ISBN: 978-3-95450-232-5 ISSN: 2673-7000 doi:10.18429/JACoW-NAPAC2022-WEPA23

SRF CAVITY INSTABILITY DETECTION WITH MACHINE LEARNING AT CEBAF*

D. Turner†, R. Bachimanchi, A. Carpenter, J. Latshaw, C. Tennant, L. Vidyaratne
Thomas Jefferson National Accelerator Facility, Newport News, VA 23606, USA

Abstract

During the operation of the Continuous Electron Beam Accelerator Facility (CEBAF), one or more unstable superconducting radio-frequency (SRF) cavities often cause beam loss trips while the unstable cavities themselves do not necessarily trip off or present a fault. Identifying an unstable cavity out of the hundreds of cavities installed at CEBAF is difficult and time-consuming. The present RF controls for the legacy cavities report at only 1 Hz, which is too slow to detect fast transient instabilities. A fast data acquisition system for the legacy SRF cavities is being developed which samples and reports at 5kHz to allow for detection of transients. A prototype chassis has been installed and tested in CEBAF. An autoencoder based machine learning model is being developed to identify anomalous SRF cavity behavior. The model is presently being trained on the slow (1 Hz) data that is currently available, and a separate model will be developed and trained using the fast (5 kHz) DAQ data once it becomes available. This paper will discuss the present status of the new fast data acquisition system and results of testing the prototype chassis. This paper will also detail the initial performance metrics of the autoencoder model, which indicate good results on the available 1Hz data. This paper will also discuss high-level software applications developed to support the project.

INTRODUCTION

The Continuous Electron Beam Accelerator Facility (CEBAF) is a 5.5-pass, 12 GeV continuous wave (CW) electron accelerator. CEBAF is comprised of two anti-parallel superconducting RF linacs connected by two sets of recirculation arcs. Electron beams are then extracted into up to four experimental halls to support nuclear physics experiments (Fig. 1).

Each linac is comprised of 200 superconducting radio-frequency (SRF) cavities, plus 18 cavities in the injector, which provide electron beam acceleration. During the operation of CEBAF, any one unstable SRF cavity can cause beam loss trips while the unstable cavities themselves do not necessarily trip off or present a fault.

The existing tools and diagnostics for identifying unstable SRF cavities at CEBAF are presently rather limited. Out of the 418 SRF cavities at CEBAF, 306 are of the original legacy CEBAF design which lacks the fast data acquisition capabilities of the newer cavities. The legacy cavity diagnostics are limited are limited to what is presented to the EPICS [1] control system which is limited to a data rate of 1 Hz, which is not fast enough to capture transient instabilities. Also, simply the number of cavities in the machine makes the search for instabilities time consuming. Work described in [2] addresses fault classification for the newer SRF cavities at CEBAF. This work addresses identifying unstable cavities of the legacy design.

To address these problems, a new fast data acquisition (DAQ) system has been designed and is being installed on the legacy SRF cavities at CEBAF. An autoencoder based machine learning anomaly detection model is being developed to identify cavity instabilities from the fast DAQ data along with archived EPICS data recorded by the MYA archiver [3].

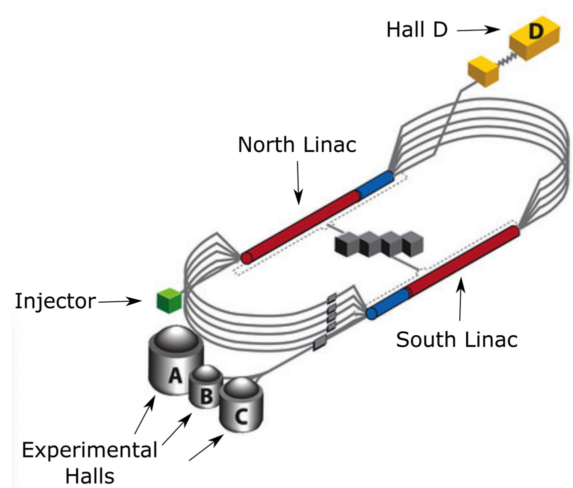

Figure 1: 12 GeV CEBAF overview. The legacy SRF cavities comprise most of the red regions of the linacs.

FAST DATA ACQUISITION SYSTEM

Each legacy SRF cavity has four analog outputs on its control module:

- GMES: Measured Gradient. Should mirror the GSET value in EPICS and should be stable to within 0.044% during normal operations. Has DC and AC components.
- PMES: Measured Phase. This is the phase in the cavity measured against the absolute 70 MHz reference. This also should mirror PSET in EPICS and should be varying less than $\pm 0.5°$.
- GASK: Gradient Drive. The change in klystron incident power needed to maintain the GMES value within

* Notice: Authored by Jefferson Science Associates, LLC under U.S. DOE Contract No. DE-AC05-06OR23177. The U.S. Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce this manuscript for U.S. Government purposes.

† dturner@jlab.org

ISBN: 978-3-95450-232-5 ISSN: 2673-7000 doi:10.18429/JACoW-NAPAC2022-WEPA23

specifications. This signal will change in response to beam detuning or beam loading.

- PASK: Phase drive. The change in klystron phase to the cavity needed to maintain the PMES value within specifications.

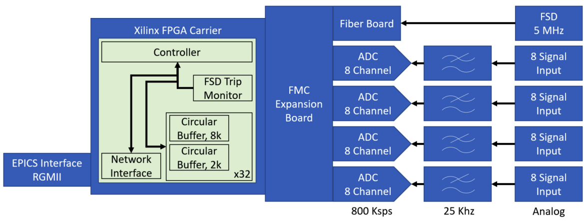

Figure 2: Fast DAQ block diagram.

Each DAQ chassis has inputs for eight cavities, for a total of 32 channels (Fig. 2). High frequency noise is filtered with analog circuitry before being piped to ADCs. Each ADC has 8 channels that samples at 800 kilosamples per second (ksps, 100 ksps per channel). ADC samples are saved into dedicated circular buffers which are the Scope and Fault Buffers. The Scope Buffer (Fig. 3, left) is a 2K buffer (2 times 1024 samples) and is essentially a network attached oscilloscope which provides real time analysis of the control signals. The Fault Buffer (Fig. 3, right) is an 8K buffer and records 1.6 seconds of samples (when the adjustable sample rate is set to 5 kHz). This Fault Buffer will fill up during either a software initiated fault or initiated by a loss of signal from the machine protection system. The buffer will “freeze” at the time of the fault signal being received. The Fault Buffer data is fed to the autoencoder model for inference.

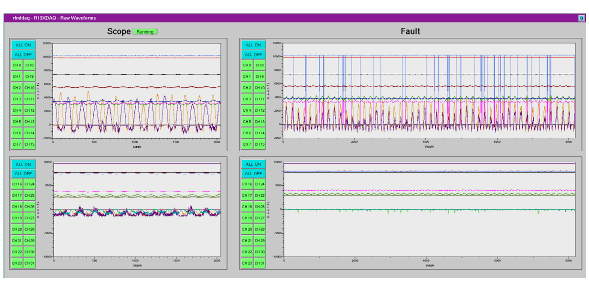

Figure 3: Fast DAQ raw waveforms screen.

Figure4 shows an example of the DAQ Fault Buffer data as it will be input into the autoencoder model.

There is presently a prototype DAQ chassis installed in one zone in the North Linac for testing. The prototype chassis performs well; some data has already been collected from it. Components for the remaining chassis are being procured and assembly has begun. Production of the DAQ chassis was delayed due to procurement problems related to COVID-19. The original budget covered the cost of outfitting both linacs with the fast DAQ chassis, but the cost of components increased due to COVID-19 to the point that only one linac can be outfitted within the allocated budget.

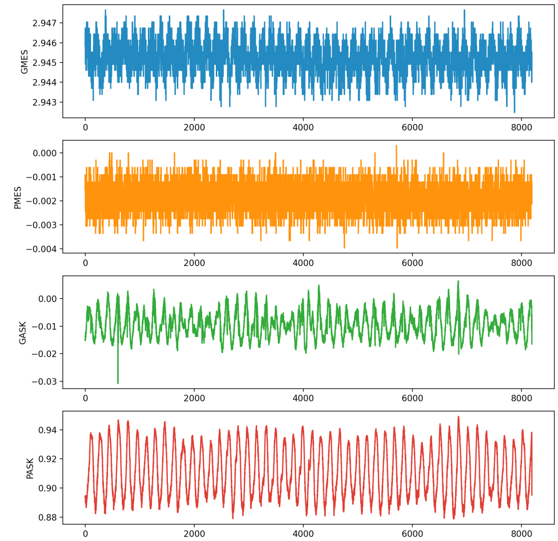

Figure 4: Fast DAQ waveform data.

ARCHIVED EPICS DATA

The MYA archiver records EPICS readbacks from the SRF cavity controls at a rate of 1 Hz. For initial training and development of the autoencoder model, data for 10 seconds prior to a beam trip and 5 seconds after are collected from the archiver. Figure 5 shows two data sets as examples. Note that the first trace in each plot is the total linac current, where at $t = 10$ s, the total current drops to zero indicating a trip. The left plot shows an example of a stable cavity. The plot on the right shows that the measured gradient around the time of the trip is unstable, therefore suggesting this cavity caused the beam loss trip. The data samples were labeled visually using such criteria as signals appearing unstable shortly before or during the time where total linac current is dropping to zero.

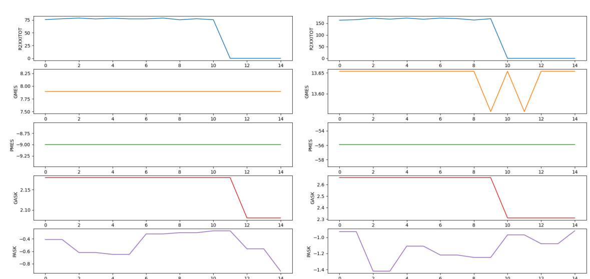

Figure 5: Examples of archived EPICS data from stable (left) and unstable (right) cavities.

AUTOENCODER MODEL

An autoencoder model [4] was chosen because most of the data will likely be of stable cavities; unstable cavities are expected to be fewer and farther between. The problem is one of anomaly detection rather than classification.

ISBN: 978-3-95450-232-5 ISSN: 2673-7000 doi:10.18429/JACoW-NAPAC2022-WEPA23

The autoencoder model is being developed using Python [5] and PyTorch [6]. Figure 6 shows the basic architecture of the autoencoder model being developed for this project.

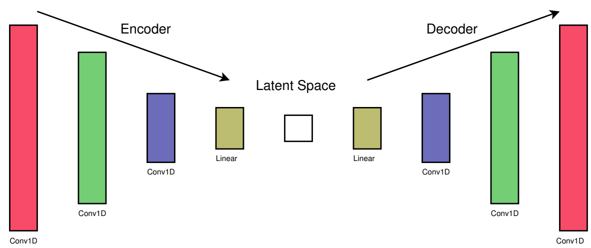

Figure 6: Autoencoder model.

The inputs to the model for archived EPICS are the same four inputs collected by the Fast DAQ, except that they are sampled at 1 Hz via EPICS and MYA rather than directly at 5 kHz. The model also includes the total linac current as an input. Early results on the archived EPICS data look promising. The initial autoencoder model was trained on 3648 examples of stable cavities. We tested autoencoder performance on a dataset of 912 stable samples and 112 unstable examples. Figure 7 shows the reconstruction losses for cavity data sets labeled "good" (i.e. stable) and cavity data sets labeled "bad" (i.e. unstable). Using 0.001 as the threshold between "good" and "bad", we get an accuracy of ~ 0.99 (7 out of 1036 samples falsely identified as "bad"). Development continues as more data is gathered and additional EPICS signals are included in the data sets. A similar model will be developed and trained using the fast DAQ data once it is available.

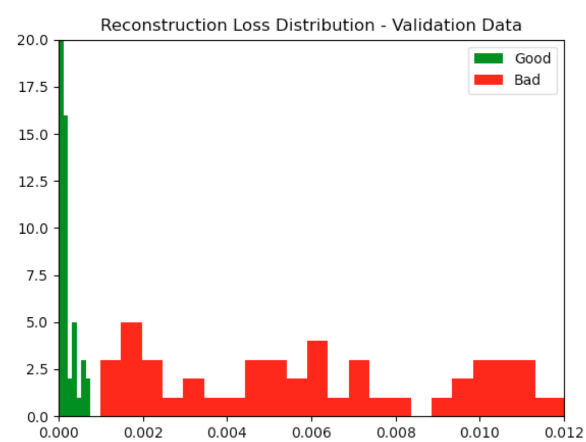

Figure 7: Initial results on archived EPICS data (zoomed in for clarity).

HIGH LEVEL SOFTWARE TOOLS

A high-level software filtering daemon is being developed to selectively harvest DAQ data (to conserve disk space) for inference by the autoencoder model. When the daemon detects a fast shutdown trip, it determines which machine protection device(s) initiated the trip to decide the nature of the trip. If the nature of the trip indicates that the trip was possibly caused by an unstable SRF cavity which did not itself present a fault, the DAQ Fault Buffer data, along with a timestamp for retrieval from the archiver, are stored on disk for inference by the autoecoder.

Another software tool (Fig. 8) is being developed to provide an interface to the autoencoder model to operators and technicians. The tool will display, for one trip event, the reconstruction loss for each of the cavities in the form of a histogram. The reconstruction loss here is interpreted as the likelihood of a cavity having caused a beam loss trip due to instability. The interface thus allows operators and technicians to quickly identify cavities that the autoencoder predicts are unstable and should be investigated further.

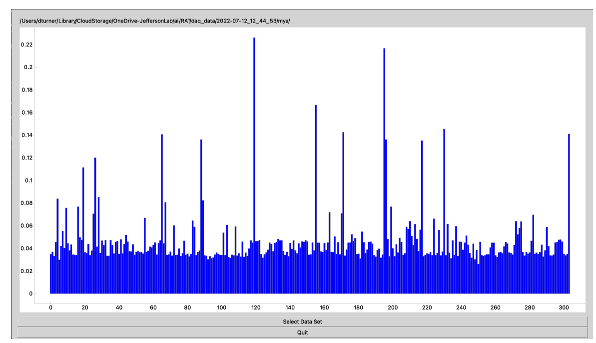

Figure 8: Operator interface.

CONCLUSION

The initial results of the autoencoder model with the archived EPICS data look promising. Data labeling, model development, and model training will continue, along with model development for the fast DAQ data and software tool development. Procurement, assembly, and installation of the fast DAQ hardware is underway.

REFERENCES

[1] Experimental Physics and Industrial Control System, `https: //epics.anl.gov/`

[2] C. Tennant *et al.*, "Superconducting radio-frequency cavity fault classification using machine learning at Jefferson Laboratory", *Phys. Rev. Accel. Beams*, vol. 23, no.11, p. 114601, 2020. `doi:10.1103/PhysRevAccelBeams.23.114601`

[3] C. J. Slominski and M. Bickley, "A MySQL Based EPICS Archiver", in *Proc. ICALEPCS'09*, Kobe, Japan, Oct. 2009, paper WEP021, pp. 447–449.

[4] S. Alla and S.K. Adari, *Beginning Anomaly Detection Using Python-Based Deep Learning*, ISBN: 978-1-4842-5176-8, Berkeley, CA, USA, Apress, 2019.

[5] V.G. Rossum and F.L. Drake, *Python 3 Reference Manual*, ISBN: 1441412697, Scotts Valley, CA, USA, CreateSpace, 1994.

[6] A. Paszke *et al.*, "PyTorch: An Imperative Style, High-Performance Deep Learning Library", *Advances in Neural Information Processing Systems 32*, pp. 8024-8035, 2019.