# Real-Time Hierarchical Neural Network Based Fault Detection and Isolation for High-Speed Railway System Under Hybrid AC/DC Grid

Qin Liu, *Student Member, IEEE*, Tian Liang, *Student Member, IEEE*, and Venkata Dinavahi, *Fellow, IEEE*

***Abstract*—Reliable and comfortable high-speed railway (HSR) has skyrocketed in popularity as a transportation medium for traveling around the world. High-voltage direct current (HVDC) electrification system has been introduced to the HSR gradually. However, the coexistence of AC and DC systems will last for a long time because AC railway systems are still in the dominant position. A detailed HSR traction system transient model operating under the hybrid AC/DC grid was established in PSCAD/EMTDC. We proposed a real-time fault detection and isolation (FDI) method for the simulated model using neural network (NN). Hierarchical structure of the monitoring system has been employed. Low-level sub-monitors supervised the conditions of their local regions and the top-level monitor collected all the feedback from sub-monitors making the final evaluation of the entire HSR system based on a voting strategy. Both off-line and real-time experiments were conducted to validate the effectiveness of the proposed method. In the experiments, the sub-monitors were designed based on Gated Recurrent Unit (GRU) algorithm and implemented on the Xilinx VCU128 FPGA board. For the off-line experiment, the sub-monitors used the training and testing dataset both from PSCAD/EMTDC to construct the architecture of their individual GRU networks and to verify how great the networks can be. For the real-time task, the sub-monitors interfaced with a real-time HSR system emulator running on the Xilinx VCU118 FPGA board to test the performance in the real-time application. The results proved that our proposed FDI method has the capability of real-time detection and can achieve better accuracy within reasonable time and resource consumption than other NN-based methods. Moreover, the method was capable of standing against noises from measured signals to some extent.**

***Index Terms*—Data-driven methods, fault detection and isolation (FDI), field-programmable gate array (FPGA), Gated Recurrent Unit (GRU), high-speed railway (HSR), high-voltage direct current (HVDC), hybrid AC/DC grid, long short-term memory (LSTM), machine learning, neural networks, real-time systems.**

Manuscript received December 10, 2019; revised May 7, 2020 and July 21, 2020; accepted September 5, 2020. Date of publication September 8, 2020; date of current version December 4, 2020. This work was supported by the Natural Science and Engineering Research Council of Canada (NSERC). The work of Qin Liu was supported by China Scholarship Council (CSC). Paper no. TPWRD-01417-2019. *(Corresponding author: Tian Liang.)*

Qin Liu is with the School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing 100044, China, and also with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada (e-mail: qin9@ualberta.ca).

Tian Liang and Venkata Dinavahi are with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada (e-mail: tian.liang@ualberta.ca; dinavahi@ualberta.ca).

Color versions of one or more of the figures in this article are available online at https://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TPWRD.2020.3022750

## I. Introduction

RAILWAY electrification systems have been in great expansion during the last decade, especially for the high-speed railway (HSR). In order to match the high demand for the HSR, the alternating current (AC) railway electrification system is commonly used in the world. An alternative is the direct current (DC) electrification system. However, the voltage level was limited to 1.5 kV and 3 kV [1] due to insulation constraints. With the rapid growth of advanced power electronic devices and material technology, higher voltage DC system has been brought to public attention, because it is able to lower losses, reduce land and material usage, improve power quality and consequently lead to lower cost compared with conventional AC system [2], [3].

Nowadays, AC electrification systems are the predominant solution for the HSR worldwide [1]. It will still take many years to unify various power supply systems into one standard form, i.e., coexistence of AC and DC systems is inevitable for a long time. Therefore, hybrid vehicles that can be energized by both AC and DC supplies are promising. ABB [4] and Bombardier [5] have provided commercially available vehicle prototypes for this kind of hybrid AC/DC application.

Safety is always a priority for railway transportation. Any fault may cause an irreversible disaster, so real-time fault detection and isolation (FDI) for HSR system is crucial. For an electrification system powered by hybrid AC/DC grid, situations are much more complicated. Aside from the common faults only existing in either AC or DC system, they can also be the result from their interaction. In AC systems, the ground fault on lines [6], open-circuit fault in converters [7] as well as internal faults in transformers [8], sensors [9] and motors [10] have been researched. While defects on lines [11] and gain fault in sensors [12] for the DC system have been discussed. In [13], the negative mutual impact between AC and DC system under parallel operation has been analyzed. The issue of mutual disturbances results in higher failure rate in the hybrid system.

Many researchers have paid attention to the FDI of railway systems. Model-based methods, such as sliding mode observer [9], state estimator [7], descriptor estimator [10] and consistency-based combining with abductive method [14], are proposed to deal with fault detection issues in railway electrification systems. However, the success of using these model-based methods highly depends on the expert knowledge base owned

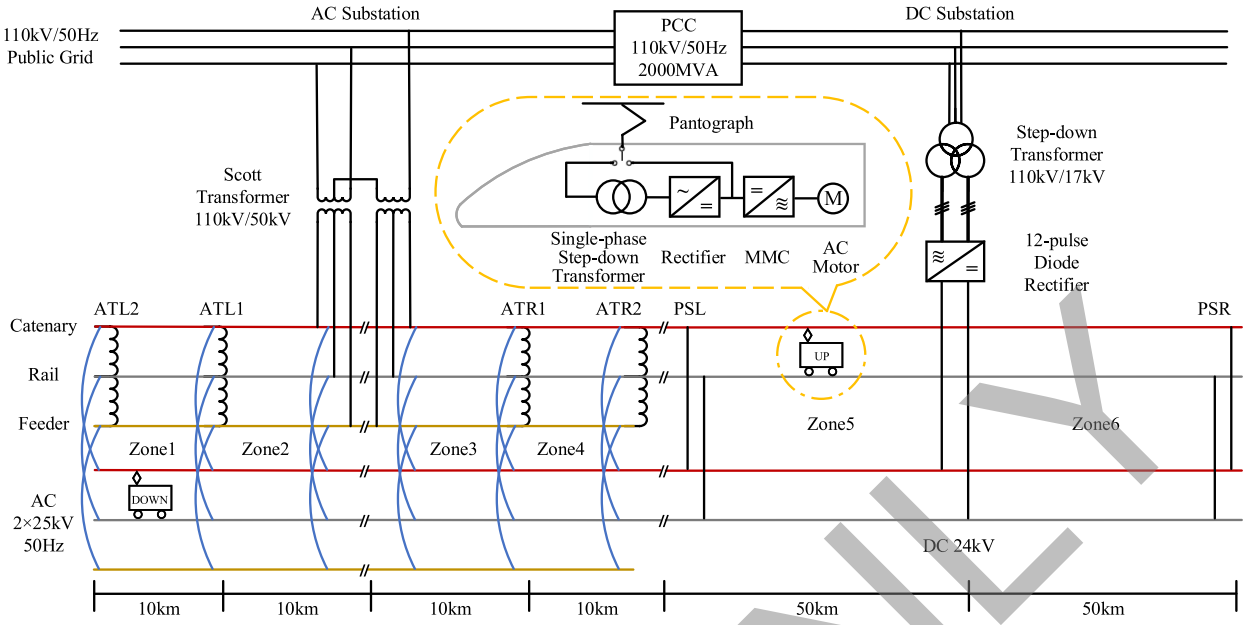


Fig. 1. Schematic of high-speed railway system under hybrid AC/DC grid.

by the designer. The model should truly and accurately express the physical laws of the object considering uncertainties and disturbances, then the fault could be diagnosed correctly from the real-world system. Another option for the FDI is data-driven approaches. Instead of creating the physical model, they collect a mass of historical data from real systems that is easily achievable in this era of big data, and then they can learn the inherent laws themselves based on the data. Machine learning especially neural network highly promotes the development of data-driven methods. A method combining kernel principal component analysis with random forest is introduced to analyze the dissolved gases from transformers to detect the abnormal condition [15]. In order to monitor the condition of the wheelset bearing, a novel deeper 1D convolutional neural network is adopted by [16]. Multiple defects of track circuits has been diagnosed by LSTM-based network in [17]. Nevertheless, there are rare researches working on monitoring the entire railway electrification system including power grid, traction substations (TSSs) and section posts (SPs), transmission lines and rolling stock, let alone the whole HSR system operating under hybrid AC/DC grid.

In this paper, we establish an HSR system model that powered by both AC and DC grid in off-line PSCAD/EMTDC and on a FPGA-based real-time system emulator, which provide the data for training and testing our method. A hierarchical Gated Recurrent Unit (GRU) based FDI method is proposed. GRU algorithm is a variant of the Long Short-Term Memory (LSTM) network that can mitigate short-term memory in long sequences confronted by Recurrent Neural Network (RNN). Furthermore, hierarchical structure of the monitoring system is created because of the large scale of the HSR system. Finally, algorithm acceleration is implemented on the field programmable gate array (FPGA) to be fully realized with a low-latency real-time capability.

The remainder of this paper is organized as follows. Section II briefly describes the model and conditions of the HSR system powered under the hybrid AC/DC grid. Section III elaborates the proposed hierarchical GRU-based FDI method, including algorithm preliminary, hierarchical structure, configuration of neural networks and hardware implementation. Section IV presents the discussion of the experimental results including data sampling approach, hierarchical structure, hardware implementation, data with noises and different neural networks, followed by the conclusions in Section V.

## II. Topology of HSR Under Hybrid AC/DC Grid

We establish a transient model of the HSR system operating under hybrid AC/DC grid on PSCAD/EMTDC in the off-line task and on an FPGA-based emulator in the real-time task. The modeled track section is supplied by two traction substations (TSSs), one is AC TSS and the other is DC TSS. They are connected to the same node of a power system (points of common coupling, PCC) in parallel, which provides three-phase 110 kV/50 Hz public voltage, as shown in Fig. 1. The total length of this track section is 140 km and the two substations are 70 km apart. For simplicity, a common double-track railway network and one and only one vehicle on each track are considered in this paper.

### A. Model Description

The voltage for the AC electrification system is $2 \times 25$ kV/50 Hz which is adopted in China railway network. An 60 MVA AC traction substation is equipped with a Scott-type transformer to decrease the voltage unbalance generated by the single-phase traction load. The transformer convert the three-phase 110 kV/50 Hz voltage of the public network to two-phase 50 kV/50 Hz voltage. Each phase supplies one of the

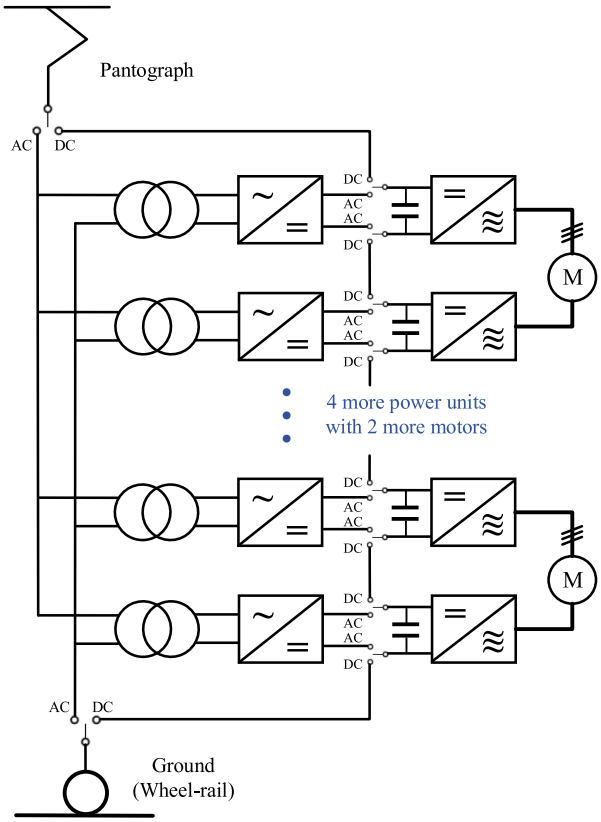


Fig. 2. Switch scheme for the vehicle operating under hybrid AC/DC grid.

two adjacent sectors and the sectors are electrically insulated by a neutral section. There are two 15 MVA autotransformers (ATs) that are evenly distributed every 10 km in one sector to increase transmission distance, reduce induced interference and cost. The AT is a signal-winding transformer connected between the catenary (the overhead contact wire) and the return feeder. The mid-point of the winding is connected to the rail (ground). With a transformation ratio of 1:1, it splits the 50 kV output voltage into two 25 kV halves, one from catenary to rail and the other from rail to feeder [18].

On the DC substation, a step-down transformer reduces the public grid voltage to 17 kV/50 Hz. The diode rectifier with 12 pulses is adopted to obtain 24 kV DC overhead line voltage and meanwhile to reduce negative impact of the TSS on the power network [13]. No neutral section is needed for DC system which substantially increases the power supply distance. The ATs are replaced by paralleling stations (PSs) that are distributed every 50 km along the line to reduce the overall circuit resistance [3].

A power unit of the rolling stock is equipped with a single-phase step-down transformer, a AC/DC rectifier and a five-level modular multilevel converter (MMC). Two converters in parallel supply an AC traction motor. We assume a vehicle with a rated power of 8 MW has eight power units supplying four motors. In order to adapt the hybrid system, the switches are also installed. An explicit switch scheme for the hybrid vehicle is displayed in Fig. 2. When AC power is available, transformers and rectifiers are activated in parallel to provide a stable DC link voltage, while the MMCs are supplied directly into the DC link in series, when DC power is available.

TABLE I
SELECTED CONDITIONS IN THE MODELED HSR SYSTEM

| Location | Descriptions | Conditions |
|---|---|---|
| None | No fault | Normal |
| AC substation | Phase-to-phase short-circuit in Scott transformer | Fault1 |
| ATs | High resistance grounding at ATL1 | Fault2 |
| | High resistance grounding at ATL2 | Fault3 |
| | Short-circuit at ATR1 | Fault4 |
| | Short-circuit at ATR2 | Fault5 |
| AC line | Ground fault on catenary at zone2 | Fault6 |
| | Ground fault on feeder at zone2 | Fault7 |
| | Ground fault on catenary at zone1 | Fault8 |
| | Ground fault on feeder at zone1 | Fault9 |
| | Ground fault on catenary at zone3 | Fault10 |
| | Ground fault on feeder at zone3 | Fault11 |
| | Ground fault on catenary at zone4 | Fault12 |
| | Ground fault on feeder at zone4 | Fault13 |
| DC substation | Phase-to-phase short-circuit in step-down transformer | Fault14 |
| | Open-circuit in transistors of rectifier | Fault15 |
| Paralleling station | Short-circuit at PSL | Fault16 |
| | Short-circuit at PSR | Fault17 |
| DC line | Ground fault on catenary at zone5 | Fault18 |
| | Ground fault on catenary at zone6 | Fault19 |
| Vehicle on up track | Short-circuit in transformer | Fault20 |
| | Open-switch in rectifier | Fault21 |
| | Open-switch in MMC | Fault22 |
| | Phase-to-phase short-circuit in motor | Fault23 |
| | Gain fault in speed sensor | Fault24 |
| Vehicle on down track | Short-circuit in transformer | Fault25 |
| | Open-switch in rectifier | Fault26 |
| | Open-switch in MMC | Fault27 |
| | Phase-to-phase short-circuit in motor | Fault28 |
| | Gain fault in speed sensor | Fault29 |

### B. Selected Conditions

Thirty conditions, consisting of 1 normal and 29 faults that may occur in transformers and converters of TSSs, ATs, transmission lines and components of vehicles, are selected to be monitored in our simulated HSR system and summarized in Table I. We divide the track section into 6 zones, as illustrated in Fig. 1. The vehicle on each track can operate in any one of the zones, and all of the 36 location permutations for the two vehicles are regarded as scenarios in a condition to ensure the evaluation is not associated with the load location.

## III. PRELIMINARIES AND PROPOSED METHOD

In this section, a hierarchical GRU-based FDI approach is elaborated. Instead of one versatile monitor collecting all the signals from sensors spread all over the 140 km track section, we design multiple sub-monitors on key nodes of the traction

network. Each monitor only requires to manage sampled data in a small range, and give immediate feedback to the upper-level monitor. While the top-level monitor on the PCC can communicate with lower-level monitors on TSSs and obtain the evaluations from all sub-monitors efficiently to make a final decision on what condition of the entire system it could be. There are two stages to build a sub-monitor: off-line construction and real-time monitoring. In off-line construction stage, the optimized configuration of the GRU network for each sub-monitor is learned from the dataset from PSCAD/EMTDC. Then the FPGA is used to implement the network for supervising the condition of its local region in real time.

### A. Gated Recurrent Unit Network

The long short-term memory (LSTM) network, a branch of RNN, is adopted in the FDI of the HSR system, since it is able to learn from the long-term historical information of the data in time series, namely, overcome the vanishing or exploding gradient problem that conventional RNN suffers from. Thanks to the "gates" mechanism in LSTM, it has the ability to figure out how important the data in a sequence is and decide to keep or throw away. The standard LSTM is expressed as following:

$$\begin{aligned}
\boldsymbol{f}_t &= \sigma\left(\boldsymbol{w}_f \cdot [\boldsymbol{x}_t, \boldsymbol{h}_{t-1}] + \boldsymbol{b}_f\right) \\
\boldsymbol{i}_t &= \sigma\left(\boldsymbol{w}_i \cdot [\boldsymbol{x}_t, \boldsymbol{h}_{t-1}] + \boldsymbol{b}_i\right) \\
\boldsymbol{o}_t &= \sigma\left(\boldsymbol{w}_o \cdot [\boldsymbol{x}_t, \boldsymbol{h}_{t-1}] + \boldsymbol{b}_o\right) \\
\tilde{\boldsymbol{c}}_t &= \tanh\left(\boldsymbol{w}_c \cdot [\boldsymbol{x}_t, \boldsymbol{h}_{t-1}] + \boldsymbol{b}_c\right) \\
\boldsymbol{c}_t &= \boldsymbol{f}_t \circ \boldsymbol{c}_{t-1} + \boldsymbol{i}_t \circ \tilde{\boldsymbol{c}}_t \\
\boldsymbol{h}_t &= \boldsymbol{o}_t \circ \tanh\left(\boldsymbol{c}_t\right)
\end{aligned} \tag{1}$$

where the current input $\boldsymbol{x}_t$ and the previous hidden state $\boldsymbol{h}_{t-1}$ are concatenated to form a vector and the output is the new hidden state $\boldsymbol{h}_t$. The cell state $\boldsymbol{c}_t$ and various gate layers (i.e. forget gate $\boldsymbol{f}_t$, input gate $\boldsymbol{i}_t$ and output gate $\boldsymbol{o}_t$ respectively in equation (1)) are the core concept in LSTM, as it will be illustrated below. $\tilde{\boldsymbol{c}}_t$ is new memory content that is regulated by the tanh activation transforming values between -1 and 1. Weights $\boldsymbol{w}$ and biases $\boldsymbol{b}$ are the optimized parameters during training the network. "." is matrix product and "$\circ$" is hadamard product or point-wise multiplication.

Gate layers contain the sigmoid activation which squishes values passing through between 0 and 1. A value of zero means "completely forget it" because any number multiplied by 0 is 0, while a value of one means "completely keep it" since the number stays the same after getting multiplied by 1. This is the way the gates learn the importance of information from data and add or remove them to or from the cell state and output. Three gates each is charged with their distinct responsibility. The forget gate decides what to keep or forget from previous steps. The input gate decides what should be kept to update the cell state from the current step. The output gate decides what information the new hidden state should carry over to the next step.

There have been numerous variants of LSTM. A newer generation of them is the Gated Recurrent Unit (GRU) [19]. Different

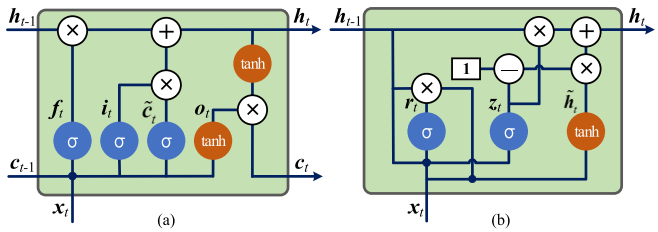


Fig. 3. Cell structures of standard LSTM and GRU. (a) LSTM. (b) GRU.

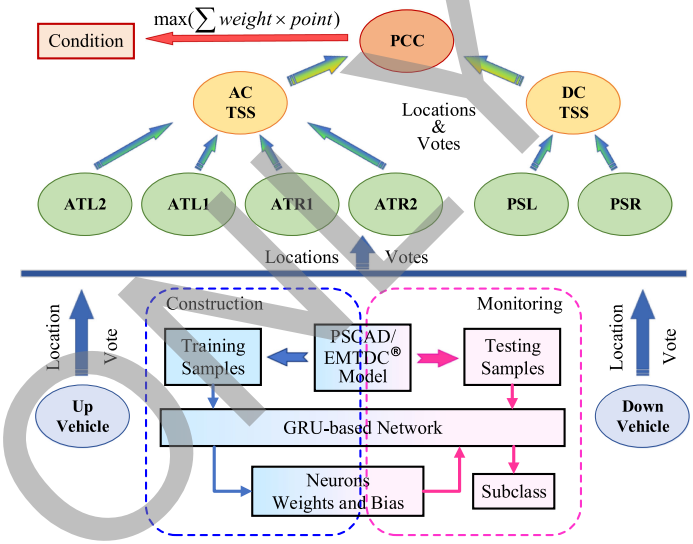


Fig. 4. Hierarchical structure of the monitoring system for FDI in HSR used in the off-line experiment.

from the standard LSTM, the forget and input gates merge into a single update gate $\boldsymbol{z}_t$ and a reset gate $\boldsymbol{r}_t$ is added in GRU. It also gets rid of the cell state and only use the hidden state to transfer information, as shown below:

$$\begin{aligned}
\boldsymbol{z}_t &= \sigma\left(\boldsymbol{w}_z \cdot [\boldsymbol{x}_t, \boldsymbol{h}_{t-1}] + \boldsymbol{b}_z\right) \\
\boldsymbol{r}_t &= \sigma\left(\boldsymbol{w}_r \cdot [\boldsymbol{x}_t, \boldsymbol{h}_{t-1}] + \boldsymbol{b}_r\right) \\
\tilde{\boldsymbol{h}}_t &= \tanh\left(\boldsymbol{w}_h \cdot [\boldsymbol{x}_t, \boldsymbol{r}_t \circ \boldsymbol{h}_{t-1}] + \boldsymbol{b}_h\right) \\
\boldsymbol{h}_t &= \boldsymbol{z}_t \circ \boldsymbol{h}_{t-1} + (\boldsymbol{1} - \boldsymbol{z}_t) \circ \tilde{\boldsymbol{h}}_t
\end{aligned} \tag{2}$$

Our sub-monitors for the FDI in HSR are created based on the GRU algorithm. We also compare the performances between these two types of LSTM in terms of accuracy, runtime and hardware resource utilization on the FPGA. Both of the two LSTM-cell structures are depicted in Fig. 3.

### B. Hierarchical Structure

As our aforementioned model, the entire system covers 140 km of high-speed tracks. It is impractical to collect the signals from sensors spread all over the system to a centralized monitor, let alone any track section between two destinations in real world. Therefor, we design a hierarchical monitoring system as presented in Fig. 4 where each sub-monitor can focus on signals and conditions of partial system in a small range. It

TABLE II
NUMBER OF LAYER NEURONS FOR DIFFERENT SUB-MONITORS

| | AC TSS | AT1 | AT2 | DC TSS | PS | Vehicle |
|---|---|---|---|---|---|---|
| **Features** | 12 | 6 | 5 | 17 | 4 | 13 |
| **a** | 4 | 8 | 16 | 8 | 8 | 16 |
| **b** | 8 | 16 | 8 | 8 | 16 | 8 |
| **Subclasses** | 6 | 9 | 6 | 9 | 6 | 11 |

follows the hierarchical nature of the electrification system of the HSR system. In this kind of topology, the monitoring system for the long track section is naturally separated into small pieces by sub-monitors spreading in various components of the system. The evaluations of sub-monitors and the vehicles' locations are the only information required to be passed through upper-level monitors. While a top-level monitor is in charge of giving a summary of all the sub-monitors in a larger range.

The monitor on the vehicle is the lowest-level one, which samples the data of voltages, currents, speeds and torques from vehicle's components. The monitor makes a judgment about whether there is anything wrong in the vehicle itself or in the zone the vehicle is within. Then the output of its location and evaluation is immediately fed back to the upper-level monitor on ATs and PSs. The voltages and currents going into and getting out from ATs and PSs respectively are inputs for these monitors. They are able to estimate the nearby zones' condition and report the conclusions of vehicles and themselves to the monitors on the AC or DC TSS. The TSS monitor can speculate the condition of the sectors covered by the independent TSS using the voltages and currents collected in the TSS. Our proposed GRU-based method is applied to build all the sub-monitors. The off-line construction for each one obtains a distinct configuration of the network including neurons for each layers and parameters of weights and bias. Then every one is capable to monitor the condition of its local area in real time by implementation on the FPGA.

The top-level monitor on the PCC aggregates all the evaluations from sub-monitors and makes the final decision for the entire track section. For the classification, a voting strategy is used. Any sub-monitor has its own classification (called subclasses to distinguish from the overall 30 classes). For instance, the sub-monitor in the vehicle reclassifies the 30 classes into 11 subclasses, as shown in Table II. Each sub-monitor is regarded as a voter who votes for a subclass. One vote counts as one point otherwise as zero. The score for a designated class is calculated by the sum of weighted points for subclasses that cover the class. A weight value is assigned to each vote according to the correlation between the voter and the class: the closer the larger. In the end, the votes from all the sub-monitors point to a class with the maximum number of the score.

### C. Network Construction

We construct sub-monitors using TensorFlow [20] that is an open-source library for machine learning developed by Google Brain team. It provides plenty and powerful Python-based high-level Application Programming Interfaces (APIs) for any-level users who intend to explore and excavate machine learning in various fields. It is flexible enough for researchers to experiment with diverse algorithms, network architectures and optimization methods.

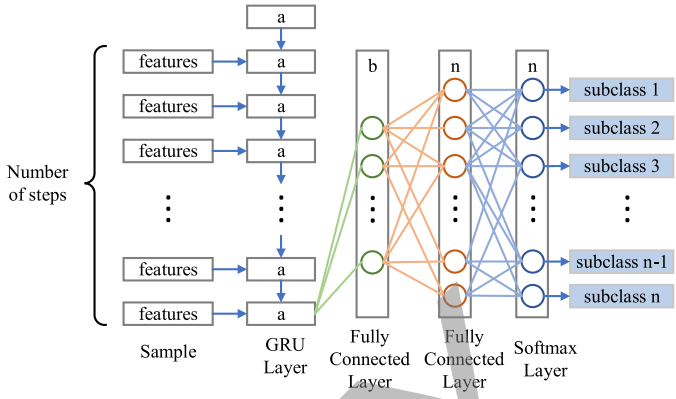


Fig. 5. Architecture of GRU-based network in sub-monitors for FDI in HSR.

In our HSR study case, GRU algorithm is elected as a winner for the FDI by trial and error using TensorFlow. The architecture of the model is designed to be simple but effective since the monitor is expected as a real-time application. The parameters of the model are somehow different for each sub-monitor on the basis of the data. The original data are normalized and reshaped into a two-dimensional matrix. In the GRU layer, cells are stacked one by one as the sequence chain and one cell processes one-time-step data. The current input and the last step hidden state are passed into a GRU-cell to calculate the next step hidden state (mathematically expressed as (2)). After iterating step-by-step, the last hidden state serves as the input to a fully connected layer filtered by Rectified Linear Unit (ReLU) activation function. The output layer is a fully connected layer directly connected to a softmax classifier. The Adam algorithm [21] is the optimizer employed to acquire optimal parameters for the network. The framework of GRU-based network and the number of layer neurons for sub-monitors are respectively shown in Fig. 5 and Table II.

### D. Implementation of Monitoring System on FPGA

The optimized GRU-based network is implemented on the FPGA to guarantee that the monitors can make reasonable evaluations of the condition of the HSR system in real time. Owing to the highly parallel capability of FPGA, the execution time of the application based on our proposed method slashed dramatically.

The real-time monitoring system is implemented on the Xilinx Virtex UltraScale+$^{TM}$ HBM VCU128 FPGA board [22], which provides the hardware environment for designs targeting the Virtex UltraScale+$^{TM}$ XCVU37P HBM FPGA with 2,852,000 programmable logic cells, 340.9Mb BRAM, and 9,024 DSP slices.

As mentioned above, the GRU layer is an iteration process of one-time-step algorithm. The detailed hardware implementation of the one-time-step is presented in Fig. 6. The vertically aligned

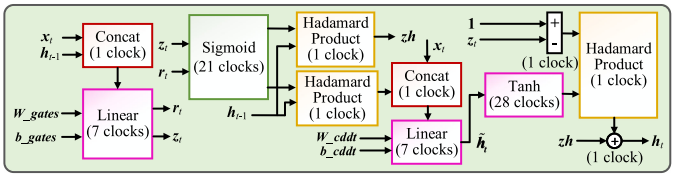


Fig. 6. Framework of GRU-based network hardware implementation for one-time-step.

operations can be executed in parallel and the latency for the operations are also displayed. The fully connected layer works as a linear function.

Additionally, in order to achieve further acceleration, we take advantage of the fixed-point calculation in the FPGA. The 24-bit word length with 18 fractional bits is chosen as the data format by trial and error aiming at minimize the runtime and resources on the FPGA without sacrificing accuracy.

Operation units shown in Fig. 6 are summarized as follows. Two arrays are merged by the concat unit:

$$\text{concat}\left(\boldsymbol{x},\boldsymbol{h}\right)=\begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{h}\end{bmatrix}. \tag{3}$$

Linear combination is calculated by the linear unit:

$$\text{linear}\left(\boldsymbol{x}\right)=\boldsymbol{w}\cdot\boldsymbol{x}+\boldsymbol{b}. \tag{4}$$

Activation functions like ReLU, sigmoid and tanh are used to regulate the values flowing through neural networks:

$$\text{Relu}(x)=\begin{cases}x & \text{if } x>0\\ 0 & \text{if } x\leq 0\end{cases}, \tag{5}$$

$$\text{sigmoid}(x)=\frac{1}{1+e^{-x}}, \tag{6}$$

$$\tanh(x)=\frac{2}{1+e^{-2x}}-1. \tag{7}$$

The elements of two matrices $\boldsymbol{A}$, $\boldsymbol{B}$ with the same dimension after the hadamard product are expressed as:

$$(\boldsymbol{A}\circ\boldsymbol{B})_{i,j}=(\boldsymbol{A})_{i,j}(\boldsymbol{B})_{i,j}. \tag{8}$$

The equation for the $i$th element of a vector $\boldsymbol{V}$ passing through the softmax operation is given by:

$$\text{softmax}\left(\boldsymbol{V}_i\right)=\frac{e^{\boldsymbol{V}_i}}{\sum_j e^{\boldsymbol{V}_j}}. \tag{9}$$

## IV. Experimental Results and Discussion

We conduct sufficient experiments to verify the effectiveness of our proposed method. As summarized in Section II B, 30 conditions each with 36 scenarios are taken into consideration. For the off-line experiment, both of the training and testing samples are collected from the model established in PSCAD/EMTDC and the details of data sampling are elaborated in this section. For the real-time task, the monitoring system directly interfaces with a real-time HSR system emulator running on the Xilinx VCU118 FPGA board. The advantages of our proposed hierarchical GRU-based FDI method for HSR system are demonstrated through comparative experiments with respect to the structure, implementation, noise resistance and neural networks.

### A. Data Sampling From PSCAD/EMTDC

We assume that the vehicle running at the speed of 300 km/h. A time-sequential sample is collected within 0.02 s at a sampling frequency of 5 kHz and the evaluation of the system condition must be derived from monitoring system in milliseconds within which the breakers can successfully cut off the fault circuit [23]. Therefore, it is reasonable if we assume that the vehicle stays at the same location during the monitoring system making one evaluation. Besides, only one power unit of a vehicle is set as having faults for the sake of simplification.

There are 36 permutations of the location for two vehicles operating on up and down track. 30 types of system conditions are considered for each permutation (except that Fault 20, Fault 21, Fault 25 and Fault 26 can be removed when the vehicle running under the DC grid). 10 and 5 samples respectively for training and testing are collected under each condition.

The training and testing samples are both collected from the model established in PSCAD/EMTDC. Faults are injected after the system reaches steady-state (that is at the time of 4 s after starting simulation in our model). The start time of a fault is between 4.1 s to 4.12 s (one period of the system) for the differences when a fault occurs at different points of an AC period.

For training samples, the time range is evenly distributed to 10 points and each point is configured as a time of the fault injection for a sample, and the data collecting begins from 4 s and ends at 4.25 s. Therefore, there are 10 data with different start fault time and each one has 1250 time steps. Then 10 samples with 100 time steps each, respectively including 10 to 100 steps after a fault occurs, are re-sampled from each datum. While for testing samples, a point is randomly picked out from the time range as a start of fault, and the data collecting begins from 4.102 s to 4.122 s so as to ensure there are at least 10 steps after fault included in the 100 time-step sample.

In summary, we obtain 103200 samples for training and 5160 samples for testing. Each one has 100 time steps with sampling frequency of 5 kHz. The measured signals comprising voltages, currents, speeds and torques serve as features in the GRU-based network listed in Table III. The feature tabs are presented in the last column. The number of features for each sub-monitor can be found in Table II which contains signals and the location of vehicles.

### B. Hierarchical Structure

There are enormous features in the simulated HSR system. If the hierarchical structure is not exploited in the monitoring system, 67 signals plus 2 locations of vehicles will be the input and 30 conditions are going to be the output for the neural network to learn from. Comparing with the most complex one in our hierarchical structure (that is the sub-monitor in the vehicle having 13 features and 11 subclasses), the computational intensity of the network without hierarchical structure is obvious. The real-time application can be guaranteed since the requirement

TABLE III
FEATURES FOR EACH SAMPLE COLLECTED FROM THE HSR SYSTEM

| Sub-monitors | Signals | Features |
|---|---|---|
| AC TSS | Input current on phase A | $i_{psa1}$ |
| | Input current on phase B | $i_{psb1}$ |
| | Input current on phase C | $i_{psc1}$ |
| | Input three-phase RMS voltage | $u_{ps1}$ |
| | Output current to left sector | $i_{\alpha 0}$ |
| | Output voltage on catenary to left sector | $u_{c\alpha 0}$ |
| | Output voltage on feeder to left sector | $u_{f\alpha 0}$ |
| | Output current to right sector | $i_{\beta 0}$ |
| | Output voltage on catenary to right sector | $u_{c\beta 0}$ |
| | Output voltage on feeder to right sector | $u_{f\beta 0}$ |
| ATL1 | Input current | $i_{\alpha 1}$ |
| | Output current | $i_{\alpha 2}$ |
| | Voltage on catenary | $u_{c\alpha 1}$ |
| | Voltage on feeder | $u_{f\alpha 1}$ |
| ATL2 | Input current | $i_{\alpha 3}$ |
| | Voltage on catenary | $u_{c\alpha 2}$ |
| | Voltage on feeder | $u_{f\alpha 2}$ |
| ATR1 | Similar to ATL1 | / |
| ATR2 | Similar to ATL2 | / |
| DC TSS | Input current of transformer on phase A | $i_{psa2}$ |
| | Input current of transformer on phase B | $i_{psb2}$ |
| | Input current of transformer on phase C | $i_{psc2}$ |
| | Input three-phase RMS voltage of transformer | $u_{ps2}$ |
| | Output current of transformer on phase A | $i_{tsa1}$ |
| | Output current of transformer on phase B | $i_{tsb1}$ |
| | Output current of transformer on phase C | $i_{tsc1}$ |
| | Output three-phase RMS voltage of transformer | $u_{ts1}$ |
| | Output current of transformer on phase A | $i_{tsa2}$ |
| | Output current of transformer on phase B | $i_{tsb2}$ |
| | Output current of transformer on phase C | $i_{tsc2}$ |
| | Output three-phase RMS voltage of transformer | $u_{ts2}$ |
| | Output DC voltage of rectifier | $u_{dc}$ |
| | Output current of rectifier to left sector | $i_{L0}$ |
| | Output current of rectifier to right sector | $i_{R0}$ |
| PSL | Input current | $i_{L1}$ |
| | Voltage | $u_{dcL}$ |
| PSR | Similar to PSL | / |
| Up Vehicle | Input current | $i_{inU}$ |
| | Output current of transformer | $i_{acU}$ |
| | Output current of rectifier | $i_{dcU}$ |
| | Input voltage | $u_{inU}$ |
| | Output voltage of transformer | $u_{acU}$ |
| | Output voltage of rectifier | $u_{dcU}$ |
| | Current of motor on phase A | $i_{maU}$ |
| | Current of motor on phase B | $i_{mbU}$ |
| | Current of motor on phase C | $i_{mcU}$ |
| | Three-phase RMS voltage of motor | $u_{mU}$ |
| | Speed of motor | $\omega_U$ |
| | Torque of motor | $T_U$ |
| Down Vehicle | Similar to up vehicle | / |

TABLE IV
LATENCY OF ALGORITHM IMPLEMENTED ON CPU AND FPGA

| Monitors | Latency ($\mu$s) | | |
|---|---|---|---|
| | CPU | FPGA | Speed-up |
| **AC TSS** | 486.03 | 83.05 | 5.9× |
| **AT1** | 409.43 | 84.51 | 4.8× |
| **AT2** | 994.83 | 96.12 | 10.3× |
| **DC TSS** | 945.43 | 99.24 | 9.5× |
| **PS** | 391.33 | 82.98 | 4.7× |
| **Vehicle** | 1295.53 | 102.17 | 12.7× |
| **PCC** | 15.1 | 0.85 | 17.8× |

of a monitor for a smaller range is significantly reduced, either from the point of view of time or resources. The comparative experiments, as shown in Table V named CENT, with respect to the consumption of time and resources on FPGA as well as accuracy can further confirm it.

Moreover, the hierarchical structure also benefits to the sharp reduction of signal transmission along the line. No longer are the signals sampled from the sensors spread all over the track section demanded to be transferred to a centralized monitor. They are able to be analyzed and processed in their local region and the results (their evaluations of the condition in the neighborhood) and the location of vehicles are the only information needed to be passed through the upper-level monitor. On the other hand, the sub-monitor can also be appointed as the top-level monitor if necessary. For example, the monitor in AC TSS can serve as the top-level monitor for the whole AC section and the accuracy of classification for AC part can remain the same.

### C. Hardware Implementation of Monitoring System

The same C codes of the FDI method for the sub-monitors in the HSR system are implemented on both FPGA using Vivado HLS 2019.2 and CPU using Visual Studio 2017. The results, shown in Table IV, prove the necessity of hardware implementation in parallel on FPGA. It may take more than 1000 $\mu$s to reach a conclusion on CPU, while the time needed is within 100 $\mu$s on FPGA. There can be more than 10× acceleration when using FPGA, which means that the evaluation of the local condition of the sub-monitor can be achieved within a sampling time (200 $\mu$s in this paper). The evaluations made by various sub-monitors are able to be calculated in parallel at the same time in their local area. Therefore, the computation time of obtaining the final conclusion can be estimated as the sum of the longest runtime of sub-monitors (100 $\mu$s), the communication delay between levels (about 200 $\mu$s for 50 km distance with optical fiber communication) and the calculation of weighted votes on PCC (1 $\mu$s). It is less than 1 ms in the study case which qualifies it for real-time application.

### D. Noise Resistance

We test the anti-noise capability of the proposed method by adding different level of noises to the measured data. The signals-to-noise ratios (SNRs) of 90, 70, 50, 30 and 10 dB are

TABLE V
COMPARISON OF DIFFERENT NNS WITH RESPECT TO ACCURACY, LATENCY, AND RESOURCE CONSUMPTION ON THE XILINX VCU128 FPGA BOARD

| NN type | Accuracy | Latency (μs) | BRAM | DSP | FF | LUT |
|---|---|---|---|---|---|---|
| GRU | 97.56% | 102.17 | 212(5%) | 3094(34%) | 270562(10%) | 511366(39%) |
| LSTM | 96.32% | 93.68 | 420(10%) | 3846(42%) | 428412(16%) | 814345(62%) |
| RNN | 86.89% | 38.66 | 212(5%) | 1075(11%) | 184730(7%) | 282952(21%) |
| CNN1D | 94.78% | 106.62 | 147(3%) | 627(6%) | 228937(8%) | 949172(72%) |
| CENT | 85.08% | 144.14 | 212(5%) | 6641(73%) | 406789(15%) | 692674(53%) |

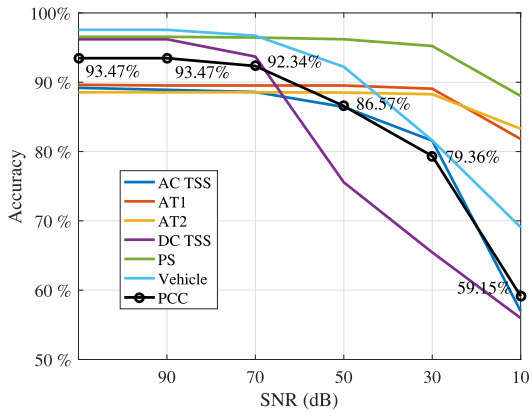


Fig. 7. FDI accuracy of monitors under different noise levels.

introduced. The accuracy of the monitors with different noise level is depicted in Fig. 7. It proves that the method can resist the noise to some extent because the hierarchical structure provides redundancy and complementation in the monitoring system. Only when the noise level reaches a very high level, for instance, when the SNR is less than 50 dB, the judgment of monitors may be unreliable. However, it would not be normal for sensors to contain such high-level noises.

### E. Comparison With Other NNs

To demonstrate the superiority of the proposed GRU-based FDI method, other prevailing neural networks such as standard LSTM, recurrent neural network (RNN) and 1D convolutional neural network (CNN1D) are experimented as comparison. Only the results of the sub-monitor in the vehicle that needs the most intense computation are displayed in Table V for simplification. The architectures of the neural networks are similar to ensure fairness; only the GRU layer is replaced by the other three layers and the principle for choosing parameters for the networks is still simplicity and effectiveness. Finally, RNN and LSTM layers are set to have 16 hidden units. For CNN1D layer, a kernel size of 2 is applied to each of the 8 filters, then a maxpool1d layer which slides a window of height 5 is directly connected. It indicates that none of the three NNs can reach the accuracy as GRU. RNN is not adequate due to the vanishing or exploding gradient problem that it suffers from in long-term sequential data. Standard LSTM and CNN1D can achieve better performance but consume tremendous resources than GRU. Above all, our GRU-based method outperforms others in terms of accuracy, the consumption of time and resources on FPGA.

### F. Off-Line Experiment

The off-line experiment is conducted based on the HSR model under hybrid AC/DC grid in PSCAD/EMTDC to validate the effectiveness of the proposed hierarchical GRU-based FDI method.

The accuracy is more than 93%. The evaluation of the condition for the entire system can be achieved within 1 ms and for the local area within one sampling time. Most of the instances are able to be predicted to their actual class. There are 337 instances in 5160 test samples that are wrongly classified; however, the error rate is reasonable. On one hand, Fault 15 is an open-circuit fault in transistors of rectifier that is a minor fault, especially when there is no load on the DC section. Therefore, it can be regarded as the same as normal condition. This kind of mistake can be corrected if more features of the rectifier are supervised. On the other hand, the reasons for major errors between Fault 4 and 10 or Fault 5 and 12 are the same. They are short-circuit faults happening at ATs and on catenary. The phenomenon can be similar if the fault on catenary is close to the AT. More information about the transmission line will also help to solve this problem.

### G. Real-Time HSR System Emulator Interfaced With FDI Monitor

Real-time emulation experiment is also conducted to validate the feasibility of the real-time FDI application. The same model of the HSR system operating under hybrid AC/DC grid is emulated on the emulator with Xilinx Virtex UltraScale+™ VCU118 FPGA board [24]. The emulator is validated by comparing the waveforms with the model simulated in PSCAD/EMTDC. Three fault scenarios are illustrated in Fig. 8. High resistance grounding at ATL2 (Fault 3), short-circuit at PSL (Fault 16) and open-switch in MMC of up vehicle (Fault 22) are shown in Fig. 8(a)(b), (c)(d) and (e)(f) respectively. All faults are injected at $t = 4.1$s. The XCVU9P FPGA on the board consists of 2,586,000 logic cells, 345.9MBb BRAM and 6,840 DSP slices. The real-time system-level model is emulated at 20 $\mu$s time-step and the resource consumption are 0.14% BRAM, 66.56% DSP, 35.08% FF and 90.55% LUT of the board. The QSFP connection is utilized to deliver the signals generated by the model to the FDI monitor, which is embedded in the dedicated VCU128 FPGA board, at a sampling frequency of 5 kHz. The lower-level sub-monitor makes an evaluation during one sampling time and outputs its result to the higher-level sub-monitor through the internal AXI protocol with desired physical on-field transmission delay inside the VCU128 board.

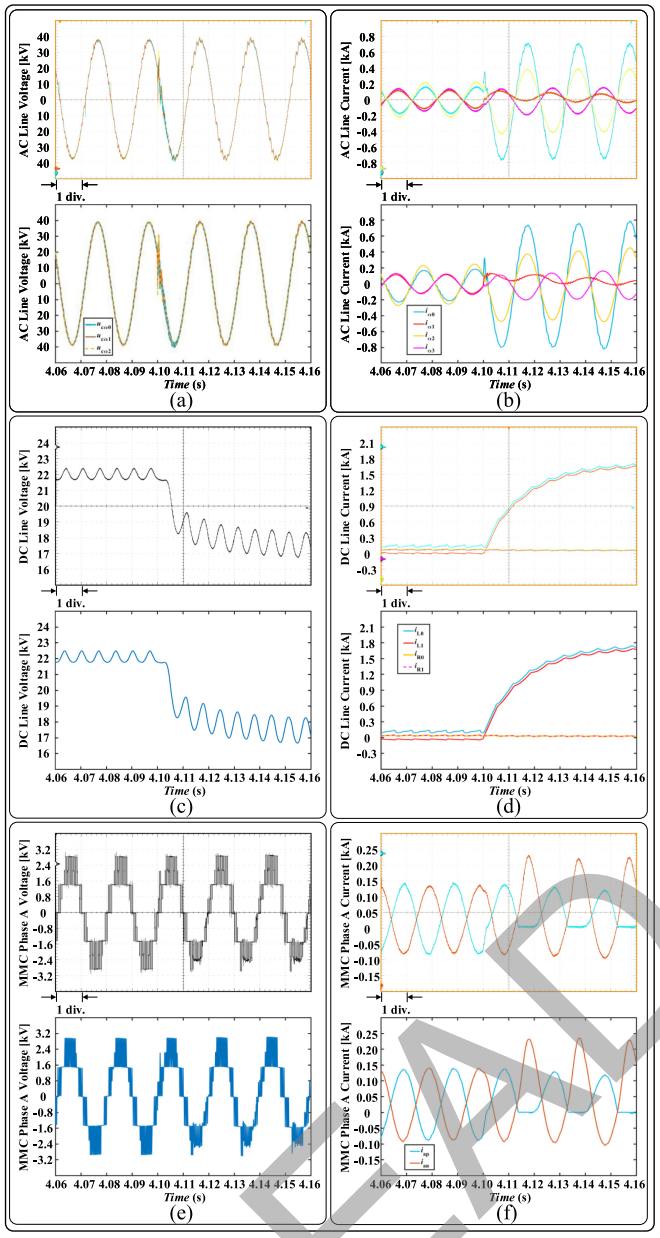


Fig. 8. Emulation results of the HSR system from real-time emulator on VCU118 FPGA (top oscilloscope sub-figure) and off-line simulation by PSCAD/EMTDC software (bottom sub-figure) for: (a) AC Line voltage, (b) AC Line current, (c) DC Line voltage, (d) DC Line current, (e) MMC phase a voltage, (f) MMC phase a current. x-axis: 10 ms/div.

The final decision made by the top-level monitor is shown on an oscilloscope by magnitude. The oscilloscope is also exploited to supervise the selected signals from the model. The hardware emulation platform is pictured in Fig. 9. The experimental results of four scenarios are demonstrated as examples in this section.

To properly take advantage of the devices we have, only the most important sub-monitors according to the specific fault type (i.e. one of the faults demonstrated below) are implemented in the VCU128 FPGA. For instance, the votes from the sub-monitors installed on AC TSS, ATR1, ATR2 and a vehicle running on Zone 4 are dominant to determine whether the

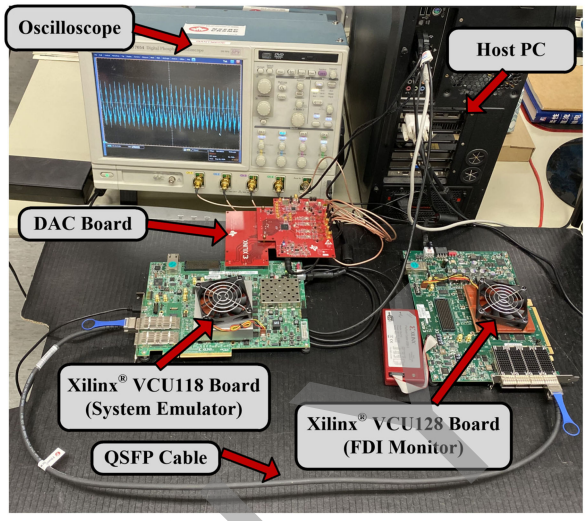


Fig. 9. Hardware emulation platform for real-time FDI experiment.

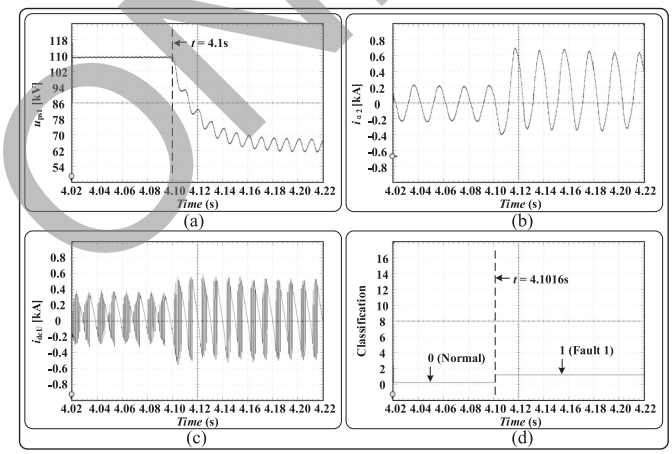


Fig. 10. Oscilloscope transient waveforms in scenario of *Fault 1* for: (a) AC TSS input three-phase RMS voltage, (b) ATL1 output current, (c) output current of rectifier of up vehicle, (d) output classification of the monitoring system.

short-circuit fault on ATR2 occurs. For this reason, the voting strategy changes a little along with different scenarios when the AC or DC TSS is appointed as the top-level monitor, but the GRU structures in sub-monitors keep the same. It also indicates the flexibility of our hierarchical method that the lower-level sub-monitor can be appointed as the top-level monitor when necessary. The labels for the fault type keep consistent with the off-line experiments in which PCC is the top-level monitor.

*1) Detectable Fault in the AC Section:* The phase A-B short-circuit fault in the primary side of the Scott transformer (Fault 1) on AC TSS can have a noticeable impact on either side sector of the AC section and may also cause some fluctuation in DC part. Fig. 10 presents the oscilloscope transient waveforms of AC TSS input three-phase RMS voltage ($u_{ps1}$), ATL1 output current ($i_{\alpha 2}$), output current of rectifier of up vehicle ($i_{dcU}$) as well as the output classification. The two vehicles are located in Zone 1 and Zone 2 respectively. The change is significant on signals from all the three-level structures. The voltage of the

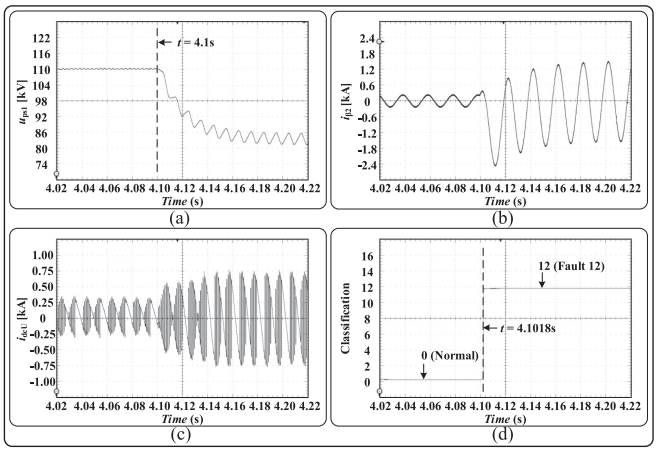


Fig. 11. Oscilloscope transient waveforms in scenario of *Fault 5* for: (a) AC TSS input three-phase RMS voltage, (b) ATR1 output current, (c) output current of rectifier of up vehicle, (d) output classification of the monitoring system.

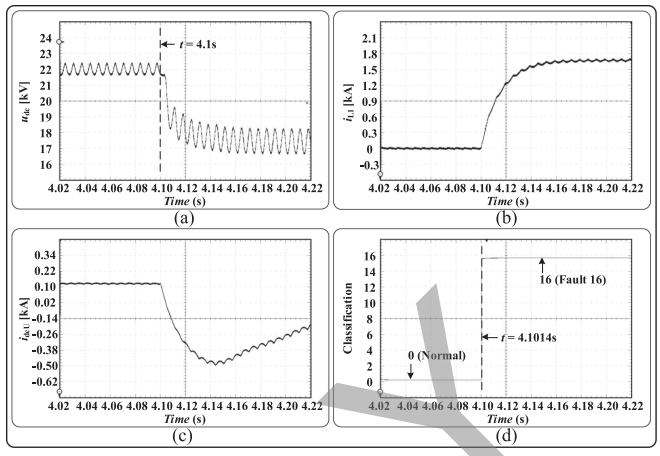


Fig. 13. Oscilloscope transient waveforms in scenario of *Fault 16* for: (a) DC TSS output voltage, (b) PSL input current, (c) output current of rectifier of up vehicle, (d) output classification of the monitoring system.

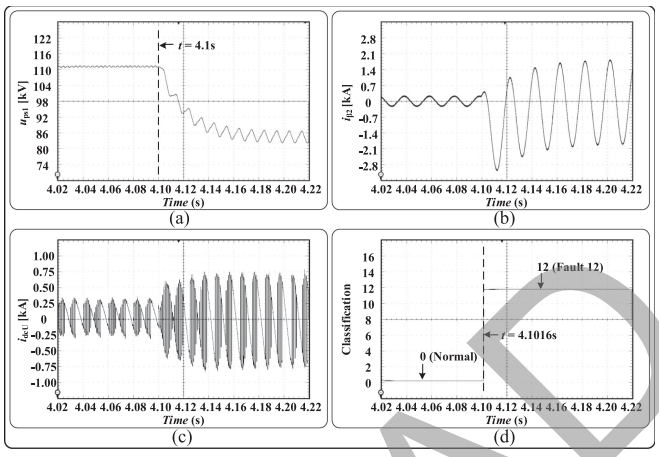


Fig. 12. Oscilloscope transient waveforms in scenario of *Fault 12* for: (a) AC TSS input three-phase RMS voltage, (b) ATR1 output current, (c) output current of rectifier of up vehicle, (d) output classification of the monitoring system.

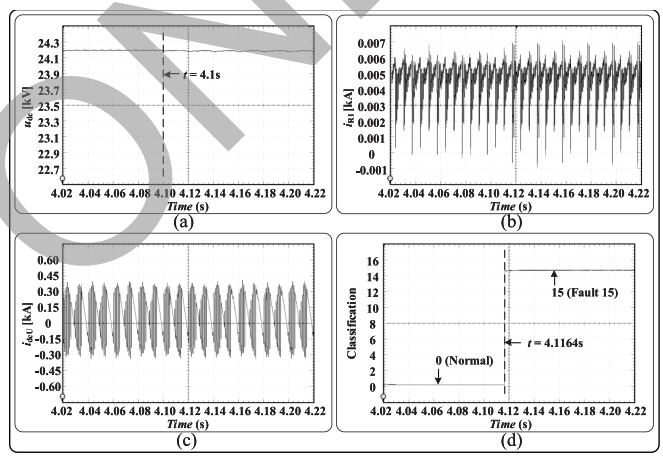


Fig. 14. Oscilloscope transient waveforms in scenario of *Fault 15* for: (a) DC TSS output voltage, (b) PSR input current, (c) output current of rectifier of up vehicle, (d) output classification of the monitoring system.

upper-level AC TSS drops suddenly and currents of lower-level ATL1 and vehicle increase after the fault occurs. The fault can be correctly detected as Fault 1 within 2 ms after fault.

*2) Undetectable Fault in the AC Section:* The short-circuit fault on ATR2 (Fault 5) and ground fault on catenary at Zone 4 (Fault 12) are going to cause huge variations of voltages and currents of the right side sector of the AC part. Fig. 11 and Fig. 12 presents the oscilloscope transient waveforms of AC TSS input three-phase RMS voltage ($u_{ps1}$), ATR1 output current ($i_{\beta 2}$), output current of rectifier of up vehicle ($i_{dcU}$) as well as the output classification. The two vehicles are located in Zone 4 and Zone 3. After the fault occurs, the voltages decrease while the currents increase in both scenarios. The curve trends of the signals resulting from these two faults are very similar, which explains why the monitoring system may make a mistake when classifying them. In this case, the Fault 5 is incorrectly classified as Fault 12, while it correctly detects the Fault 12 within 2 ms after fault.

*3) Detectable Fault in the DC Section:* The short-circuit fault on PSL (Fault 16) leads to a dramatic influence on the left side sector of the DC section. Fig. 13 presents the oscilloscope transient waveforms of DC TSS output voltage ($u_{dc}$), PSL input current ($i_{L1}$), output current of rectifier of up vehicle ($i_{dcU}$) as well as the output classification. The two vehicles are located in Zone 5 and Zone 6 respectively. The DC voltage $u_{dc}$ becomes unstable and the amplitude reduces suddenly. Meanwhile, $i_{L1}$ and $i_{dcU}$ apparently change in the opposite direction. The fault is correctly detected as Fault 16 within 2 ms after fault.

*4) Undetectable Fault in the DC Section:* The open-circuit in the transistor of rectifier (Fault 15) is a minor fault which may not result in obvious changes through signals we sampled here, especially when there's no load on the DC section. Fig. 14 and Fig. 15 presents the oscilloscope transient waveforms of DC TSS output voltage ($u_{dc}$), PSR input current ($i_{R1}$), output current of rectifier of up vehicle ($i_{dcU}$) as well as the output classification when Fault 15 occurs at $t = 4.1$s and when it's

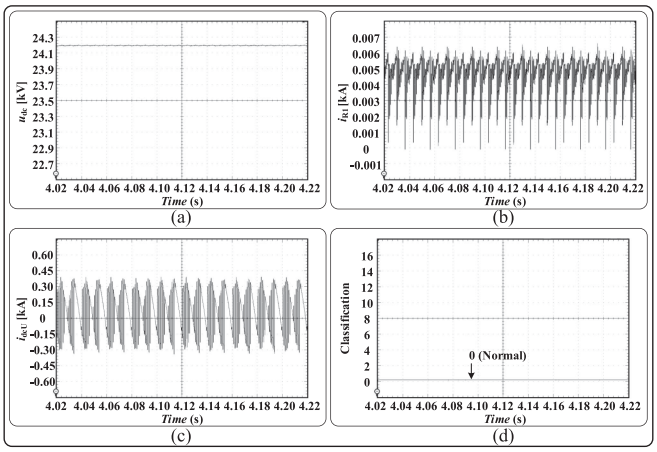


Fig. 15. Oscilloscope transient waveforms in scenario of *Normal Condition* for: (a) DC TSS output voltage, (b) PSR input current, (c) output current of rectifier of up vehicle, (d) output classification of the monitoring system.

normal condition. The two vehicles are located in Zone 2 and Zone 3 respectively. From the waveforms it can be seen that there are not much changes in these signals, especially without enough time steps collected after fault. Therefore, the monitor cannot classify it as Fault 15 immediately, but regards it as the same as normal condition.

## V. Conclusion

In this paper, we proposed a hierarchical GRU-based FDI method for the HSR system and implemented it on the FPGA for real-time monitoring. The data used for training and testing the NN were obtained from the HSR transient system model simulated in PSCAD/EMTDC and emulated on the FPGA in real time. The system was powered by both AC and DC substations, i.e. working under hybrid AC/DC grid. A hierarchical structure was adopted in the monitoring system. The lower-level monitor sampled and analyzed the signals in the neighborhood and immediately gave feedback to the upper-level monitor. The evaluations of the condition in a small range from sub-monitors were delivered level-by-level. Eventually, the top-level monitor assembled all the results from lower-level monitors and made the final decision of the condition of the entire HSR system using a voting strategy. The contrast experiments demonstrated that the evaluation of the real-time condition of the entire HSR system can be achieved by applying GRU which outperformed other NNs in terms of accuracy, runtime and hardware resource utilization on FPGA. The total runtime for an evaluation of the condition for the entire HSR system is within 1 ms and for a local area within one sampling time. Furthermore, the real-time experiment manifests the feasibility and capability of the proposed method in the real-time application, in which the correct classification of the system condition can be made within 2 ms after a fault occurs. The hierarchical structure was proved to be a better option for this kind of large-scale system. In addition, different level of noises have been added into measured signals, and the results revealed that our proposed method also had an excellent anti-noise performance.

Our experiments are based on the simulated HSR model in PSCAD/EMTDC and the FPGA-based real-time system emulator. For the real-world situation, more challenges, such as the imbalanced data problem between normal and faulty conditions, remain to be solved for the data-dependent method. Future work will consider how to generate reliable samples in accordance to the existing field data using advanced technologies.

## References

[1] H. Hayashiya *et al.*, "Reliability analysis of DC traction power supply system for electric railway," in *Proc. EPE'17 ECCE Eur.*, Sep. 2017, pp. P.1–P.6.

[2] X. Zhu, H. Hu, H. Tao, and Z. He, "Stability analysis of PV plant-tied MVDC railway electrification system," *IEEE Trans. Transp. Electrific.*, vol. 5, no. 1, pp. 311–323, Mar. 2019.

[3] A. Verdicchio, P. Ladoux, H. Caron, and C. Courtois, "New medium-voltage DC railway electrification system," *IEEE Trans. Transp. Electrific.*, vol. 4, no. 2, pp. 591–604, Jun. 2018.

[4] ABB, "Traction systems for locomotives and high-speed applications," 2018. [Online]. Available: https://library.e.abb.com/public/896cf517fcce4406b7a4facb6d6b7d0c/Traction_systems_high%20power_RevB_180916_web.pdf

[5] Bombardier, "MITRAC hybrid: The dual power propulsion chain," 2009. [Online]. Available: https://www.bombardier.com/content/dam/Websites/bombardiercom/supporting-documents/BT/Bombardier-Transportation-ECO4-MITRAC_Hybrid-EN.pdf

[6] G. Cho *et al.*, "A novel fault-location algorithm for AC parallel auto-transformer feeding system," *IEEE Trans. Power Del.*, vol. 34, no. 2, pp. 475–485, Apr. 2019.

[7] D. Xie and X. Ge, "A state estimator-based approach for open-circuit fault diagnosis in single-phase cascaded H-bridge rectifiers," *IEEE Trans. Ind. Appl.*, vol. 55, no. 2, pp. 1608–1618, Mar./Apr. 2019.

[8] J. Jiang, L. Zhou, S. Gao, W. Li, and D. Wang, "Frequency response features of axial displacement winding faults in autotransformers with split windings," *IEEE Trans. Power Del.*, vol. 33, no. 4, pp. 1699–1706, Aug. 2018.

[9] K. Zhang, B. Jiang, X. Yan, and J. Shen, "Interval sliding mode observer based incipient sensor fault detection with application to a traction device in China railway high-speed," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 2585–2597, Mar. 2019.

[10] Y. Wu, B. Jiang, and N. Lu, "A descriptor system approach for estimation of incipient faults with application to high-speed railway traction devices," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 49, no. 10, pp. 2108–2118, Oct 2019.

[11] K. R. Li, K. Y. See, F. Fan, and N. B. Narampanawe, "Condition monitoring of high power DC feeder with multiple parallel cables," *IEEE Trans. Power Del.*, vol. 35, no. 2, pp. 787–793, Apr. 2020.

[12] F. Garramiola, J. Poza, J. D. Olmo, and T. Nieva, "Hardware-in-the-loop performance analysis of a railway traction system under sensor faults," *J. Eng.*, vol. 2019, pp. 3797–3801, Jun. 2019.

[13] A. SzelÄg and M. Patoka, "Some aspects of impact analysis of a planned new 25kV AC railway lines system on the existing 3kV DC railway system in a traction supply transition zone," in *Proc. Int. SPEEDAM*, pp. 796–800, Jun. 2014.

[14] Z. Liu and K. Hu, "A model-based diagnosis system for a traction power supply system," *IEEE Trans. Ind. Informat.*, vol. 13, no. 6, pp. 2834–2843, Dec 2017.

[15] C. Dai, Z. Liu, K. Hu, and K. Huang, "Fault diagnosis approach of traction transformers in high-speed railway combining kernel principal component analysis with random forest," *IET Elect. Syst. Transp.*, vol. 6, pp. 202–206, Sep. 2016.

[16] D. Peng, Z. Liu, H. Wang, Y. Qin, and L. Jia, "A novel deeper one-dimensional CNN with residual learning for fault diagnosis of wheelset bearings in high-speed trains," *IEEE Access*, vol. 7, pp. 10278–10293, 2019.

[17] T. D. Bruin, K. Verbert, and R. Babuška, "Railway track circuit fault diagnosis using recurrent neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 3, pp. 523–533, Mar. 2017.

[18] A. Gómez-Expósito, J. M. Mauricio, and J. M. Maza-Ortega, "VSC-based MVDC railway electrification system," *IEEE Trans. Power Del.*, vol. 29, no. 1, pp. 422–431, Feb 2014.

[19] K. Cho *et al.*, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," 2014. [Online]. Available: https://arxiv.org/abs/1406.1078,
[20] M. Abadi *et al.*, "Tensorflow: A system for large-scale machine learning," 2018. [Online]. Available: http://arxiv.org/abs/1605.08695
[21] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014. [Online]. Available: https://arxiv.org/abs/1412.6980
[22] Xilinx, "VCU128 evaluation board user guide," 2018. [Online]. Available: https://www.xilinx.com/support/documentation/boards_and_kits/vcu128/ug1302-vcu128-eval-bd.pdf
[23] A. Rojek, "Parameters of DC high-speed circuit-breakers," in *Proc. 13th Int. Conf. MET*, Jul. 2018, pp. 1–5.
[24] Xilinx, "VCU118 evaluation board user guide," 2018. [Online]. Available: https://www.xilinx.com/support/documentation/boards_and_kits/vcu118/ug1224-vcu118-eval-bd.pdf

**Qin Liu** (Student Member, IEEE) received the B.Eng. degree in measurement and control technology and instrument from Beijing Jiaotong University, Beijing, China, in 2013. She is currently working toward the Ph.D. degree with the School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University and being a Visiting Ph.D. Student in electrical and computer engineering with the University of Alberta, Edmonton, AB, Canada. Her research interests include real-time simulation and control of power systems, hardware acceleration, and system-on-chip.

**Tian Liang** (Student Member, IEEE) received the B.Eng. degree in electrical engineering from Nanjing Normal University, Nanjing, Jiangsu, China, in 2011. He received the M.Eng. degree in biomedical engineering from Tsinghua University, Beijing, China, in 2014. He is currently working toward the Ph.D. degree in electrical and computer engineering with the University of Alberta, Edmonton, AB, Canada. His research interests include real-time emulation of power systems, power electronics, and system-on-chip.

**Venkata Dinavahi** (Fellow, IEEE) received the B.Eng. degree in electrical engineering from the Visveswaraya National Institute of Technology, Nagpur, India, in 1993, the M.Tech. degree in electrical engineering from the Indian Institute of Technology Kanpur, Kanpur, India, in 1996, and the Ph.D. degree in electrical and computer engineering from the University of Toronto, Toronto, ON, Canada, in 2000. He is currently a Professor with the Department of Electrical and Computer engineering, University of Alberta, Edmonton, AB, Canada. His research interests include real-time simulation of power systems and power electronic systems, electromagnetic transients, device-level modeling, large-scale systems, and parallel and distributed computing.