ISBN: 978-3-95450-215-8 ISSN: 2226-0366 doi:10.18429/JACoW-LINAC2022-MOPORI05

APPLICATION OF VIRTUAL DIAGNOSTICS IN THE FEBE CLARA USER AREA

J. Wolfenden*[1], C. Swain[1], C. P. Welsch[1], University of Liverpool, Liverpool, UK
T. H. Pacey[1], A. E. Pollard[1], J. Jones[1], D. Dunning[1], ASTeC/STFC, Daresbury, UK
[1]also at Cockcroft Institute, Daresbury, UK

Abstract

Successful user experiments at particle beam facilities are dependent upon the awareness of beam characteristics at the interaction point. Often, properties are measured beforehand for fixed operation modes; users then rely on the long-term stability of the beam. Otherwise, diagnostics must be integrated into a user experiment, costing resources and limiting space in the user area. This contribution proposes the application of machine learning to develop a suite of virtual diagnostic systems. Virtual diagnostics take data at easy to access locations, and infer beam properties at locations where a measurement has not been taken, and often cannot be taken. Here the focus is the user area at the planned Full Energy Beam Exploitation (FEBE) upgrade to the CLARA facility (UK). Presented is a simulation-based proof-of-concept for a variety of virtual diagnostics. Transverse and longitudinal properties are measured upstream of the user area, coupled with the beam optics parameters leading to the user area, and input into a neural network, to predict the same parameters within the user area. Potential instrumentation for FEBE CLARA virtual diagnostics will also be discussed.

INTRODUCTION

The interaction point (IP) in a particle accelerator is the focal point at which the attention of users and operators converge. At this point users require certain beam parameters in order to achieve their desired output, whilst operators monitor this location and tune the machine settings accordingly. The balance of these two sets of goals is key to any successful exploitation plan. If the user blocks the operators diagnostic efforts then they cannot be certain of the beam parameters their instrumentation receives; likewise if operator diagnostics interfere with the users ability to receive the beam in a manner suited to their needs, their output is affected. A utilitarian approach is therefore required to proceed. Standard practise is therefore to operate in nominal "user modes". These are machine settings which provide a stable beam with known parameters. These parameters are measured in great invasive detail ahead of user operation, and the machine stability utilised to reliably reproduce these parameters once invasive high resolution diagnostics are removed. In general, this approach works well. Other lower resolution non-invasive measures can be used to monitor the beam away from the IP, and machine jitter can be quantified and converted into an uncertainty for users, which can then be baked directly into their output. However, issues can arise when users require non-standard beam parameters or when the facility is using a novel acceleration scheme, such as plasma wakefield acceleration [1, 2]. The former requires operators to tune the machine settings on-the-fly, meaning either the beam is tuned with less resolution due to a lack of diagnostics, or the user loses beam time as the required diagnostics are inserted and then removed from the IP. The latter suffers, at this time, from a fundamental shot-to-shot instability, which means a much larger error is introduced when relying upon machine stability with certain machine parameters. An obvious solution one might suggest is the implementation of several novel non-invasive diagnostics that have been developed in recent years [3–6]. Unfortunately these methods would still fall foul of the user-operator balance described above as the instrumentation would need to be placed close to the IP, and hence user instrumentation. It is here that the concept of a virtual diagnostic (VD) can be deployed. A VD is a technique based upon machine learning which uses beam measurements from one location on a beamline to infer, with high accuracy, beam parameters at another location. This practise could therefore be used to move high resolution IP diagnostics away from the IP, freeing the space for users, whilst still providing shot-to-shot beam measurements, even in exotic operational modes.

Presented in this contribution is a case study into the application of such VDs in the framework of the full energy beam exploitation (FEBE) CLARA (STFC, UK) [7] user area IP. This simple example focuses solely on transverse beam size measurements from particle tracking simulations (Elegant [8]), but other beam parameters will be discussed. The goal is to maintain an alignment with experimental plans for the facility, providing actionable VD implementations. Two VD options will be discussed, along with the accompanying diagnostic instrumentation.

SIMULATION

As with any application of machine learning, a significant quantity of quality input and output data is required in order to facilitate model training. To produce this data, the particle tracking code Elegant was used. A lattice file for a nominal FEBE CLARA operational mode was chosen. This lattice had been tuned to maximise the beam current at the IP; however, the actual absolute values of the machine settings are unimportant at this stage. In order to produce a random sampling of the transverse beam parameters at the IP, the K1 values of several quadrupole magnets in the beamline were chosen, presented in Fig. 1, and randomly

* joseph.wolfenden@cockcroft.ac.uk

ISBN: 978-3-95450-215-8 ISSN: 2226-0366 doi:10.18429/JACoW-LINAC2022-MOPORI05

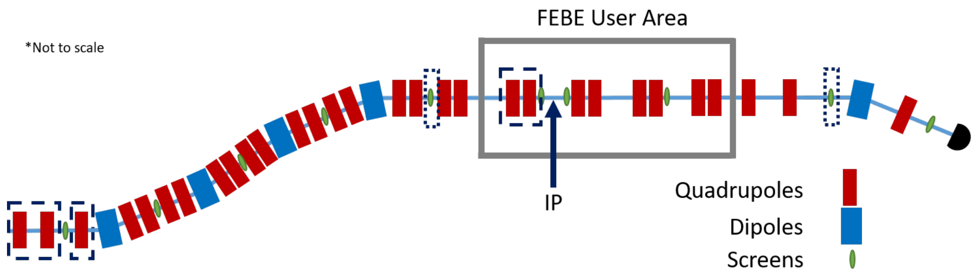

Figure 1: A latter section of the FEBE CLARA beamline. Quadrupoles varied and screens used in simulations are indicated.

varied within a range of ±20 % around the nominal value. These quadrupoles were chosen as they produced the largest change in transverse beam parameters at the IP, whilst maintaining beam transport. This was specific to this beam setup and other operational modes may require different element variations to achieve the same effect. The 6D phase parameters of the ensuing macro-particles were then captured at planned diagnostic stations; again, keeping a practical focus on long-term implementation. These 6D measurements could then be used to produce a variety of inputs and outputs for the machine learning model. In this test case, a simple scenario of 2D beam profile was chosen.

Two implementation methods were targeted in this study. The first was a "Pre-IP" method, placing a non-invasive beam profile diagnostic upstream of the IP to produce beam profile measurements at the IP. For this model, a screen at the end of the FEBE transfer line was chosen as an input, with the IP as an output, indicated in Fig. 1. The second implementation was a "Post-IP" method, placing an invasive beam profile diagnostic downstream of the IP close to the beam dump, to produce upstream beam profile images of the IP. This instance required the beam dump screen as an input, with the IP as an output once again, shown in Fig. 1. Simulations were used to generate two databases of 10,000 pairs of beam profile images, with their associated machine settings (i.e. quadrupole K1 values).

CONVOLUTIONAL NEURAL NETWORK

The machine learning model chosen in this instance was based upon a convolutional neural network (CNN) architecture [9]. The model takes the images described in the above section (48 px × 48 px) and the machine settings (varied quadrupole K1 values in this instance) as inputs and produces beam profile images at the IP as an output. The model was initially trained roughly by hand before conducting hyperband tuning [10]. Hyperband tuning produces a large array of models instances with stochastically varied hyperparameters. These models are trained for several epochs, the accuracy produced is then evaluated, and the worst performing models are dropped. This process then repeats for the remaining models until only one optimal model remains. This process was conducted separately for each of the two test cases. The model structure, with associated tunable hyperparameters, is presented in Fig. 2. The first CNN sub-structure is comprised of several convolution and maxpooling layers, the second CNN sub-structure is comprised of several convolution and upsampling layers. The fully connected (FC) dense layers at the centre join the two sub-structures together and provide an input for the machine settings.

RESULTS

The models produced by the tuning process were then passed example input images and machine settings to evaluate the IP images produced. Example images of the Pre-IP (top) and Post-IP (bottom) methods are presented in Fig. 3. The left images are the input images, the centre images are the simulated images from Elegant, and the right images are the model prediction. It is clear from Fig. 3 that the beam profile can be reproduced to extremely fine detail, with even

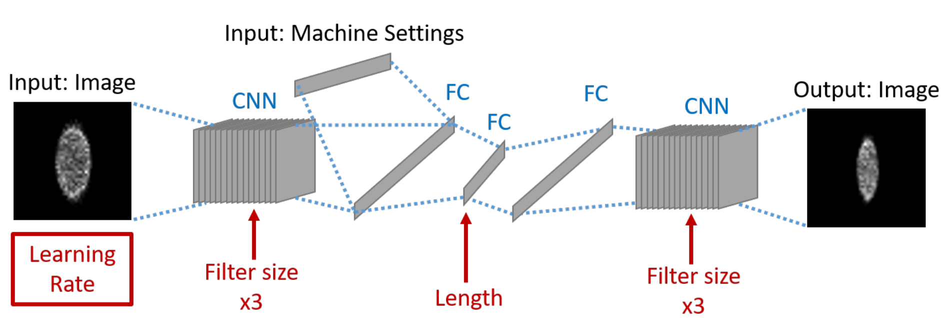

Figure 2: The model structure of the full tunable CNN. Tuned parameters are in red, model components in blue. CNN = Convolutional Neural Network; FC = Fully Connected.

ISBN: 978-3-95450-215-8 ISSN: 2226-0366 doi:10.18429/JACoW-LINAC2022-MOPORI05

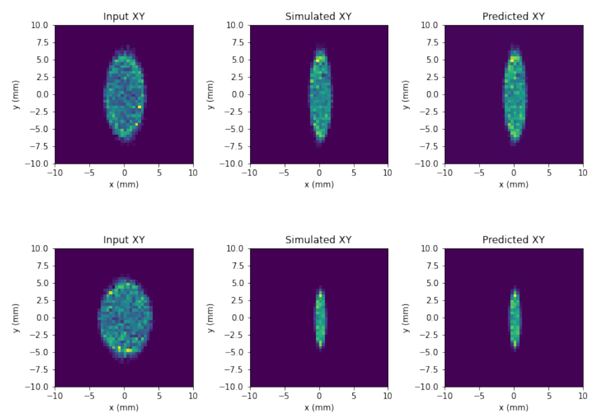

Figure 3: Example inputs, simulated outputs, and predicted outputs for the two models. Top: Pre-IP method. Input 2D beam profile from FEBE CLARA transfer arc, output 2D beam profile at IP. Bottom: Post-IP method. Input 2D beam profile from FEBE CLARA beam dump, output 2D beam profile at IP.

small charge density fluctuations within the profile being predicted correctly. The RMS error generated by the training and tuning process on these predicted images in comparison to the simulated ground truth is on average ~ 0.01 %.

There are negligible differences in performance between the Pre-IP and Post-IP methods. This implies that the reconstruction process is reversible, despite the model missing several parameters which would traditionally be viewed as critical to a particle tracking code. Therefore, which of the two methods to implement would be dependent upon the practical scenario in question, and would not be reliant on, or suffer from, variations in accuracy or resolution.

CONCLUSIONS

This contribution has demonstrated a test case for the utilisation of VD techniques to predict IP beam profiles using measurements away from the IP. The high accuracy of these results could be linked to the simplicity of these test cases; from Fig. 1, which is the latter section of FEBE CLARA, it is clear that the number of optical elements which have been varied is a small quantity of the possible options. The link to longitudinal profile variations has also not been included. Therefore, further studies are required to increase the complexity of the variations within the training data and to include longitudinal effects. This framework has been constructed with this in mind, and will serve as a solid foundation for these works. The simulations and CNN architecture are in no way linked or dependent upon the specifics of the test cases studied here, and would function equally well for other measurable quantities beyond beam profile; although it is likely that the accuracy would vary from that found here.

An secondary outcome of this work has also been the predictive capability of the CNN model. In the Pre-IP case, for a given input image, the final focusing quadrupoles can be varied offline, and the IP beam profile predicted ahead of time. This could serve as a operational tuning tool during user beam periods, where non-standard operational modes have been requested.

This study has been driven by a goal of simple implementation. Beam profiles are measured using an array of standardised instrumentation, both invasive and non-invasive, meaning the experimental measurements required to drive the work presented here would be simple to put into practise. Future work looking at other beam parameters must take this into account. Including experimental and hardware limitations into these models in often a simple, yet overlooked, task. This will often ease the transition from training and prediction with simulated results, to practical measurables.

ACKNOWLEDGEMENTS

This work is supported by the AWAKE-UK phase II project funded by STFC and the STFC Cockcroft core grant No. ST/G008248/1.

REFERENCES

[1] E. Adli, A. Ahuja, *et al.*, “Acceleration of electrons in the plasma wakefield of a proton bunch", *Nature*, 2018. doi:10.1038/s41586-018-0485-4

[2] A. J. Gonsalves, K. Nakamura, *et al.*, “ Petawatt laser guiding and electron beam acceleration to 8 gev in a laser-heated capillary discharge waveguide", *Phys. Rev. Lett.*, Feb 2019. doi:10.1103/PhysRevLett.122.084801

[3] M. Bergamaschi, A. Aryshev, *et al.*, “Noninvasive Micrometer-Scale Particle-Beam Size Measurement Using Optical Diffraction Radiation in the Ultraviolet Wavelength Range", *Phys. Rev. Applied 13*, Jan. 2020. doi:10.1103/PhysRevApplied.13.014041

[4] N. Kumar, A. Salehilashkajani, *et al.*, “ Non-invasive beam profile monitor for medical accelerators", *Phys Med.*, vol. 73, pp. 173–178, May 2020. doi:10.1016/j.ejmp.2020.04.023

[5] R. Fiorito, A. Shkvarunets, *et al.*, “Noninvasive emittance and energy spread monitor using optical synchrotron radiation", *Phys. Rev. ST Accel. Beams 17*, vol. 17, Iss. 12, Dec. 2014. doi:10.1103/PhysRevSTAB.17.122803

[6] R. Kieffer, L. Bartnik, *et al.*, “Direct Observation of Incoherent Cherenkov Diffraction Radiation in the Visible Range", *Phys. Rev. Lett. 121*, 2018. doi:10.1103/PhysRevLett.121.054802

[7] A. R. Bainbridge, D. Angal-Kalinin, J. K. Jones, T. H. Pacey, Y. M. Saveliev, and E. W. Snedden, “The Design of the Full Energy Beam Exploitation (FEBE) Beamline on CLARA”, in *Proc. IPAC’22*, Bangkok, Thailand, Jun. 2022, pp. 2594–2597. doi:10.18429/JACoW-IPAC2022-THPOPT015

[8] M. Borland, “Elegant: a flexible SDDS-compliant code for accelerator simulation", in *Proc. ICAP’00*, INSPIRE, Darmstadt, 2000. doi:10.2172/761286

ISBN: 978-3-95450-215-8 ISSN: 2226-0366 doi:10.18429/JACoW-LINAC2022-MOPORI05

[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks", *Advances in neural information processing systems*, pp. 1097–1105, 2012. doi:10.1145/3065386

[10] L. Lisha, K. Jamieson, *et al.*, “ Hyperband: A Novel Bandit-Based Approach to Hyperparameter Optimization", *Journal of Machine Learning Research 18*, 2018. doi:10.48550/arXiv.1603.06560