional arrays. If we treat these data as multiple images and deal with it with deep learning multi-layer convolutional neural networks, we can directly extract the synchrotron damping oscillation parameters of the injection transients, which will greatly streamline the data processing process and provide a viable solution for instant online processing [3].

CONVOLUTIONAL NEURAL NETWORK

Convolutional Neural Networks (CNN) is a type of feed-forward neural network with convolutional computation and deep structure. It is one of the representative algorithms of deep learning [4-5]. Convolutional neural networks have the ability to represent learning, and can shift-invariant classification of input information according to their hierarchical structure. Therefore, it is also called " Shift-Invariant Artificial Neural Networks, SIANN" [6].

Convolutional neural networks are inspired by the visual organization of living things. Visual cortical cells receive signals from photoreceptors on the retina, but a single visual cortical cell does not receive all of the signals from the photoreceptor, but only accepts the signal from the stimulus region it dictates. Only by feeling the stimulation in the field can the neuron be activated. Multiple visual cortical cells systematically superimpose the receptive field, completely receiving signals transmitted by the retina and establishing a visual space.

*Work supported by National Natural Science Foundation of China (No.11575282) and Ten Thousand Talent Program and Chinese Academy of Sciences Key Technology Talent Program
†xuxingyi@sinap.ac.cn
#lengyongbin@sinap.ac.cn

8th Int. Beam Instrum. Conf. IBIC2019, Malmö, Sweden JACoW Publishing
ISBN: 978-3-95450-204-2 ISSN: 2673-5350 doi:10.18429/JACoW-IBIC2019-WEPP021

DATASET FOR CNN

The data set consists of 482 sets of BPM electrical signals captured at the instant of accelerator injection. We use a threshold trigger algorithm to automatically capture the injection process. In order to realize the acquisition of BPM electrical signals, we chose a high-speed data acquisition board, and with the fixed delay line, we realized that each single bunch collects two data points every turns when it goes through the BPM. These two data points are derived from the same electrode of BPM, separated by 100ps. We use channel A and channel D to refer to these two data channels. The stored data format is 512*2080*2. Among them, 512 is more than 500 bunches in the storage ring of Shanghai light source. We collect more than 2000 data in each injection, in which the injection process starts from the first hundred circles, and 2 represents the two data channels A and D. In order to facilitate the processing and speed up the training process, we cut and normalize the data. In order to highlight the injected beam, we expanded the injected bunch by 300. The data finally introduced into the algorithm is 482*800*1800*2 in the range of 0-1. The dataset is shown in Table 1.

Table 1: Dataset for CNN

Margin	Original data	Pre-processed data
Shape	482*512*2080*2	482*800*1800*2
Sets	482	482
bunch	512	500
Turns	2080	1800
Channel	2	2
range	-15000-20000	0-1

APPLICATION OF CONVOLUTIONAL NEURAL NETWORK IN DATASET

After the data pre-processing is completed, the data of the two channels A and D in each group of data can be drawn into images which shown in Fig. 1 and Fig. 2. We divide the data set into a training set and a test set according to 3:1, and use the synchrotron damping oscillation parameters obtained by a series of methods such as look-up table method as labels. The goal is to extract the amplitude and damping time in synchronous damped oscillations by a convolutional neural network (shown in Fig. 3).

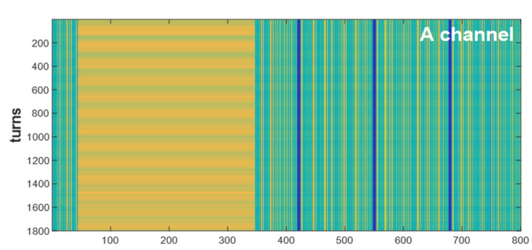

Figure 1: A group data of A channel drawn into an image.

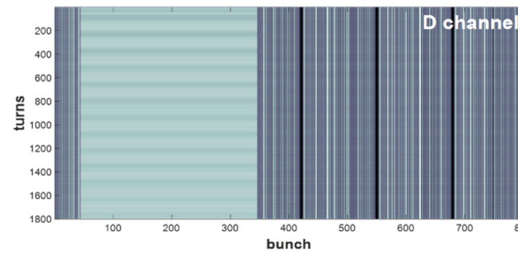

Figure 2: A group data of D channel drawn into an image. D channel is similar to A channel. The difference between them is caused by the delay(100ps). With both of them, we can remove the influence of different charge quantities to get the phase oscillation parameters.

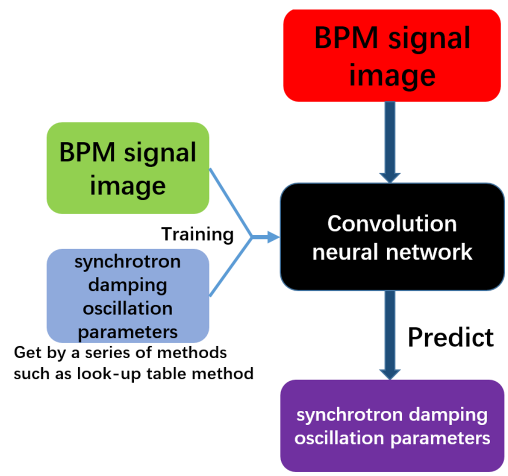

Figure 3: Machine learning schematic to predict synchronous damped oscillation parameters.

In each prediction, we have two images composed of A and D channels. The corresponding position data of the two images is the BPM data of the same bunch on the same turn. The two have 100 ps delay. Using this two-channel data, we can remove the effects of different charge quantities between different bunch and extract the phase oscillation parameters. Considering the relationship between the two pictures, we refer to the three-channel simultaneous processing method of the three primary colors of the color picture, and treat each training picture as a color picture with two primary colors. That is, the training data is a picture of 1800*800 with a thickness of 2.

We have established a convolutional neural network consisting of one input layer, five convolutional layers, six pooling layers, two fully-connected layers, and one output layer to deal with this problem (shown in Fig. 4). The convolution kernel in the convolutional layer can respond to the coverage data, and the convolutional layer extracts the different characteristics of the input by the movement of the convolution kernel on the input layer. By pooling the data, the pooling layer downsamplings a large matrix into a small matrix, reducing the amount of computation and preventing overfitting. The fully-connected layer is similar to the common BP artificial neural network. Each neuron is connected by a weight parameter and activated by the activation function after the weighted average.

ISBN: 978-3-95450-204-2 ISSN: 2673-5350 doi:10.18429/JACoW-IBIC2019-WEPP021

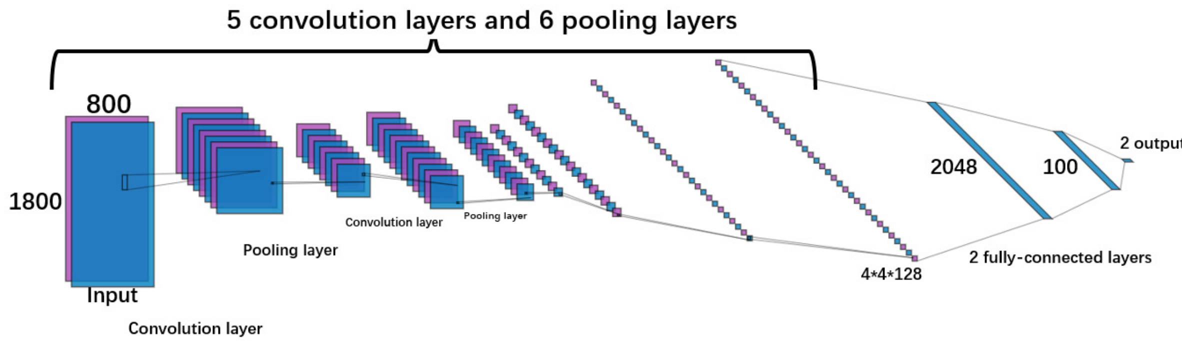

Figure 4: The structure of the convolutional neural network built in this experiment which consisting of one input layer, five convolutional layers, six pooling layers, two fully-connected layers, and one output layer

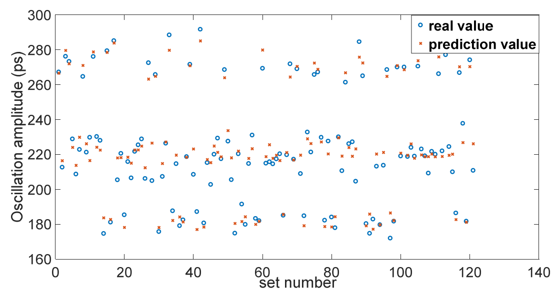

Figure 5: Predicted value versus real value about oscillation amplitude.

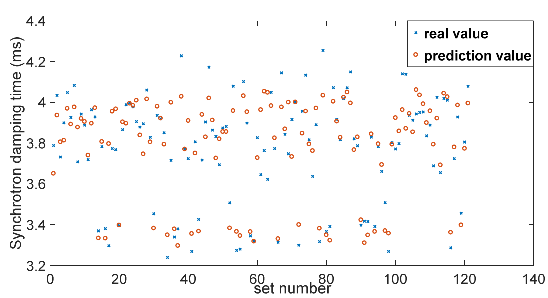

Figure 6: Predicted value versus real value about Synchrotron damping time.

After training the convolutional neural network model with 361 sets of training data, our model's cost function quickly converges. After that, we applied the trained model to the 121 sets of test data, and obtained the predicted value for each set of test data. We show these 121 sets of prediction value and test(real) value in Fig. 5 (oscillation amplitude) and Fig. 6 (Synchrotron damping time). It can be seen that the predicted value is in good agreement with the real value, wherein the prediction error for the oscillation amplitude is about 5 ps, and the prediction error of the Synchrotron damping time is about 0.08 ms. This margin of error can be covered by experimental accuracy and is within our acceptance.

The error between the predicted value and the true value is analyzed. On the one hand, because the capture time period of the injection process is long, there are only hundreds of sets of data in this training, which leads to the model not being adequately trained. On the other hand, the so-called real value is obtained by a series of processes such as look-up table method, and the error introduced by this process is large. This will result in inaccurate labels value in the training data.

CONCLUSION

In this paper, the machine learning image processing technology, convolution neural network(CNN), is used to process the BPM electrical signal data stored in the form of multidimensional matrix. The innovation lies in the processing of two-channel two-dimensional matrix data as images of two primary colors. Here we use the characteristics of translation invariance and weight sharing of convolutional neural networks, which solves the problem that the refilled bunches appear randomly in 512 bunches and realizes high-speed processing of large-scale data.

It can be seen that through application of machine learning technology, we have achieved extraction of synchronous damping oscillation parameters directly from the BPM raw data. In subsequent experiments, we intend to capture more training data to improve prediction accuracy. In addition, we know that the BPM signal actually contains almost all the information of the bunch, and we need to mine the data efficiently. Our future goal is to extract a series of beam state parameters we need from a multi-dimensional BPM data, such as position, bunch size, and charge.

REFERENCES

[1] E.Fol, J. Coello de Portugal, and R. Tomas, "Application of machine learning to beam diagnostics", in *Proc. 7th International Beam Instrumentation Conf. (IBIC'18)*, Shanghai, China, Sep.2018, pp. 169-176. doi:10.18429/JaCoW_IBIC2018-TUOA02

[2] Zhou, Y.M. et al., "Bunch-by-bunch longitudinal phase monitor at SSRF", in *Proc. Nuclear Science and Techniques*, 2018. 29: 113

[3] X.Y.Xu, Y.M.Zhou, and Y.B.Leng, "Machine learning application in bunch longitudinal phase measurement", in *Proc. 10th Int. Particle Accelerator Conf. (IPAC'19)*, Melbourne, Australia, May, 2019, pp. 2625-2628. doi:10.18429/JACoW-IPAC2019-WEPGW064

ISBN: 978-3-95450-204-2 ISSN: 2673-5350 doi:10.18429/JACoW-IBIC2019-WEPP021

[4] Goodfellow, I., Bengio, Y., and Courville, A., Deep learning (Vol. 1)． Cambridge：MIT press，2016：326-366

[5] Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, L., Wang, G. and Cai, J., 2015. “Recent advances in convolutional neural networks”, arXiv preprint arXiv:1512.07108

[6] Zhang, W., 1988. “Shift-invariant pattern recogni-tion-neural network and its optical architecture”, in *Proc. annual conference of the Japan Society of Applied Physics*.