INITIAL STUDIES OF CAVITY FAULT PREDICTION AT JEFFERSON LABORATORY*

L. S. Vidyaratne†, A. Carpenter, R. Suleiman, C. Tennant, D. Turner, Jefferson Laboratory, Newport News, VA, USA
K. Iftekharuddin, Md. Monibor Rahman, ODU Vision Lab, Department of Electrical and Computer Engineering, Old Dominion University, Norfolk, VA, USA

Abstract

The Continuous Electron Beam Accelerator Facility (CEBAF) at Jefferson Laboratory is a CW recirculating linear accelerator (linac) that utilizes over 400 superconducting radio-frequency (SRF) cavities to accelerate electrons up to 12 GeV through 5-passes. Recent work has shown that given RF signals from a cavity during a fault as input, machine learning approaches can accurately classify the fault type. In this paper, we report initial results of predicting a fault onset using only data prior to the failure event. A dataset was constructed using time-series data immediately before a fault ("unstable") and 1.5 seconds prior to a fault ("stable") gathered from over 5,000 saved fault events. The data was used to train a binary classifier. The results gave key insights into the behaviour of several fault types and provided motivation to investigate whether data prior to a failure event could also predict the type of fault. We discuss our method using a sliding window approach. Based on encouraging initial results, we outline a path forward to leverage deep learning on streaming data for fault type prediction.

INTRODUCTION

The Continuous Electron Beam Accelerator Facility (CEBAF) at Jefferson Lab is a high power, continuous wave recirculating linac servicing four different experimental nuclear physics end stations [1]. CEBAF completed an energy upgrade in 2017 with a goal of effectively doubling its energy reach from 6 GeV to 12 GeV. The upgrade required the installation of 11 additional cryomodules, named C100s denoting their capability for providing 100 MV of energy gain. A schematic of CEBAF with locations of the new C100 cryomodules is provided in Figure 1. Each cryomodule is composed of 8 superconducting radio-frequency (SRF) cavities. In addition, a digital low-level radio frequency system (LLRF) is implemented to regulate the new cryomodules.

CEBAF experiences frequent short machine downtime trips caused by SRF system faults, particularly when the cavity gradients are pushed to their upper limits. A significant portion of the SRF system faults occur within the C100 cryomodules. Consequently, a data acquisition system is implemented to record data from these cryomodules to investigate the nature and the origin of the SRF faults. The system is configured to record time-series waveform data when any of the C100 cavities fault. These recorded waveform data are analyzed by a subject matter expert (SME) to determine the cavity that caused the trip, and the type of fault. In previous studies we have investigated the possibility of using artificial intelligence (AI) techniques to automate this highly tedious waveform analysis process [2, 3]. It is quite helpful to expedite the beam recovery process with fast automated cavity fault identification after an event. However, it may be further beneficial to investigate the possibility of using AI to predict RF failures beforehand in order to reduce certain faults from occurring. While the data acquisition system is being upgraded for compatibility with such predictive models, we conduct a feasibility study for RF fault prediction using currently available data.

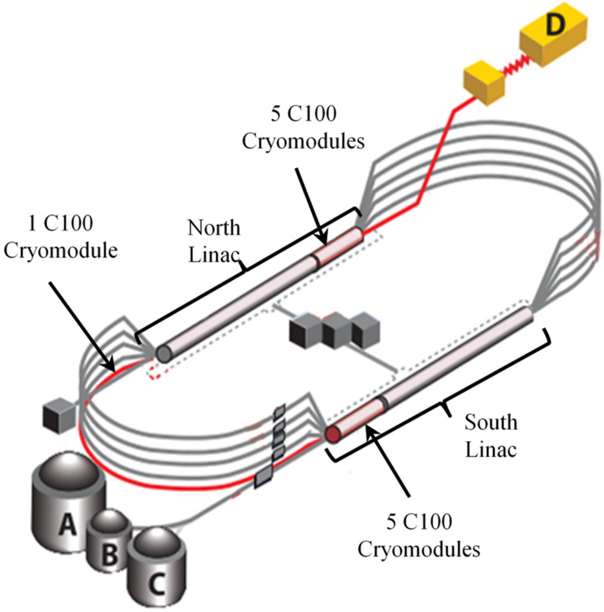

Figure 1: CEBAF schematic with locations of C100 cryomodules.

CAVITY FAULT CLASSIFICATION

The cavity fault classification is posed as a supervised machine learning (ML) problem, with ground truth fault labels for recorded data provided by SMEs. The data used for the classification task is the full time-series waveforms pertaining to each fault event recorded by the data acquisition system. The entire time-series waveform represents approximately 1638.4 ms (from t = −1536 ms to

* This work is supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics under Contract No. DE-AC05-06OR23177.

† lasithav@jlab.org

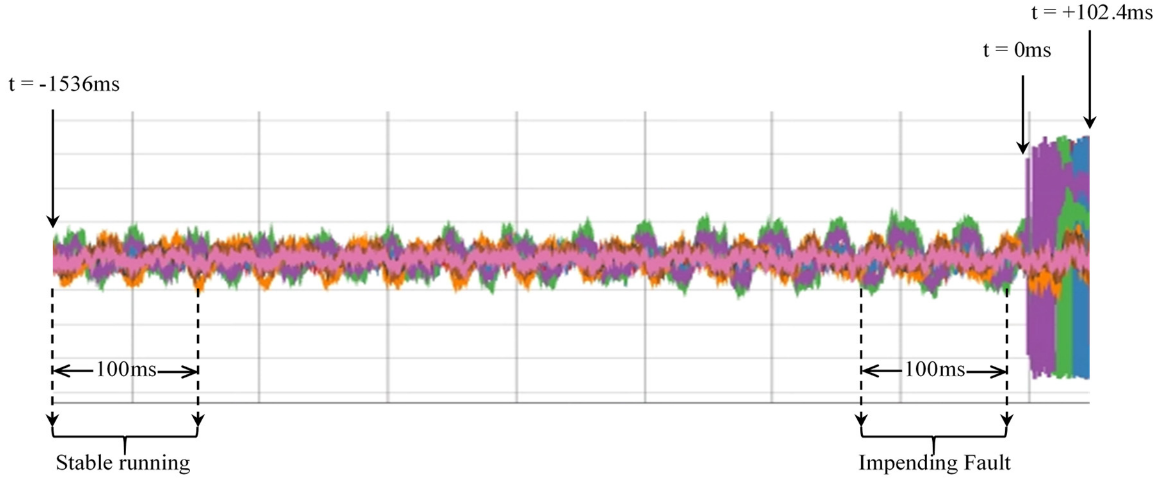

Figure 2: Waveform captured by the data acquisition system. The total duration of the waveform is 1.64 seconds.

t = +102.4 ms) comprised of both pre-fault and post-fault signals as shown in Figure 2. Waveforms are sampled at a constant rate of 5 kHz. We record 17 different RF signals from each C100 cavity, and a selection of these are used for analysis. In effect, the classification task uses both pre-fault and post-fault waveforms to identify the cavity that failed and the type of fault in a given cryomodule. Detailed discussions on the classical ML and deep learning (DL) models developed for cavity fault classification and associated performance characteristic can be found in [3, 4]. Table 1 summarizes the performance of ML and DL models for cavity fault classification with testing data. The best performing ML models are deployed within CEBAF.

Table 1: Cavity and Fault Type Classification Accuracy

	Cavity Identification	**Fault Classification**
ML Model (%)	88.0	86.9
DL Model (%)	87.8	81.3

Note that the classification task exploits both pre-fault and post-fault data to achieve the model accuracy shown in Table 1. However, a fault prediction model can only utilize data prior to a fault event to make predictions on an impending fault. Moreover, fault prediction must be performed at the cavity level as prediction of a fault would inherently identify the associated cavity. We therefore conduct the feasibility study for fault prediction as a two-step process as follows: 1) we define a binary classification task to identify waveforms describing impending faults, and 2) a moving window based multi-class classification task to predict the type of failure before onset.

BINARY CLASSIFICATION TASK

Binary classification task is designed to investigate the possibility of distinguishing waveforms representing imminent faults from waveforms representing stable running conditions.

Dataset

The dataset utilized for this study consists of a total of 5,047 fault events. Table 2 summarizes the dataset composition with respect to fault types.

Table 2: Fault Prediction Dataset

Fault Type	**Number of Events**
Single Cavity Turn Off	885
Microphonics	710
100 ms Quench	608
Controls Fault	847
Electronic Quench	673
3 ms Quench	542
Heat Riser Choke	720
Unknown	62

The current data acquisition system is set up to collect data only in the event of a RF cavity trip (see Figure 2). Note that approximately 94% of the captured waveform (t < 0) represents pre-fault activity, with t = 0 the fault onset, and t > 0 the post-fault data. Therefore, we utilize 100 ms segments extracted from the waveforms to represent stable running and impending fault classes as follows: for stable running, extract a 100 ms segment from the earliest possible window of the captured waveform (annotated as “Stable running” in Figure 2, with a window of [−1536 ms, −1436 ms]). Segments of this region are verified by SMEs as sufficient representations of stable running conditions. For the impending fault, we extract a 100 ms window as close as possible to the fault onset (annotated as “Impending fault” in Figure 2). Accordingly, we extract impending fault segments with a window of [−105 ms, −5 ms], with a 5 ms buffer to ensure that fault onset activity is not inadvertently captured. As a result, for the binary classification task we obtain 5,047 examples. Additionally, we retain labels of the underlying fault type in the waveforms representing the impending fault phase to perform a follow-up analysis on the feasibility of predicting fault types.

Prediction Model and Training

We develop a deep recurrent neural network (DRL) architecture to process the time-series for the binary classification task. The detailed architecture is shown in Figure 3.

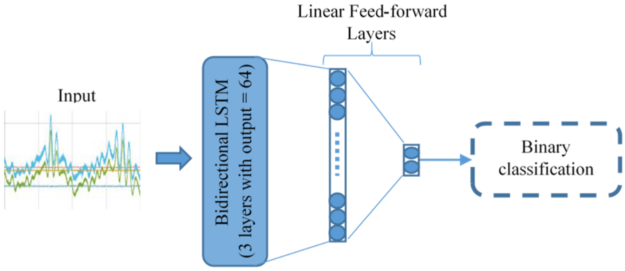
Linear Feed-forward Layers
Input
Bidirectional LSTM (3 layers with output = 64)
Binary classification

Figure 3: DRL architecture for binary classification of pre-fault data.

The 100 ms segments extracted for this study represents time-series data with each 500 time steps long. The Long Short Term Memory (LSTM) [5] layers in the above model processes the time-series input, and the corresponding features are classified at the last layer (softmax activation). The architecture and the training scheme is developed using the PyTorch deep learning library [6]. The data is divided with fault based stratification to obtain 60% for training, 20% for validation, and the remaining 20% for testing.

Binary Classification Results

The confusion matrix obtained from the testing data is shown in Figure 4. The overall testing data classification accuracy is 74.74%. Detailed testing results are summarized in Table 3. However, the confusion matrix shows a distinct issue in the models ability to recognize many impending fault events correctly, highlighted by the large false negative percentage (23.48%). This is also observed in Table 3 where the recall rate for impending fault, and the precision for stable running classes are both low.

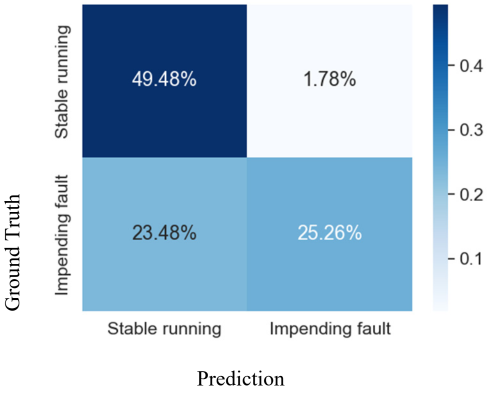

Figure 4: Confusion matrix for classification of stable running versus impending fault.

Table 3: Binary Classification Performance

Class	Precision	Recall	F1-score
Stable running	67.8%	96.5%	79.6%
Impending fault	93.4%	51.8%	66.7%

This behavior indicates that certain impending fault segments may closely resemble stable running conditions prior to fault onset, to a degree that the model is unable to distinguish between the two classes. We speculate that certain fault types identified by SMEs may not present sufficient precursors within the waveforms we use for the analysis. In order to further investigate this issue, we take a closer look at the fault types that represent the false negatives. Figure 5 shows the distribution of false negative events in terms of the ground truth fault type.

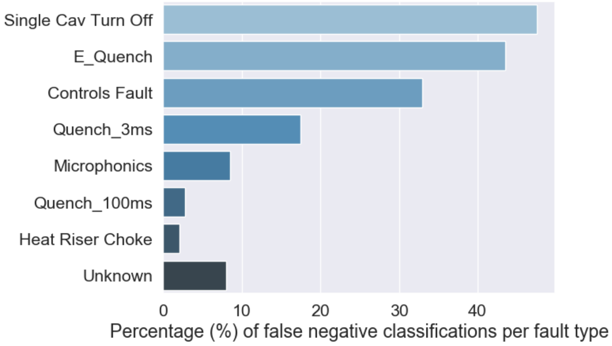

Figure 5: Percentage of events with false negative classification according to fault type.

It is evident from Figure 5 that the false negatives are dominated by events from ‘Single Cav Turn Off’, ‘E_Quench’, and ‘Controls Fault’ fault types. Upon discussion with domain experts, it was confirmed that these fault types oftentimes require ancillary data for correct identification due to the lack of precursors present in the waveforms used for the study. Therefore we conducted a secondary experiment where we discard events from the three aforementioned fault types and obtain a secondary dataset of 2,642 fault events. We use the same model architecture, and follow the same training procedure. The model achieves a classification accuracy 92.1% for this dataset.

FAULT TYPE PREDICTION TASK

Given encouraging results from the binary classification task, we investigate the possibility of predicting the fault type before the onset of beam trip. In order to quantify the ability to predict the fault type with respect to the onset, we define a window-based analysis on the waveform data that we have collected. Specifically, a window of a specific

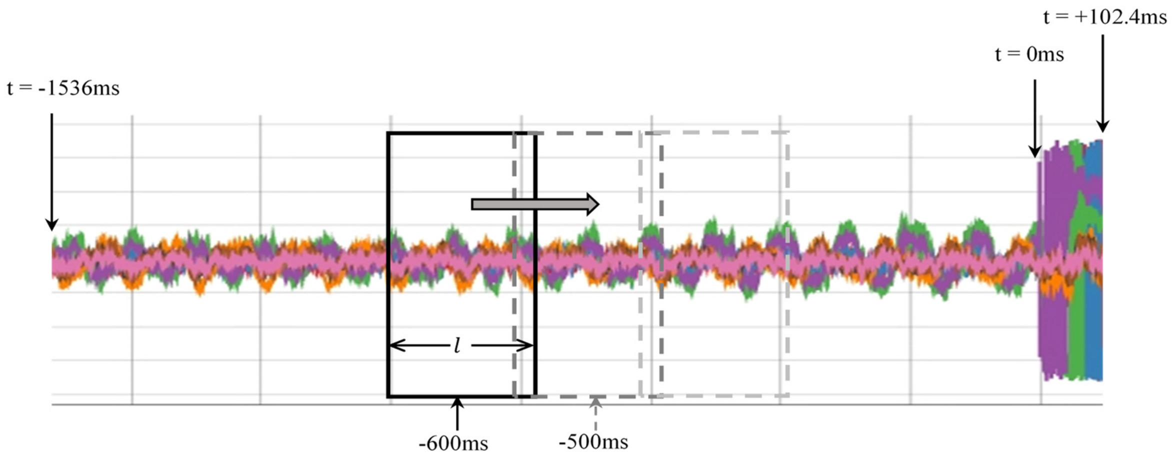

Figure 6: Window based fault prediction analysis scheme.

length l is positioned over the waveform data, centered on a time stamp t before the onset of the fault, as illustrated in Figure 6. The data that falls within the window is extracted to perform analysis on fault prediction. This is repeated over window locations that incrementally move from t = –600 ms to 0 ms in 50 ms intervals. The value of –600 ms is chosen as the furthest time stamp prior to the fault event based on suggestions of domain experts, and 0 ms denotes the fault event onset.

Dataset

We utilize the same dataset described in Table 2 as our base dataset. The window-based analysis in Figure 6 results in multiple experiments conducted over each window location. In essence, for a given window location t, data that falls within the window is extracted from all examples to create training, validation, and testing sets specific to location t. Analysis using this data yields fault prediction performance for t ms prior to fault onset. In addition to multiple window locations, we also experiment with the effect of window size on prediction performance using lengths of l = 100 ms, 200 ms, 300 ms.

Prediction Model and Training

We utilize the same general deep LSTM architecture shown in Figure 3 for this analysis. However, fault type prediction is essentially a multiclass classification task. Therefore, we augment the classification layer of the network to accommodate 8 fault classes. The same training and testing procedures used in the binary classification task are leveraged for each window dataset. Model development, training, and testing is performed using the PyTorch library [6].

Fault Type Prediction Results

Although the prediction network is trained and tested for each window experiment in a multiclass manner, we inspect the prediction performance for each class separately. This allows for observing the prediction efficacy for each fault more efficiently, considering the insights obtained through the binary classification task. Figure 7 shows performance of the prediction model plotted as a function of time prior to onset of “Microphonics”, and “E_Quench” fault types. The horizontal axis of the plots shows the center of the window used to extract data for the prediction task. For instance,

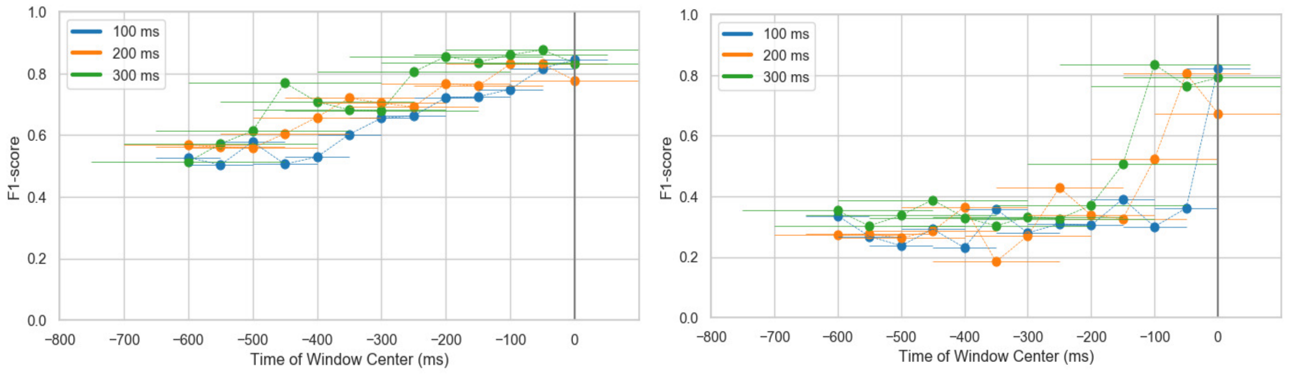

Figure 7: F1-score plots for the window based fault type prediction of microphonics and electronic quench fault types.

ISBN: 978-3-95450-221-9 ISSN: 2226-0358 doi:10.18429/JACoW-ICALEPCS2021-WEPV025

the F1-score at –600 ms for a given fault type indicates the model's ability to predict the fault type 600 ms prior to the fault onset. We also observe that the predictive power of the model is highly dependent of the type of fault. For instance, the model is able to predict microphonics faults with over 0.5 F1-score from –600 ms with significant improvements with window locations closer to onset. Conversely, the model is unable to predict E_Quench events with acceptable accuracy until the window overlaps the fault onset. Figure 8 shows a snapshot of the fault dependent prediction performance for the case of 300 ms data window centered at –200 ms (200 ms before the fault onset).

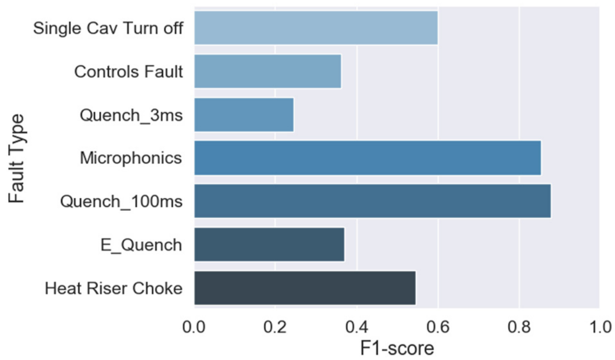

Figure 8: Fault specific model prediction performance for the case of 300 ms data window centered at –200 ms.

This analysis further demonstrates that certain faults such as microphonics show precursors prior to onset, which can be used to predict the faults prior to a trip. The analysis also provides insights into several fault types that may not show precursors due to their nature, or may not be sufficiently captured in the waveforms used in this study.

FUTURE WORK

We have discussed our initial studies into the possibility of automated prediction of RF faults in C100 cavities using ML models. The investigations have yielded promising results, with insights into the possible prediction time ranges for multiple fault types. However, this study leverages static RF fault data. An effective fault prediction framework would require a data acquisition system that is able to stream the data signals in real-time during CEBAF operation. The C100 modules, and the associated LLRF systems, are currently undergoing a firmware upgrade to allow the collection of streaming data. This will provide valuable information regarding variations encountered in stable running conditions, and possible non-RF related fault events in CEBAF.

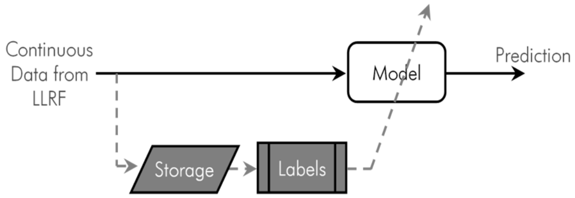

Figure 9: Preliminary framework for using machine learning models with continuously streaming data.

Streaming data provides an opportunity to further improve and adapt the machine learning models to predict faults before they occur. To that end, we envision a data processing framework as shown in Figure 9. According to the preliminary framework, the prediction model gets direct access to the continuous data to make predictions on stable running versus impending fault, with probable fault type in real-time. The framework also allows storing select subset of the data for the purposes of routine model performance evaluation, developing new models, re-training and fine-tuning of the deployed models. We also envision the C100 firmware upgrades and the proposed streaming data processing framework will generate significant amounts of information-rich data that may be useful to diagnose other CEBAF machine events beyond SRF cavity faults.

REFERENCES

[1] C. E. Reece, "Continuous wave superconducting radio frequency electron linac for nuclear physics research," *Physical Review Accelerators and Beams,* vol. 19, no. 12, p. 124801, 12/28/ 2016, doi: 10.1103/PhysRevAccelBeams.19.124801

[2] A. Solopova *et al.*, "SRF cavity fault classification using machine learning at CEBAF," in *10th Int. Particle Accelerator Conf.(IPAC'19), Melbourne, Australia, 19-24 May 2019*, 2019: JACOW Publishing, Geneva, Switzerland, pp. 1167-1170, doi:10.18429/JACoW-IPAC2019-TUXXPLM2

[3] C. Tennant, A. Carpenter, T. Powers, A. Shabalina Solopova, L. Vidyaratne, and K. Iftekharuddin, "Superconducting radio-frequency cavity fault classification using machine learning at Jefferson Laboratory," *Physical Review Accelerators and Beams,* vol. 23, no. 11, p. 114601, 11/30/ 2020, doi: 10.1103/PhysRevAccelBeams.23.114601

[4] L. Vidyaratne, T. Powers, C. Tennant, K. M. Iftekharuddin, Md M. Rahman and A. Shabalina, "Deep Learning-Based Superconducting Radio-Frequency Cavity Fault Identification at Jefferson Laboratory," *Frontiers Operational Intelligence (under review),* 2021.

[5] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation,* vol. 9, no. 8, pp. 1735-1780, 1997.

[6] A. Paszke *et al.*, "Pytorch: An imperative style, high-performance deep learning library," *Advances in neural information processing systems,* vol. 32, pp. 8026-8037, 2019.