Identifying Overlapping Successive Events Using a Shallow Convolutional Neural Network

Wenting Li, Meng Wang

Rensselaer Polytech Institute

Dec. 4 2018

Motivations of Event Identification

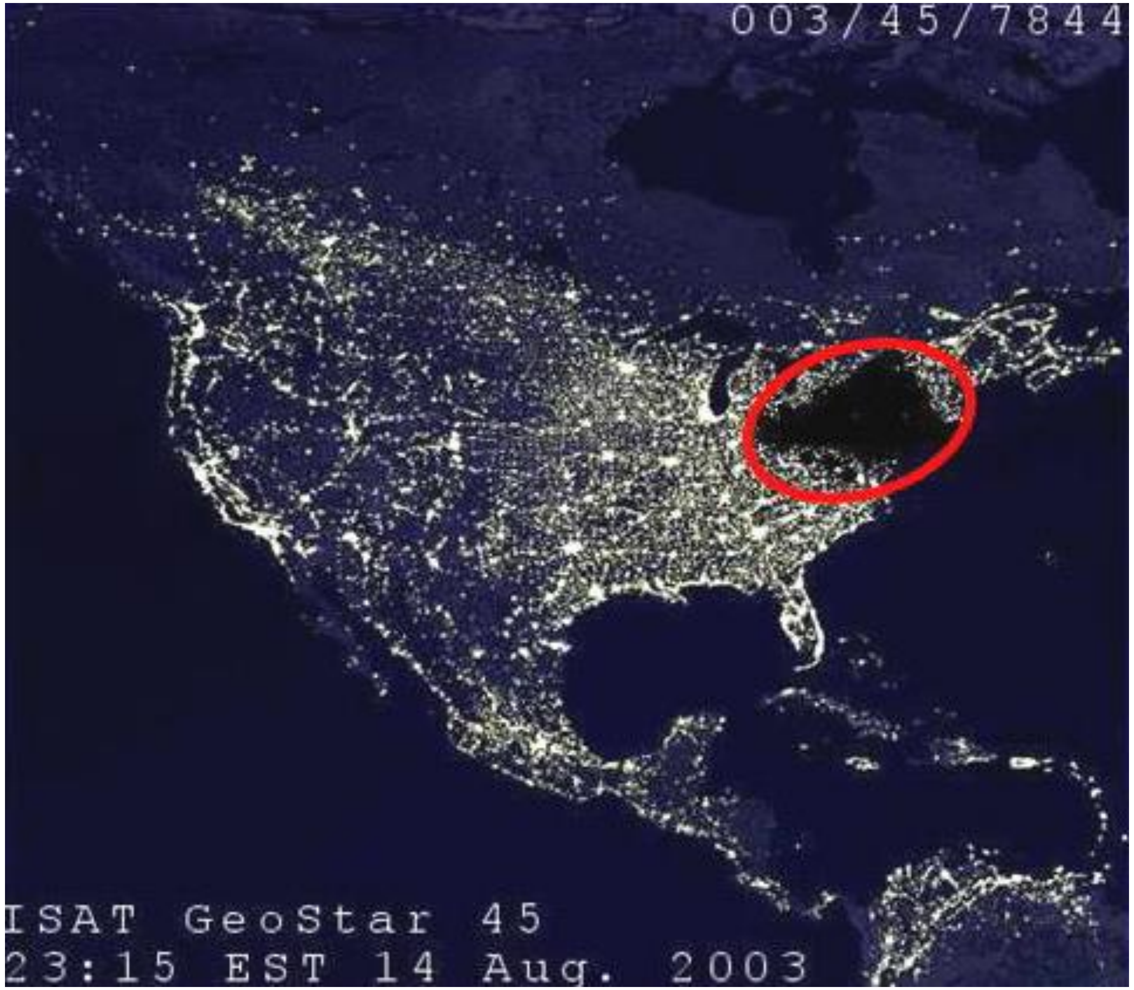


Fig. 1 The satellite photograph shows the blackout in 2003
https://www.snopes.com/fact-check/blackout-2/

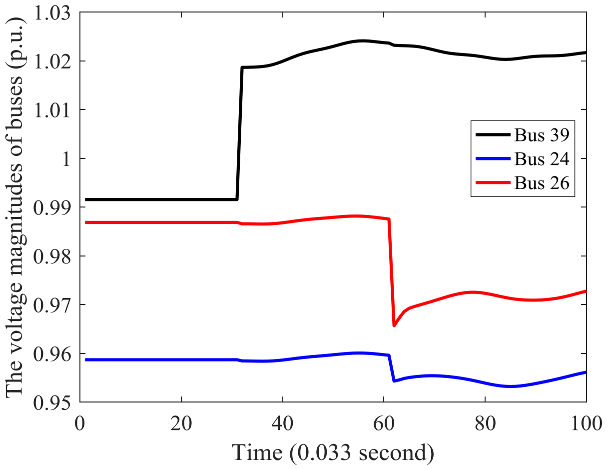
- Efficiently identifying large events are beneficial to **avoid blackout** and improve power system **stability and reliability**;
- More than **2000** Phasor measurement units (PMUs) have been installed in the United States;
- A number of PMU data motivate people to monitor and control power system with **data-driven** methods.

CURENT

Existing Approaches

- Recent development of event identification methods[1-5]:
- Advantages:
  - model-free, robust to model errors.
- Limitations:
  - **Single events** or multiple events with long time intervals and minor overlapping;
  - A **large number** of training datasets;
  - No clear **physical interpretations;**
  - High training **complexity.**

Challenges of Identifying Successive Events



(a)

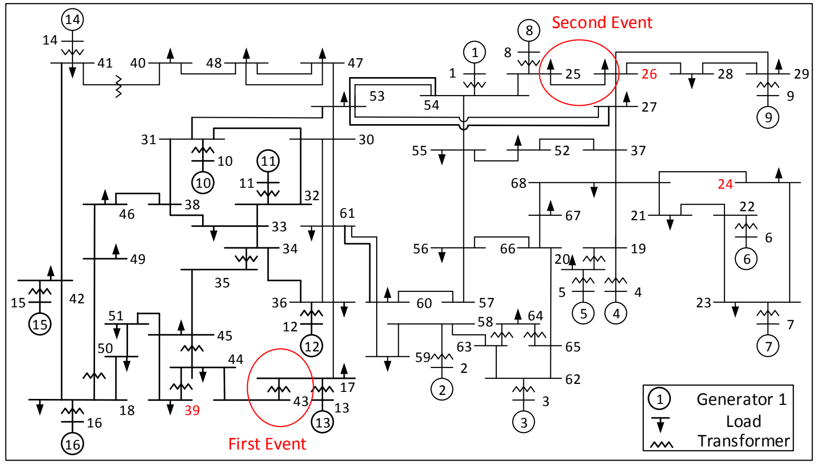
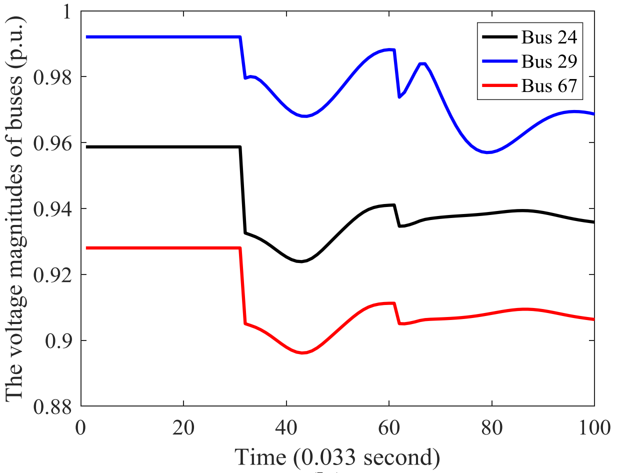


Fig. 3 (a) Line 17-43 tripped at 1 s followed by line 25-26 tripped event at 2 s in the IEEE 68-bus power system.

- The two successive events in (a) are almost independent,

Challenges of Identifying Successive Events



(b)

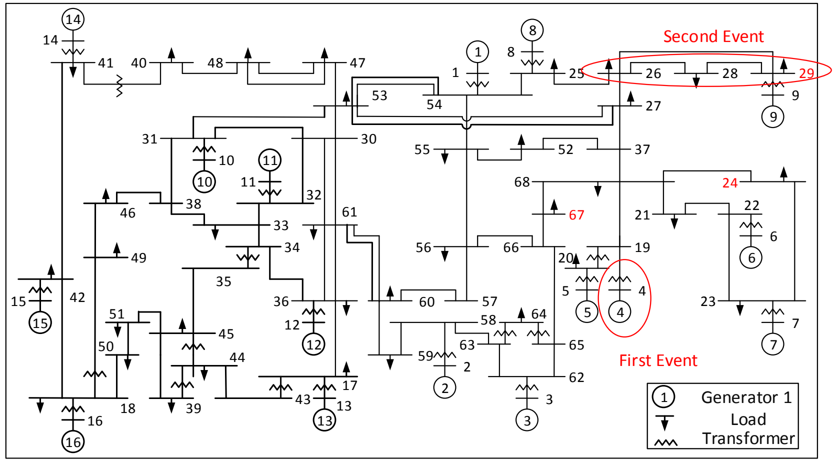


Fig. 4 (b) Generator at bus 4 tripped at 1 s followed by line 26-29 tripped at 2 s in the IEEE 68-bus power system.

- The two successive events in (a) are almost independent, but the successive events in (b) **are overlapping**;
- **Insufficient** number of historical successive events for training;
- **Goals**
    - Offline training of **single events**
    - Online identification of **overlapping** successive events

Definition of Dominant Features

- Dominant Features:
  - $M_1^T = [y_1, \ldots, y_T]$ is a data matrix of m PMU channels in T,
    $$M_1^T \approx U_r \Sigma_r V_r^\dagger$$
    - $U_r, V_r \in C^{m \times r}$ $\Sigma_r = diag(\sigma_1, \cdots, \sigma_r)$ are singular vectors and values.
  - Since $y_t = A\, y_{t-1}$, the two data matrices in sequential windows:
    $$M_1^T = A M_2^{T+1}$$
  - Define the eigenvalues of A and $\sigma_1, \cdots, \sigma_r$ as the $\boldsymbol{dominant\ features}$.
- Classification:
  - The extracted dominant features are input to a 2-layer CNN classifier.

Offline Training

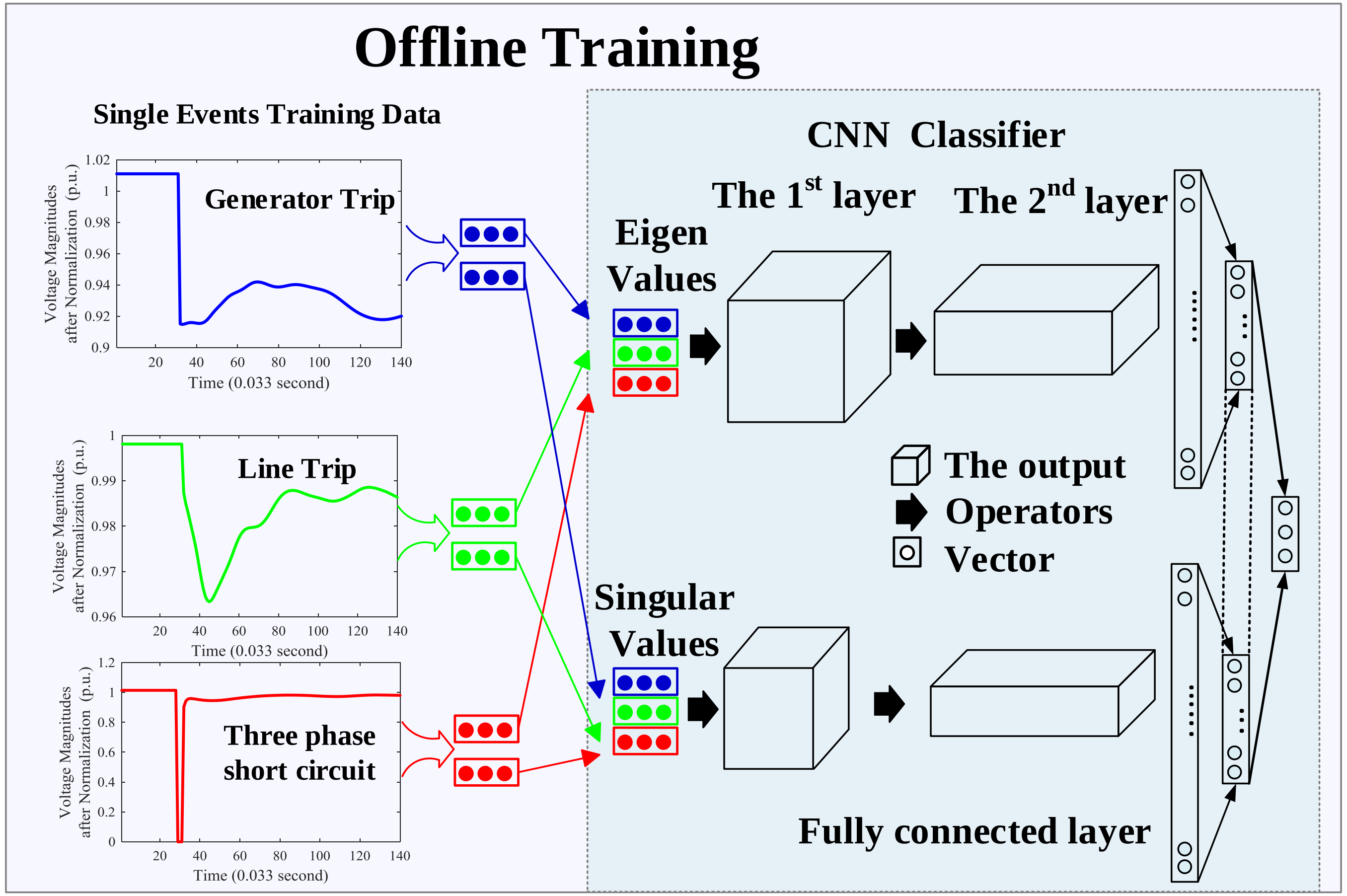


Fig. 5 Offline training on single events of Generator Trips (**GT**s), Line Trips (**LT**s), and Three phase short circuit (TPs)

- Extract dominant **Eigen values** $\boldsymbol{\lambda}$ of the state matrix and **singular values** $\boldsymbol{\sigma}$ of the recorded single events as features;
- Input featrures (λ, σ) to a 2-layer CNN of **two-path**;
- Train the **2-layer CNN classifier** to identify the type of an event.

Online Testing

Given the two **successive events** occurring at T1 and T2 respectively, there are **three steps** to identify the type of second event:

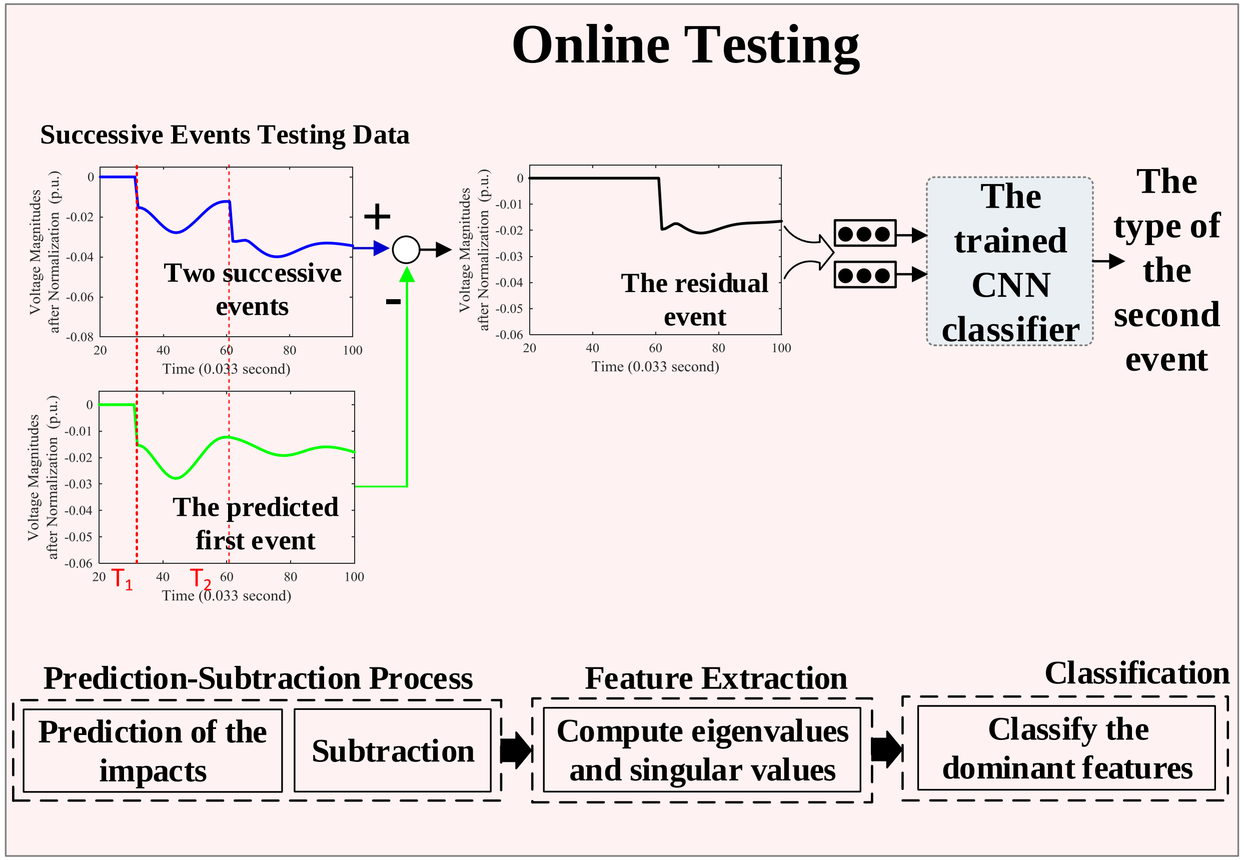


Fig. 5 Online testing of successive events through three components

- Step 1: **predict** the impact of the first event after T_2 and **subtract** it;

Online Testing

Given the two **successive events** occurring at T1 and T2 respectively, there are three steps to identify the type of second event:

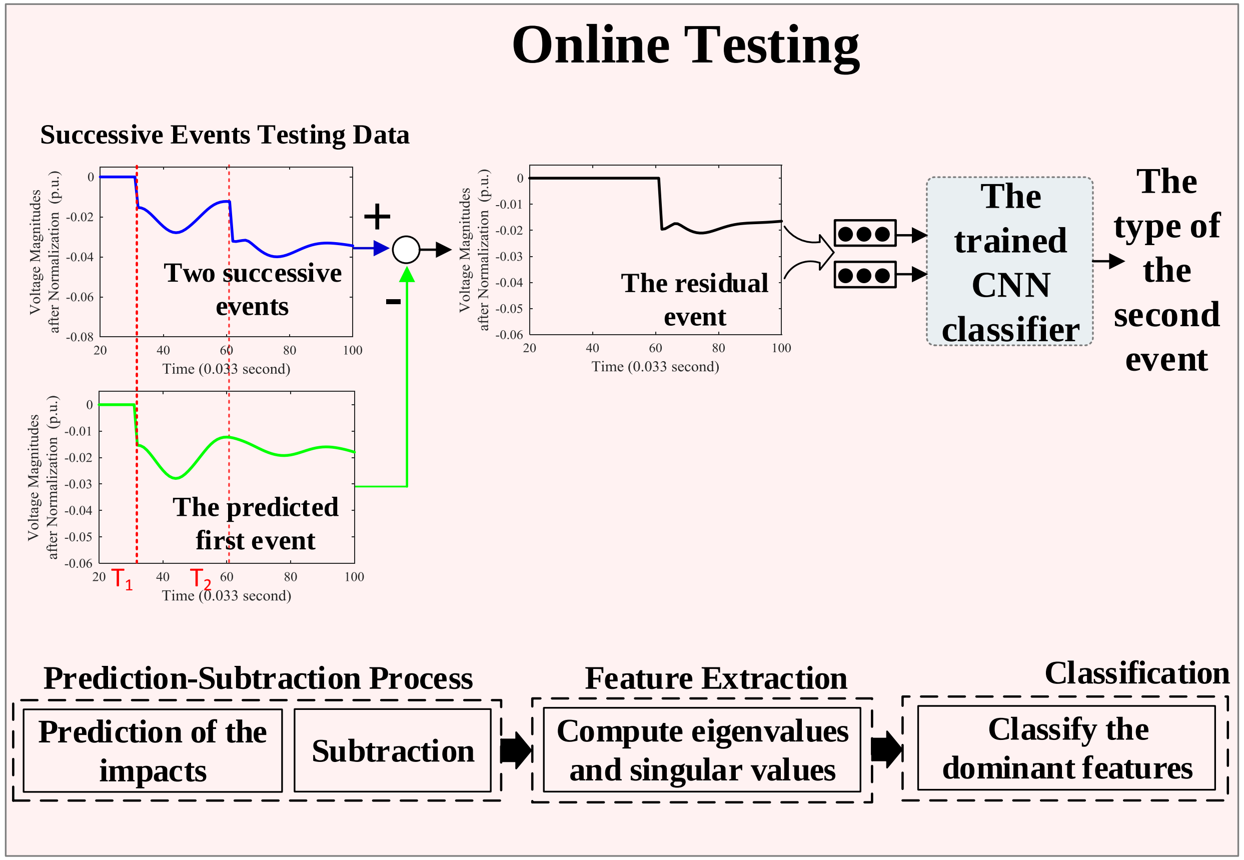


Fig. 5 Online testing of successive events through three components

- <u>Step 1</u>: **predict** the impact of the first event after T_2 and **subtract** it;
- Various methods can be employed to predict the measurements, such as time series analysis and Hankel matrix based methods [6-7];

Online Testing

Given the two **successive events** occurring at T1 and T2 respectively, there are three steps to identify the type of second event:

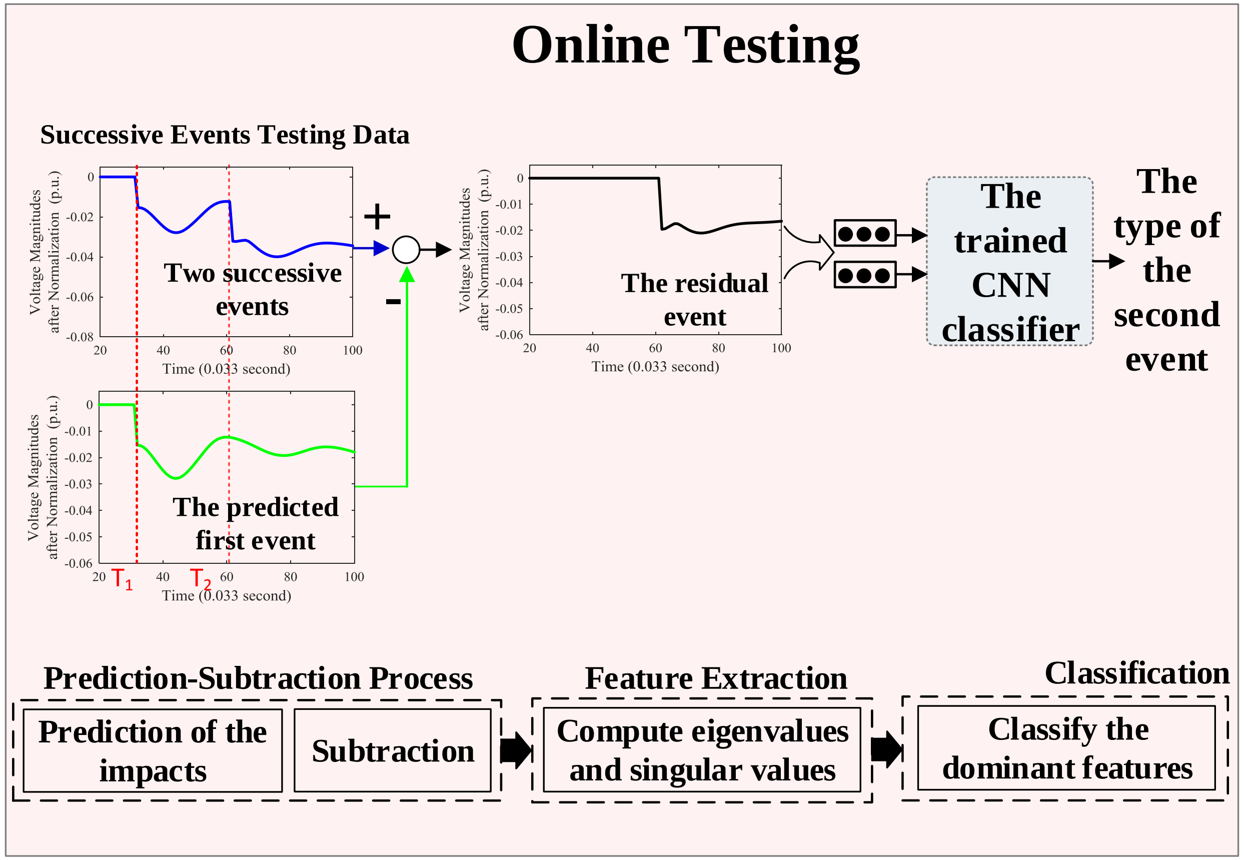


Fig. 5 Online testing of successive events through three components

- Step 1: **predict** the impact of the first event after T_2 and **subtract** it;
- Various methods can be employed to predict the measurements, such as time series analysis and Hankel matrix based methods [6-7];
- We apply the prediction method based on the **dominant eigenvalues.**

Online Testing

Given the two successive events occurring at T1 and T2 respectively, there are **three steps** to identify the type of second event:

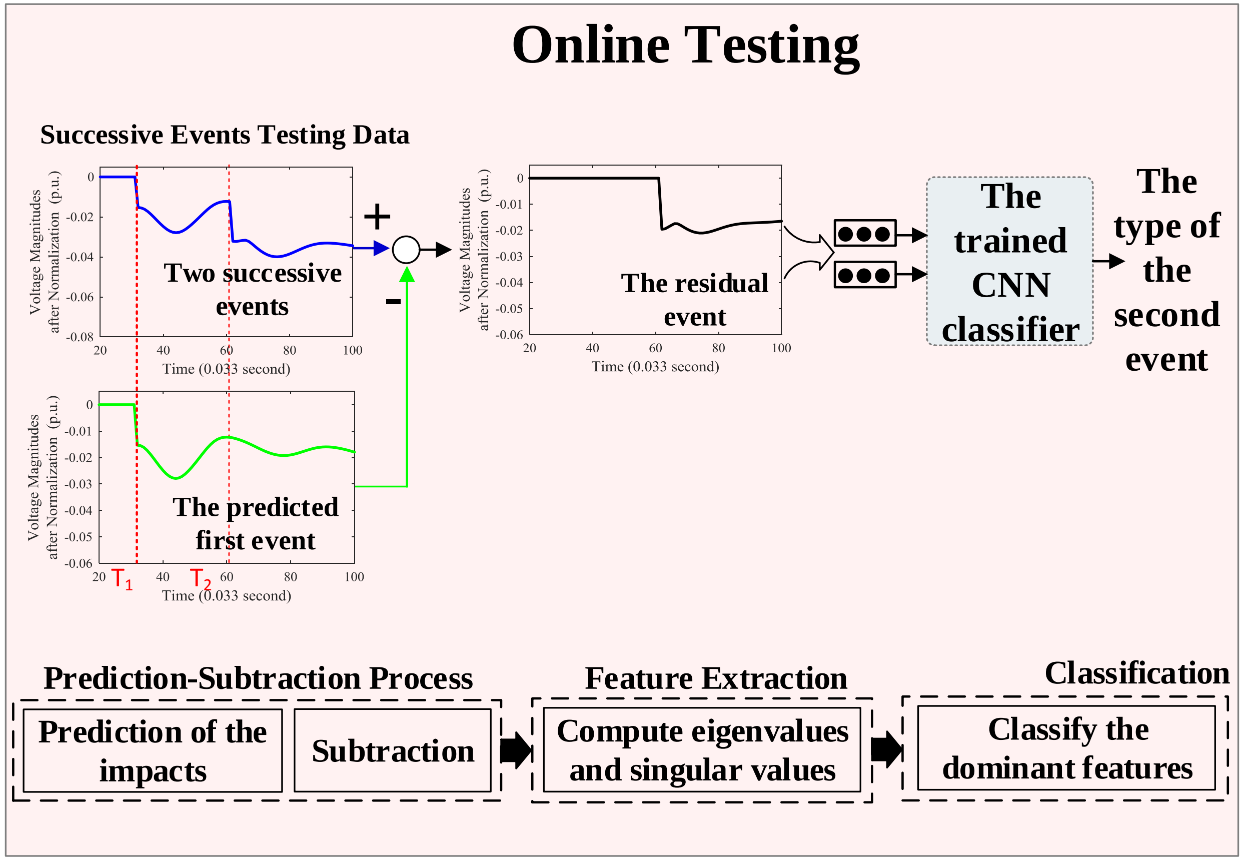


Fig. 5 Online testing of successive events through three components

- <u>Step 2</u>: The dominant features **(λ, σ) are extracted** from the residual measurements;

Online Testing

Given the two successive events occurring at T1 and T2 respectively, there are **three steps** to identify the type of second event:

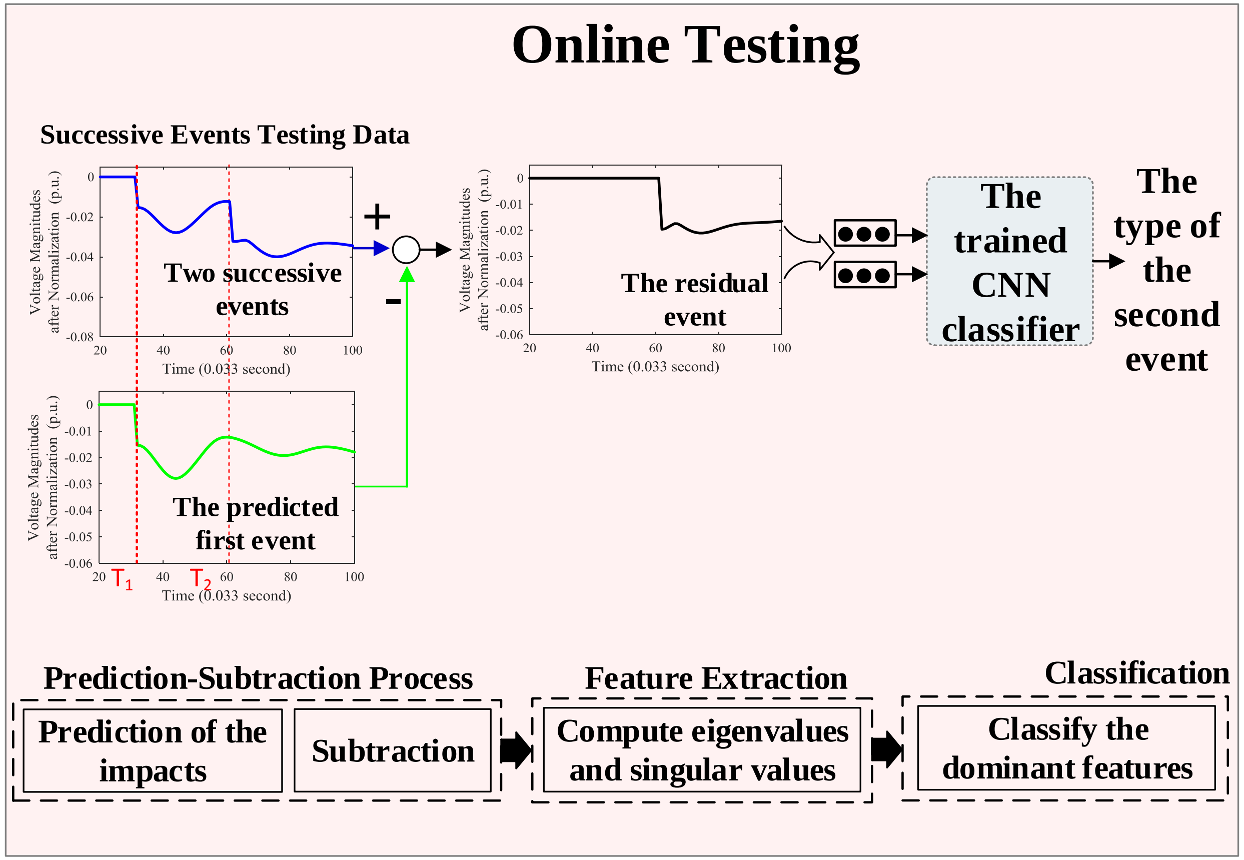


Fig. 5 Online testing of successive events through three components

- <u>Step 2</u>: The dominant features **(λ, σ) are extracted** from the residual measurements;
- <u>Step 3</u>**:** The trained **CNN outputs the type** of the second event in real time.

Numerical Results

Table 1.The Identification Accuracy Rate (**IAR**[1]) of 380 simulated datasets

IAR %	LT + GT	LT + TP	LT + LT	GT + GT	GT + TP	TP +GT
CNN-F[2]	**86.4**	**97.8**	**81.1**	**87.3**	**77.5**	**91.3**
CNN-T[3]	63.8	65.3	80.5	63.1	26.5	85.8

- Simulated datasets are generated by **PSS/E** in the IEEE 68-bus power system.
- 967 single events are employed to train the classifier, and the CNN identifies 1136 successive events of different combinations as shown in Table , where "**LT+GT**" denotes a line trip event is followed by a generator trip event.

[1]**IAR** is the ratio of the number of correctly classified events to the total number of events.

[2]**CNN-F** represents the CNN with **dominant features** as inputs.

[3]**CNN-T** represents the CNN with **time series** as inputs.

Performance with small training datasets

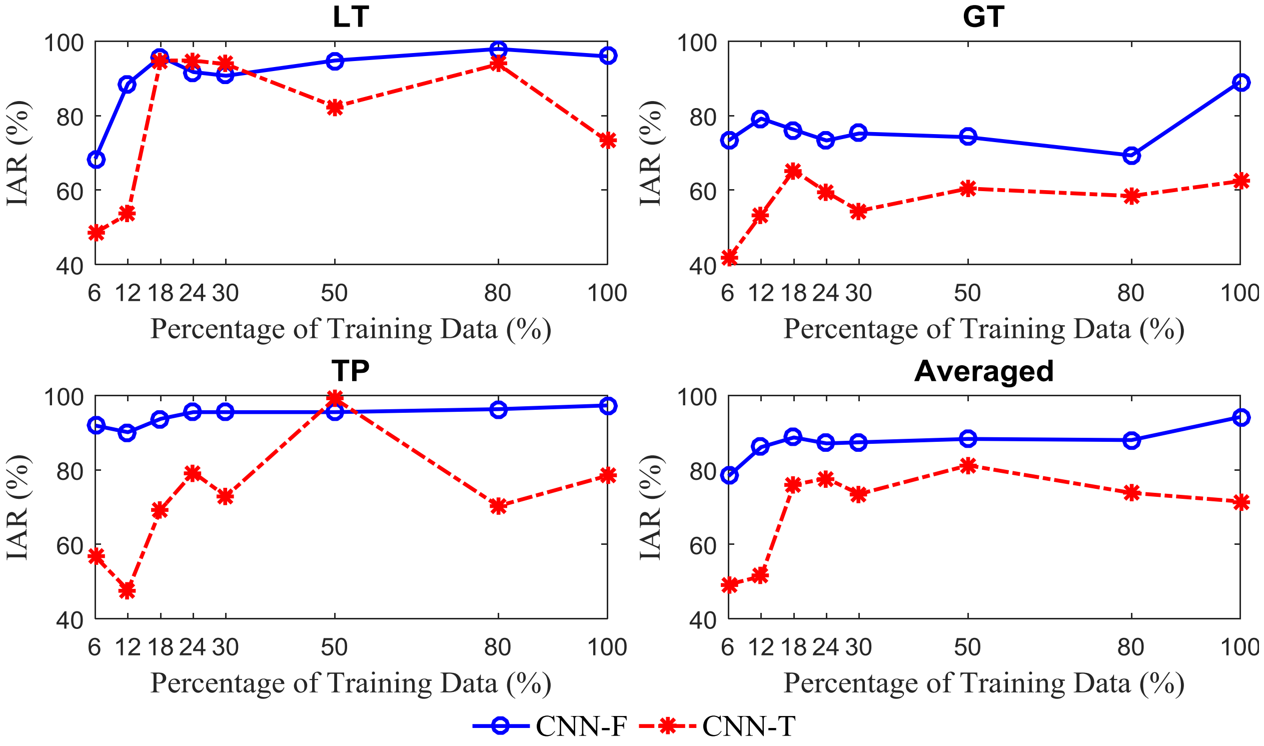


Fig. 6 The IAR of CNN-F and CNN-T when partial training datasets are available

- The dominant features are robust to initial conditions, and the **CNN-F** achieves **a higher IAR** with small training datasets.
- The IAR of **CNN-T is sensitive** to the size of the training data.

The Impact of Subtracting the First Event

Table 1. The performance comparison with and without subtracting the first event ($\Delta T = 1s$)

Classifier	Process	LT %	GT %	TP %	Overall %
CNN-F	NS	79.4	68.3	96.4	81.9
CNN-F	**SP**	**95.9**	**89.1**	**97.3**	**94.2**
CNN-T	NS	94.8	83.2	78.4	85.1
CNN-T	SP	73.2	62.4	78.4	71.5

- “Not Subtract (**NS**)” :using the data after the second event directly.
- “Subtract the Prediction (**SP**)” means using the residual data after subtracting the first event.
- Subtracting the impact of the first event can **enhance** CNN-F’s performance significantly.
- CNN-F is **robust to inaccuracies** in the measurements.

Conclusions

- **Compact features**, like the dominant eigenvalues of the state matrix, are effective to characterize events and robust to initial conditions.
- The proposed **CNN classifier** trained on single events is able to identify successive events in real time.
- The CNN-F using the dominant features is robust to **the measurement inaccuracies and small training datasets**.
- The proposed **prediction-subtraction** process can reduce the impacts of earlier events and enhance the identification accuracy.

CURENT

Reference

[1] W. Li and M. Wang and J. Chow “Real-time Event Identication through Low-dimensional Subspace Characterization of High-dimensional Synchrophasor Data,” IEEE Trans. Power Syst., 2018, 4937-4947.

[2] Y. Song, W. Wang, Z. Zhang, H. Qi, and Y. Liu, “Multiple event detection and recognition for large-scale power systems through cluster-based sparse coding,” IEEE Trans. Power Syst., 2017.

[3] Valtierra-Rodriguez, Martin, et al. “Detection and classification of single and combined power quality disturbances using neural networks,” IEEE Transactions on Industrial Electronics, 61.5 (2014): 2473-2482.

[4] Jiang, Huaiguang, et al. “Fault detection, identification, and location in smart grid based on data-driven computational methods,” *IEEE Transactions on Smart Grid* 5.6 (2014): 2947-2956.

[5] Rafferty, Mark, et al. "Real-time multiple event detection and classification using moving window PCA." IEEE Transactions on Smart Grid 7.5 (2016): 2537-2548.

[6] Hao, Yingshuai, et al. “Modelless Data Quality Improvement of Streaming Synchrophasor Measurements by Exploiting the Low-Rank Hankel Structure.” IEEE Transactions on Power Systems 33.6 (2018): 6966-6977.

[7] Elsner, James B. "Analysis of time series structure: SSA and related techniques." (2002): 1207-1208.

Acknowledgements

This work was supported primarily by the ERC Program of the National Science Foundation and DOE under NSF Award Number EEC-1041877 and the CURENT Industry Partnership Program, and NSF Grant 1508875

Other US government and industrial sponsors of CURENT research are also gratefully acknowledged.

CURENT