Email: editor@ijerst.com or editor.ijerst@gmail.com

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 21, Issue 2, 2025

Advanced islanding Detection for Power Grids using Transformer Neural Networks

[1]S. Imran Khan, [2]Mandala Bhavya Sree, [3]Chakali Ajith, [4]Kandlapalli Sowjanya, [5]Kundhu Shireesha, [6]Shaik Naseem Akhtar,

[1]Assistant Professor, Department Of EEE, Ananthalakshmi Institute Of Technology And Sciences, Itikalapalli, Near Sk University,Ananthapur.

[2,3,4,5,6] Student, Department Of EEE, Ananthalakshmi Institute Of Technology And Sciences, Itikalapalli, Near Sk University,Ananthapur.

Abstract

The increasing use of solar PV systems is causing distribution companies to worry about accidental islanding more and more. Due to unintended energization and low power quality, these accidental islands endanger workers and sensitive equipment. There are non-detection zones in conventional detection technologies. Due to the need for human feature extraction from data, new machine learning-based approaches to accidental island detection are difficult to predict. Our proposed end-to-end solution automates the feature extraction portion of the modeling process using Time Series Transformers, which simplifies the process overall. Time Series Transformer provides a simple but trustworthy detection model for islanding detection, as shown by our findings, which surpass those of existing machine learning approaches. Jhoanna Rhodette Pedrasa is a professor at the University of the Philippines' Diliman campus in Quezon City, Philippines, who specializes in electrical and electronics engineering. jipedrasa@up.edu.ph a straightforward and precise model for islanding detection based on the features extracted from the data. The structure of this article is as follows. Part II provides a summary of the most up-to-date approaches to islanding identification using machine learning and deep learning. In Section III, the dataset, TST architecture, additional models, and assessment measures are detailed. Discussion and outcomes of the experiments are presented in Section IV. The article is concluded in section V.

Keywords—deep neural networks, feature extraction, islanding, machine learning, protection

INTRODUCTION

Unintentional islanding is a major issue for electric utilities, especially in regions where PV is widely used, because of the growing number of residential solar photovoltaics (PV) systems [1]. When a section of the power grid gets disconnected from the main grid—typically as a result of maintenance or faults—but is still powered by local DER like solar PVs, this phenomenon is called accidental islanding. The absence of grid support services that regulate power quality makes unintended islands potentially dangerous places for sensitive equipment to be located, hence this is obviously not desired. Utility personnel doing maintenance might be put in danger by unintended islands because they could think the wires are de-energized. Inverters have both active and passive detection techniques that can identify islanded circumstances, which helps to avoid accidental islands. Bypass inverters work by keeping an eye on a certain parameter, such voltage or frequency, and turning off the device when it goes beyond a certain threshold. There are non-detection zones using passive approaches, but [2]. In contrast, active approaches intentionally introduce small disturbances to enhance detection accuracy, but this comes at the expense of power quality degradation [2]. Machine learning approaches are now being used in research to solve these problems since they are very accurate and do not compromise power quality [4]-[10]. Because they rely on human feature extraction, however, existing machine learning approaches are challenging to model [3].

Our study's key proposal is to automate Machine learning and deep learning are the two primary approaches to data-driven islanding identification that are covered in this section. In machine learning, there are two main steps: first, you extract features using a Fourier or Wavelet transform. Then, you design a classifier utilizing those features. Deep learning approaches, on the other hand, include feature extraction into the model itself, therefore it is not necessary to separate the two. The Fourier Transform [A] In order to extract frequency-domain harmonic characteristics, most mainstream machine learning

ISSN 2319-5991 www.ijerst.com

Vol. 21, Issue 2, 2025

approaches to islanding detection use the Fourier transform. For the purpose of modeling the 2nd order harmonic characteristics collected from the signal, prior research has used K-Nearest Neighbors [4] and Long Short-Term Memory (LSTM) [5]. The 3rd order harmonics and overall harmonic distortion levels in the voltage and current signals were also included in the models used by Convolutional Neural Network (CNN) [6] and Long Short-Term Memory (LSTM) [7]. Two major problems, nevertheless, plague this method. To start with, non-stationary disturbances may not be well captured by the Fourier transform because of its frequency domain limitation. Secondly, non-linear loads and transformer energization may produce comparable harmonics, which might lead to false positives if harmonics are the only metric used [4]. Section B: The Wavelet Transform To overcome the shortcomings of the Fourier transform, the Wavelet transform is used to capture the signal's temporal and frequency properties. Wavelet energy per decomposition level have been extracted in previous research. Level 3 wavelet decomposition is suitable for the Extreme Gradient Boosting (XGB) method [8], but level 5 wavelet decomposition using the Doubechies family yielded good results for Deep Neural Networks [9] and Multi-Layer Perceptron [10]. Two crucial factors were missing from these investigations' systematic procedures: the choice of wavelet family and the amount of decomposition. This procedure may be lengthy and complicated due to the indefinite number of decomposition layers and the wide variety of wavelet families that are accessible. Section C: Deep Learning-Based End-to-End Models Utilizing just the raw time series signal as input data, an end-to-end technique automates feature extraction inside the model. So far, end-to-end modeling has not been used in any of the available research on islanding detection. Due to their sequential processing and recurrent nature, LSTM architectures are slow to train, yet they were formerly thought to be state-of-the-art for time series data [11].

Meanwhile, convolutional neural networks (CNNs) may be used for time series data in the same way they are used for picture feature extraction. When it comes to cardiac signal anomaly identification, for example, state-of-the-art InceptionTime CNN models have shown effective [12][13]. The InceptionTime model, on the other hand, is computationally costly and overfitting-prone because of the intricacy of its ensemble network [18]. The implementation of a self-attention mechanism in the transformer design is another new discovery that enables data to be processed in parallel [17]. Several studies have shown that transformers outperform LSTM and CNN in terms of efficiency when it comes to computer vision and natural language processing. Since transformers do not engage in sequential or recursive data processing, they are inherently quicker than LSTMs [11]. Transformers outperform state-of-the-art CNN models in terms of robustness and training resource consumption [14, 19]. It has been shown that transformers exhibit good precision in signal processing applications with respect to performance. When compared to LSTM and CNN models, the TST has shown to be more accurate in identifying brain signals and heartbeats [15]. Therefore, one possible approach to end-to-end islanding detection that is simple, efficient, and precise is to apply a model based on transformers to time series voltage signals.

METHODOLOGY

The experiment was carried out by creating a dataset by simulating events in the power system and capturing voltage waveforms at the same time. The dataset was then partitioned into three parts: training, validation, and testing. A variety of models were built utilizing the training and validation data, including TST, InceptionTime, LSTM, and XGB. The models were then tested on the testing set and assessed using detection time and binary classification metrics. A. Creating the Dataset The data set used for this research was produced by running simulations on a SIMULINK-based modified IEEE 13-bus test system. Bus 634, which is where the home PV system is linked, is where the voltage waveform data was captured. Fig. 1 shows that distinct power system events were injected at different places.

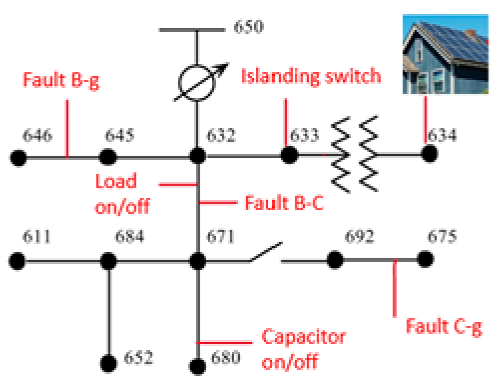

Fig. 1. Modified IEEE 13-bus test system.

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

Since the island may already be stopped by the existing over and undervoltage protection, the settings for the simulations were chosen to keep the voltages below acceptable levels. An outline of the various power system events and their random features may be found in Table I.

TABLE I. POWER SYSTEM EVENTS AND CORRESPONDING RANDOMIZED PARAMETERS

Event	Parameter Range
Load switching on/off (LoadOn, LoadOff)	Load (MVA) = [2.980, 5.950] operating at 0.8 power factor
Line-to-line and line-to-ground fault (fault_bc, fault_b, fault_c)	Fault resistance (Ω) = [0.100, 1.000]
Capacitor switching on/off (CapOn, CapOff)	Reactive power (MVar) = [3.50, 3.70]
Islanding (island)	Residential load (kVA) = [3.000, 3.300] at 0.95 power factor

Using a sample rate of 200 kHz borrowed from the high-frequency non-intrusive load monitoring system at the National Renewable Energy Laboratory (NREL) [20], each simulated event lasted for 0.05 seconds. Because the event switching timings were randomly generated, they may happen at any point in the cycle. The following number of samples were used to construct 1,000 islanding voltage waveforms: 125 for each of the seven non-islanding power system events, plus an extra 125 for normal operating circumstances. Two thousand distinct voltage waveforms are therefore included in the dataset. Figure 2 shows an example of a dataset batch.

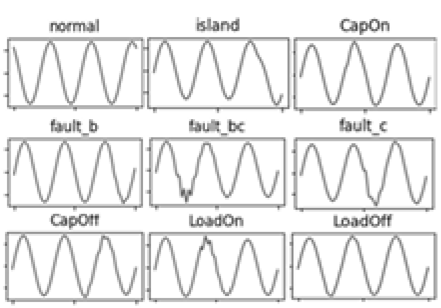

Fig. 2. Examples of event waveforms.

To determine the precision, we looked at both the original dataset and a noisy one with 40 dBSNR of Gaussian white noise. Because noise may be present in observed voltage signals, the dataset with noise was the one that was used for the remainder of the experiment. One more step: we divided the data into three sets: one for training, one for validation to keep an eye on overfitting, and one for evaluation.

B. Transform for Time Series The TST architecture's layers are shown in Fig. 3. First, the raw voltage signals are mapped to initial embeddings or feature vectors in the TST architecture. Following this, positional encodings are applied to the embeddings in order to record the correlations and timing of the observations. The model is able to pay attention to various portions of the voltage signals and capture relationships between observations at different points thanks to the encoder layer that contains the self-attention mechanism [15]. Predictions of islanding or non-islanding are made by the output layer. The TST was constructed using the tsai python module. The hyperparameters used were 16 attention heads, 128 maximum features, 0.1 dropout rate, and 0.001 learning rate, all of which were necessary for this binary classification modeling.

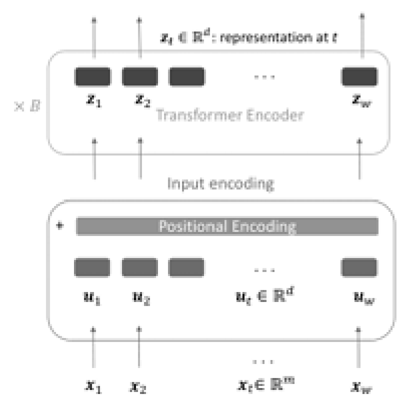

Fig. 3. TST Architecture [15].

We tested TST's efficacy against that of XGB, LSTM, and InceptionTime Coin CNN. Based on previous research, we used the xgboost python module to simulate the XGB, employing energies taken from the Daubechies wavelet family with five stages of decomposition [9]. The tsai python library's default hyperparameters were used in an end-to-end model consisting of LSTM and InceptionTime CNN. Only the hyperparameters of XGB—alpha, gamma, and maximum depth—were tuned because of the massive volume of data. Training the TST, LSTM, and InceptionTime CNN took 40 epochs, while the

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 21, Issue 2, 2025

XGB needed 500 trials to reach optimal performance. Criteria for Assessment (C.) Here are the models that were tested for accuracy, precision, and recall: (1), (2), and (3), in that order.

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \quad (1)$$

$$Precision = TP / (TP + FP) \quad (2)$$

$$Recall = TP / (TP + FN) \quad (3)$$

True negatives (TN), false positives (FP), and false negatives (FN) are defined here. The metrics described earlier were calculated using a cutoff of 0.5, which means that an event is considered islanding if its estimated likelihood score is at least 0.5. The models' performance was evaluated at different thresholds using Receiver Operating Characteristic (ROC) curves. On one side, we have the false positive rate, and on the other, we have the genuine positive rate, for different thresholds from 0 to 1. The ability of a model to differentiate between islanding and non-islanding is inversely proportional to its Area Under the Curve (AUC). Furthermore, in order to ensure compliance with the IEEE standard 1547-2018 [16], we assessed the average detection time of each model and compared it to 2 seconds. We calculated the detection time for TST, InceptionTime, and LSTM by adding the model inference code execution time on the test set to the delay from the 0.050 seconds time frame. The detected time for XGB was the sum of the timestamps for the wavelet transform, feature scaling, model inference, and 0.050 seconds time frame taken together.

RESULTS AND DISCUSSION

Here is the presentation of this section: 1) Testing how well the model performs with both clean and noisy datasets, using identical settings for the model in both instances. 2) Looking into the false positives produced by the TST and comparing the efficiency and accuracy of the models. 3) A review of the true positive-false positive tradeoff as it pertains to the ROC plot and AUC of the models may be found in the discussion. 4) Verifying how different pieces of hardware affect the model's detection timings. A. Silent vs. Surprising Signals Figure 4 displays the results of comparing the accuracy of the models trained with and without noise. Both sets of models were tested using the testing set and used the identical model setup and hyperparameters.

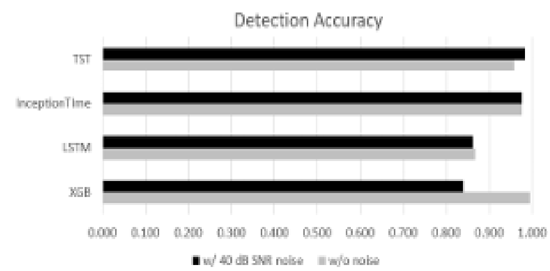

Fig. 4. Comparison of the detection accuracies of the models.

With a test accuracy of 99.5%, the XGB model demonstrated remarkable performance when trained without noise. Nevertheless, after being re-trained with noise, its performance plummeted to 84%. Because the significance of a given feature could diminish as a result of variations in the noise level, this highlights the drawback of feature extraction by hand. Under these noisy circumstances, the islanding and non-islanding were inaccurately represented by the Daubechies wavelet energies with level 5 decomposition. Both instances demonstrated consistent performance from the TST, InceptionTime, and LSTM models, demonstrating the benefits of an end-to-end strategy. Even when faced with noise, these models are able to autonomously extract features. While InceptionTime achieved an impressive 97.5% accuracy, the LSTM model fared the poorest with just around 87%. Training the TST with the noisy dataset increased its accuracy from 95.8% to 98.3%, suggesting that the model trained on clean data may have been overfit. Better generalization and increased accuracy were the outcomes of adding noise, which inhibited the model from learning complicated patterns. When all models trained with the noisy dataset were compared, the TST obtained the best accuracy. B. Measures for Assessment In Figure 5, we can see the models' confusion matrices, where islanding is represented by 1 and non-islanding by 0. Among all the models, the TST was the only one to accurately identify every

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 21, Issue 2, 2025

instance of islanding—there were no false negatives. The absence of false positives in LSTM and InceptionTime, however, suggests that they accurately identified every instance of non-islanding. XGB outperformed LSTM in identifying real islanded circumstances, despite its mixed false positive and negative performance.

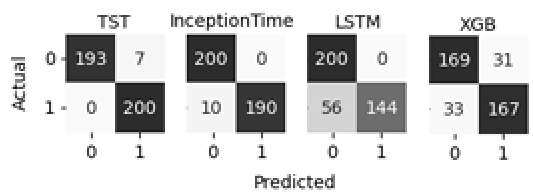

Fig. 5. Confusion matrix of the four models.

Each model's recall and accuracy, as determined by equations (2) and (3) using Fig. 5, are shown in Table II. The LSTM and InceptionTime models both hit 100% accuracy, which means they won't mistakenly label a normally linked grid situation as an isolated one. If this were to happen, the inverter would be unable to export electricity to the grid because of the false positive. The TST, in contrast, had zero false negatives and a perfect recall rate. A false negative occurs when the inverter does not identify an isolated situation and keeps the grid powered, endangering the safety of utility personnel. The TST is the superior model because it is more likely to avoid islanding-related dangers, and false negatives have a much more disastrous impact.

TABLE II. COMPARISON OF PRECISION AND RECALL METRICS

Model	Precision	Recall
TST	96.6%	**100.0%**
InceptionTime	100.0%	95.0%
LSTM	100.0%	72.0%
XGB	84.3%	83.5%

Two types of fault events—line-to-ground phase C and line-to-line phase B-C—led to seven false positives on the TST. The lower probability ratings (0.509 to 0.638) for these occurrences suggest that the TST had trouble accurately classifying them, in contrast to the mean probability score of 0.948 for all the successfully categorized events. With a likelihood value of 1.000, the phase C ground fault would be an exception. During the 0.0500 second interval, this error happened at the 0.0494 second point. The time remaining for the distortion to fluctuate was so little that the TST failed to accurately detect the incident. Increasing the number of failure occurrences included in the training dataset as the time frame approaches its end might help reduce this effect. C. ROCK True positive and false positive rates are shown in relation to each other when the threshold setting is reduced from 1 in the bottom left corner to 0 in the top right corner in Fig. 6.

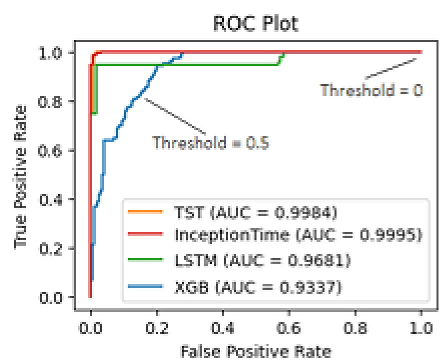

Fig. 6. ROC and AUC of the four models.

Even when the threshold is set extremely high, the ROC of TST and InceptionTime models demonstrates that these models keep the true positive rate high and the false positive rate low. This is due to the fact that these models' high probability scores for properly identified occurrences demonstrate their level of confidence. XGB had the lowest performance, to the point that a 0.25 threshold setting would be necessary to achieve a 90% true positive rate, while incurring a false positive rate of around 25%. When comparing the models' area under the curve, it is clear that TST and InceptionTime can distinguish between islanding and non-islanding situations with almost flawless accuracy. For example, the TST achieves a remarkable accuracy rate of 99.84% when presented with two randomly chosen events: one islanding and one non-islanding. I. Detection Duration When tested on different hardware, including (A) a cluster of high-performance computers, (B) a laptop, and (C) a mobile device, the average detection times for each model are shown in Table III. The 2-second threshold for detection was met by all models with the

ISSN 2319-5991 www.ijerst.com

Vol. 21, Issue 2, 2025

exception of the LSTM model that was used in a mobile device. Despite having poor accuracy, the XGB's quick detection time is a result of its training on low-dimensional extracted data. The TST had the quickest detection time out of the three end-to-end models. This means the TST can put a stop to inadvertent islands more quickly, which means less time for dangerous repairs and better power quality overall.

TABLE III. COMPARISON OF MODEL DETECTION TIMES

Model	Average detection time (ms)		
	AMD EPYC 7742 64-core 2.25GHz 692GB RAM (A)	Intel i5-3320M 2-core 2.60GHz 16GB RAM (B)	ARM Cortex-A76 8-core 2.20GHz 4GB RAM (C)
TST	73	120	131
InceptionTime	132	618	827
LSTM	188	959	2403
XGB	53	57	98

CONCLUSION

Finally, this research lays forth a comprehensive method for identifying distribution networks that have unintentionally experienced islanding by using the TST. As a first-of-its-kind, straightforward, and dependable detection model, TST trumps XGB, LSTM, and InceptionTime when it comes to islanding detection. Specifically when trained with noisy datasets, the research shows that the TST model is more resilient and reliable, reaching an accuracy of 98.3%. Additionally, the TST model has the best recall (100 percent) according to the experiment, which means that all islanding scenarios are more likely to be recognized. The TST's short detection time makes it a promising alternative to more conventional security measures, which might find use in less powerful edge devices. Inverter microcontrollers and other power-constrained devices could benefit from the TST in future research when implemented using the small ML framework.

REFERENCES

[1]. A. Hoke, “Smart inverter utility experience in Hawaii,” National Renewable Energy Laboratory, Atlanta, Georgia, August 2019.

[2]. D. Narang, S. Gonzalez and M. Ingram. “A Primer on the Unintentional Islanding Protection Requirement in IEEE Std 1547 2018,” National Renewable Energy Lab., United States, Tech. Rep. NREL/TP-5D00-77782, April 2022.

[3]. A. Hussain, C. -H. Kim and A. Mehdi, "A comprehensive review of intelligent islanding schemes and feature selection techniques for distributed generation system," in IEEE Access, vol. 9, pp. 146603 146624, 2021.

[4]. A. Ezzat, B. E. Elnaghi, and A. A. Abdelsalam, ‘‘Microgrids islanding detection using Fourier transform and machine learning algorithm,’’ Electric Power Systems Research, vol. 196, Elsevier, Jul. 2021.

[5]. A. A. Abdelsalam, A. A. Salem, E. S. Oda and A. A. Eldesouky, "Islanding detection of microgrid incorporating inverter based DGs using Long Short-Term Memory network," in IEEE Access, vol. 8, pp. 106471-106486, 2020.

[6]. K. Maresch, G. Marchesan and L. F. Freitas-Gutierres, "Passive islanding detection based on angular velocity harmonic patterns with perceptron neural network," in IEEE Latin America Transactions, vol. 19, no. 10, pp. 1665-1673, Oct. 2021.

[7]. A. K. Ozcanli and M. Baysal, “Islanding detection in microgrid using deep learning based on 1D CNN and CNN-LSTM networks,” Sustainable Energy, Grids and Networks, vol. 32. Elsevier BV, p. 100839, Dec. 2022

[8]. M. A. Ajith and R. M. Shereef, "Islanding detection of inverter based distributed generations using Wavelet transform and KNN," 2022 IEEE International Conference on PESGRE, Trivandrum, India, pp. 1-6, 2022.

[9]. M. A. Khan, V. S. Bharath Kurukuru, A. Haque and S. Mekhilef, "Islanding classification mechanism for grid-connected photovoltaic systems," in IEEE Journal of Emerging and Selected Topics in Power Electronics, vol. 9, no. 2, pp. 1966-1975, April 2021.

[10]. A. Najar, H. K. Karegar and S. Esmaeilbeigi, "Intelligent islanding

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 21, Issue 2, 2025

detection scheme for microgrid based on deep learning and Wavelet transform," 2020 10th Smart Grid Conference (SGC), Kashan, Iran, 2020, pp. 1-5

[11]. A. Orvieto, S. L. Smith, A. Gu, A. Fernando, C. Gulcehre, R. Pascanu, S. De, "Resurrecting recurrent neural networks for long sequences," arXiv, 2023

[12]. H. J. Crocker and A. W. Costall, "An InceptionTime-inspired convolutional neural network to detect cardiac abnormalities in reduced-lead ECG data," 2021 Computing in Cardiology (CinC), Brno, Czech Republic, 2021, pp. 1-4

[13]. D. Borra, A. Andaló, S. Severi and C. Corsi, "On the application of convolutional neural networks for 12-lead ECG multi-label classification using datasets from multiple centers," 2020 Computing in Cardiology, Rimini, Italy, 2020, pp. 1-4

[14]. Y. Bai, J. Mei, A. Yuille, C. Xie, "Are transformers more robust than CNNs?" arXiv, 2021

[15]. G. Zerveas, S. Jayaraman, D. Patel, A. Bhamidipaty, C. Eickhoff, "A transformer-based framework for multivariate time series representation learning," Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, August 2021, pp. 2114-2124.