Preprints.org

Article

Advanced ENF Region Classification Using UniTS-SinSpec: A Novel Approach Integrating Sinusoidal Activation Function and Spectral Attention

Yujin LI , Tianliang LU [*] , Gaojun ZENG , Kai ZHAO , Shufan PENG

Article

Advanced ENF Region Classification Using UniTS-SinSpec: A Novel Approach Integrating Sinusoidal Activation Function and Spectral Attention

Yujin Li, Tianliang Lu *, Gaojun Zeng, Kai Zhao and Shufan Peng

School of Information Network Security, People's Public Security University of China, Beijing, China

* Correspondence: lutianliang@ppsuc.edu.cn

Abstract: The electric network frequency (ENF), often referred to as the industrial heartbeat, plays a crucial role in the power system. In recent years, it has found applications in multimedia evidence identification for court proceedings and audio-visual temporal source identification. This paper introduces a ENF region classification model named UniTS-SinSpec within the UniTS framework. The model integrates the si-nusoidal activation function and spectral attention mechanism while also redesigning the model frame-work. Training is conducted using a public dataset on the open science framework platform (OSF), with final experimental results demonstrating that after parameter optimization, the UniTS-SinSpec model achieves an average validation accuracy of 97.47%, surpassing current state-of-the-art and baseline models. Accurate classification can significantly aid in ENF temporal source identification. Future re-search will focus on expanding dataset coverage and diversity to verify the model's generality and ro-bustness across different regions, time spans, and data sources. Additionally, it aims to explore extensive application potential of ENF region classification in preventing crimes such as telecommunications fraud, terrorism, and child pornography.

Keywords: electric network frequency; ENF classification; sinusoidal activation function; spectral attention; UniTS-SinSpec model

1. Introduction

The ENF is produced by the operation of alternating current generators[1], and the nominal values of ENF exhibit variation across different countries. Table 1 provides a selection of nominal voltage and ENF statistics for various countries globally[2].

Table 1. Nominal voltage and ENF of the world's nominal values extract schematic table

Region	Voltage	Frequency
Brazil	110/220 V	60 Hz
China	220V	50 Hz
Germany	230V	50 Hz
Japan	100V	50/60 Hz
USA	120V	60 Hz
Myanmar	230V	50 Hz

Due to the limited capacity of power grids to store electricity, inconsistencies between production power and consumption power are primarily caused by factors such as load changes, generator instability, transmission line faults, or dispatch issues[1,3–8]. Adi Hajj-Ahmad et al .[5] have pointed out that the fluctuation range of ENF signals is generally related to the size of the power

grid. Therefore, the fluctuation patterns of power grids differ across countries and even within different regions of the same country, leading to a widespread interest in studying the regional classification of ENF based on these differences. The ENF varies among different regions and due to its unforgeable characteristics; these differences cannot be fabricated. Notably, researchers like Wei-Hong Chuang et al.[9] attempted to manipulate ENF through specific editing techniques; however, experimental results indicated that these editing attempts could be readily detected.

Contemporary approaches for regional classification heavily leverage signal processing and machine learning technologies. While conventional signal processing techniques' efficacy hinges upon operator expertise, machine-learning methodologies enhance automation and precision in categorization but remain reliant on manually engineered features which hinder capturing intricate data insights. The advent of deep-learning technology has ushered neural network-based methodologies into ENF regional classifications by harnessing sophisticated feature extraction capacities for automatic acquisition of complex data representations. Nonetheless, extant deep-learning approaches encounter constraints when leveraging frequency domain information alongside periodic signal-processing capabilities.

In addressing the challenges posed by traditional signal processing techniques and the limitations of machine learning in ENF regional classification tasks, we refined and advanced the UniTS model through both training and testing using the OSF public dataset.

- We conduct a comparative analysis of the sinusoidal activation function in relation to commonly utilized activation functions. Our investigation focuses on the impact of activation functions on processing periodic signals, revealing that the sinusoidal activation function leverages the periodic nature of the sinusoidal function to optimize retention and extraction of periodic features within ENF data, thereby enhancing its efficacy in processing periodic signals.
- Develop a spectral attention module and integrate it into the UniTS. The spectral attention mechanism is utilized to effectively capture crucial frequency information in the original ENF data within the frequency domain, thereby addressing the model's limited noise resistance to raw data and inadequacies in extracting and learning frequency domain information.
- We propose a standardized dataset configuration and preprocessing approach for ENF classification based on the categorization of ENF regions. Due to the diverse research objectives in previous studies on ENF, variations exist in both the datasets utilized and the processing methods employed. Our proposal aims to establish a unified dataset configuration and preprocessing approach that is better suited for the task of classifying ENF regions.
- In order to validate the efficacy of the enhanced model and investigate the factors influencing its performance, a total of three experiments were devised in this study, with each experiment being further subdivided into multiple groups according to its specific objectives. These objectives encompassed an assessment of the fundamental classification performance of the UniTS model, an ablation experiment on the novel enhancements, and a hyperparameter search for the refined model. The ultimate average classification accuracy achieved by the UniTS-SinSpec model was 97.47%.

2. Related Work

Conventional approaches for frequency domain classification of the power grid primarily encompass signal processing and machine learning methodologies. Signal processing techniques entail the extraction of frequency domain features followed by manual comparison and identification. For instance, employing band-pass filters to extract fundamental ENF and harmonic information, as well as utilizing spectral analysis to scrutinize the ENF signal [10]. Furthermore, investigations have employed high-resolution spectral estimation methods based on eigenvalue decomposition, such as Multiple Signal Classification (MUSIC), as well as parameter estimation methods based on the signal subspace rotational invariance principle, such as Estimation of Signal Parameters via Rotational Invariance Techniques (ESPRIT), to analyze the characteristics of the ENF signal for grid classification and localization [11,12]. While these approaches can partially extract features from the ENF signal

and conduct regional classification, their efficiency and accuracy are heavily reliant on operator expertise.

Machine learning-based approaches have been utilized for ENF classification. These methods leverage empirical knowledge to extract features from data and subsequently employ classification algorithms. While support vector machines (SVM), random forests (RF), and K-nearest neighbors (KNN) are common machine learning techniques [12,13], their reliance on manually designed features and combination strategies poses challenges in fully capturing the long-term and cyclical nature of ENF as well as achieving efficient and accurate classification, despite some improvements in automation and accuracy.

The advancement of deep learning technology has led to the integration of neural network-based approaches in ENF classification. These methods excel at automatically extracting intricate features from data and demonstrating strong performance in classifying tasks. Notably, Georgios Tzolopoulos et al. [14] achieved outstanding results through a multi-classifier fusion framework for regional classification. This approach leverages convolutional neural network (CNN) to extract spatial features via convolutional layers; long short-term memory network (LSTM) to capture long-term dependencies within time series data; and SVM for handling high-dimensional data. The collective performance is commendable with a final accuracy rate of 96%.

In February 2024, Shanghua Gao et al. introduced the UniTS model [15]. This model is a unified multi-task time series model suitable for prediction, classification, anomaly detection, and interpolation tasks. It can handle homogeneous time series with different variables and lengths. Illustrated in Figure 1, the UniTS model consists of multiple stacked repetitive modules known as UNITS blocks and a lightweight tower designed for both mask generation and classification. Each UNITS block integrates components such as multi-head self-attention (MHSA), variable MHSA, a dynamic multi-layer perceptron (MLP), and a gating module. The mask tower is responsible for generating task-specific masks, while the classification tower (CLS) is dedicated to handling classification tasks. Cross-attention mechanism processes tokens and performs classification predictions. Tested on 38 datasets, the UniTS model outperformed other models and demonstrated zero-shot learning as well as prompt-based learning capabilities. Its design features encompass task tokenization for enhanced flexibility and adaptability by converting task specifications into uniform token representation; a unified time series architecture incorporating dynamic linear operators and dynamic MLP modules to effectively capture relationships between sequence points of varying lengths; joint training of generation and prediction tasks utilizing masked reconstruction pre-training strategy to enhance feature learning ability.
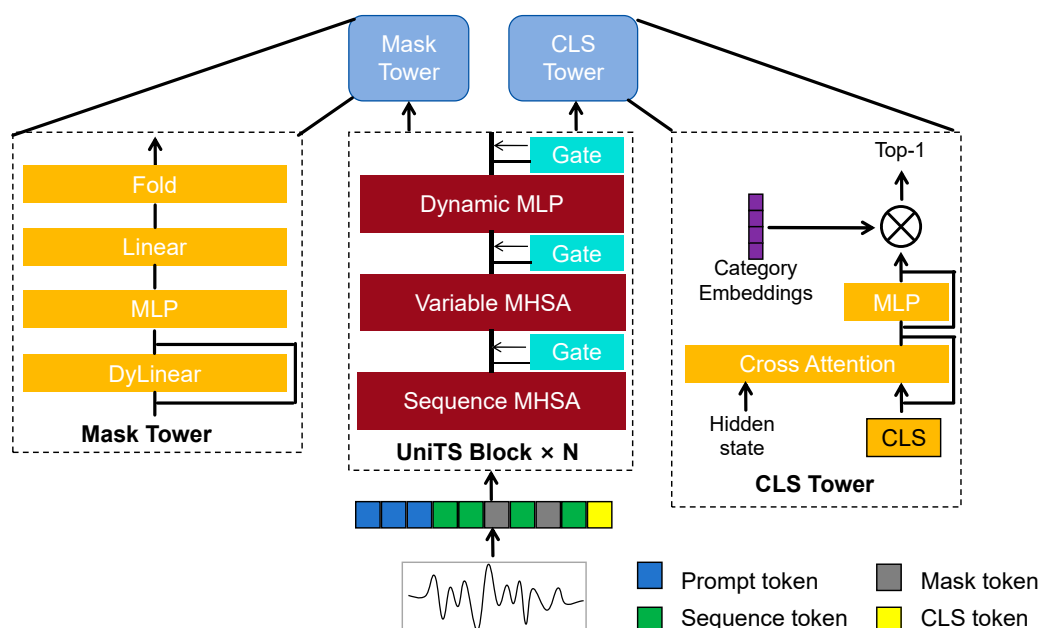

Figure 1. UniTS architecture.

The early definition of time series data can be traced back to 1927, when G. U. Yule defined time series data as a sequence of observed values recorded in chronological order [16]. This definition laid the foundation for subsequent research on time series analysis methods.

However, a more comprehensive and accurate definition of time series data was proposed by John Cochrane in 2021, who considered time series data as a sequential recording of data points within specific time intervals, enabling analysts to study variable variations over time, identify trends, seasonal patterns, and other temporal dependence structures [17]. ENF data exhibits typical characteristics of time series analysis, which were extensively discussed in the study by Jumar et al[18]. Furthermore, Kruse et al.[19] emphasized that ENF data can serve as valuable input for prediction and analysis using advanced time series models to enhance grid reliability and predictive accuracy . While ENF data falls under the category of time series data and is amenable to utilization by relevant models for specific tasks; it differs from general time series datasets in several aspects: (1) Incorporation of both temporal and frequency domain information: ENF data encompasses characteristics from both domains whereas general time-series datasets primarily focus on temporal changes and patterns; (2) Cyclical and regional characteristics: ENF records exhibit pronounced cyclical patterns, such as diurnal and seasonal variations, as well as distinct regional features. Different regions may experience varying electricity demand and generation capacity at different times. Conversely, general-time-series datasets periodicity varies based on the nature of the dataset such as seasonal sales or meteorological observations. These distinctions underscore that while ENF records are categorized under Time Series Data; their unique attributes necessitate tailored applications and analytical approaches.

3. Proposed Approaches

The UniTS model demonstrates strong performance in handling multi-task time series data. However, it exhibits limitations when applied to tasks involving ENF data: 1) The model lacks a specialized design for extracting periodic features from time series data with significant periodicity, resulting in suboptimal performance in capturing the periodic features of ENF; 2) Noise is easily introduced during the collection, transmission, and processing of ENF data. The UniTS model does not adequately consider noise resistance when processing such noisy data, consequently impacting classification accuracy; 3) ENF data encompasses not only temporal information but also rich spectral information. While the UniTS model primarily focuses on temporal features, its ability to extract and utilize spectral features is relatively limited.

To address the limitations of the UniTS model in classifying frequency data and enhance its performance, this study implemented innovative enhancements. Firstly, the original gaussian error linear unit (GELU) activation function was substituted with the sinusoidal activation function (SAF) to leverage its periodic nature for improved capture and representation of periodic features within input signals, thereby enhancing the processing efficacy of frequency data. Secondly, a spectral attention module was introduced to extract features and compute spectral attention weights in the frequency domain, enabling identification and emphasis on crucial components of frequency data. This heightened the model's sensitivity to frequency domain information and classification performance. Experimental results demonstrate that the enhanced model achieved a peak classification accuracy of 97.47% on the OSF public dataset.

3.1. Sinusoidal Activation Function

Authors should discuss the results and how they can be interpreted from the perspective of previous studies and of the working hypotheses. The findings and their implications should be discussed in the broadest context possible. Future research directions may also be highlighted.

The GELU activation function used in the UniTS model is a nonlinear activation function based on the Gaussian error function, proposed by Hendrycks and Gimpel in 2016 [20]. Its definition is given by (1).

$$GELU(x) = 0.5x(1 + tanh(\sqrt{\frac{2}{\pi}}(x + 0.044715x^3) \quad (1)$$

The GELU activation function smoothly approaches zero for negative values and gradually increases for positive values. This characteristic enables GELU to outperform traditional rectified linear unit (ReLU) activation functions in many deep learning models, as it effectively handles the negative input data, thereby enhancing model performance. The formula of the ReLU activation function is presented in (2). Nair and Hinton proposed the ReLU activation function in 2010 [21]. The ReLU activation function maintains linearity for positive values and directly outputs zero for negative values. This feature makes the ReLU activation function computationally efficient due to its simple calculation method, which reduces model complexity during training.

$$ReLU(x) = max(0, x) \qquad (2)$$

However, the GELU and ReLU activation functions have limitations in capturing periodic data features, thereby hindering the extraction of repeating patterns in such data. Specifically, periodic data often exhibits conspicuous cyclic fluctuations and repetitive structures, which may be attenuated when processed through the GELU and ReLU activation functions. This could result in a diminished capacity of the model to discern these crucial features. Hence, for datasets containing significant periodic characteristics, it might be necessary to integrate alternative methods or activation functions to more effectively capture these features.

Parascandolo and Virtanen explored the potential of the SAF in deep learning [22]. Despite being less commonly utilized compared to other activation functions, SAF has demonstrated superiority in capturing periodic features in specific scenarios. In a study by Michael S. Gashler, Stephen C. Ashmore, and their colleagues [23], researchers investigated the application of sinusoidal activation functions in deep neural networks for time series modeling problems. They initialized the network with fast fourier transform (FFT) and implemented dynamic parameter optimization, concluding that SAF effectively captures periodic patterns in data and outperforms traditional activation functions in certain tasks. Additionally, Sitzmann et al. discovered that neural networks with periodic activation functions (such as SAF) are particularly adept at representing complex natural signals and their derivatives [24]. These networks excel at fitting and representing details of time series data, making them especially suitable for processing periodic signals such as ENF.

The mathematical representation of SAF is depicted in (3). This function directly applies a sinusoidal transform to the input value x, converting the linear input into a periodic output, enabling the model to effectively capture the periodic characteristics of the input signal. Figure 2 illustrates the schematic diagrams of ReLU, GELU, and SAF.
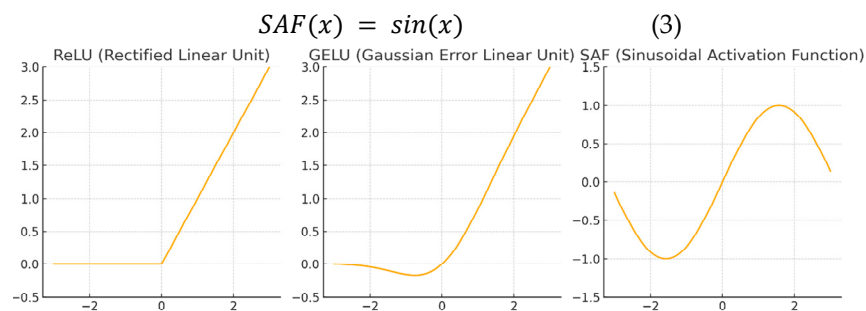
$$SAF(x) = sin(x) \qquad (3)$$

Figure 2. Illustration of activation functions.

To compare the impact of different activation functions on processing the original periodic signal, we conducted simulations using periodic signals and applied them suc-cessively to SAF, ReLU, and GELU as depicted in Figure 3. The unique sinusoidal activation mechanism of SAF ena-bles it to more effectively capture and replicate the periodic variations in the signal, thereby preserving a greater pro-portion of its original periodic characteristics during pro-cessing.
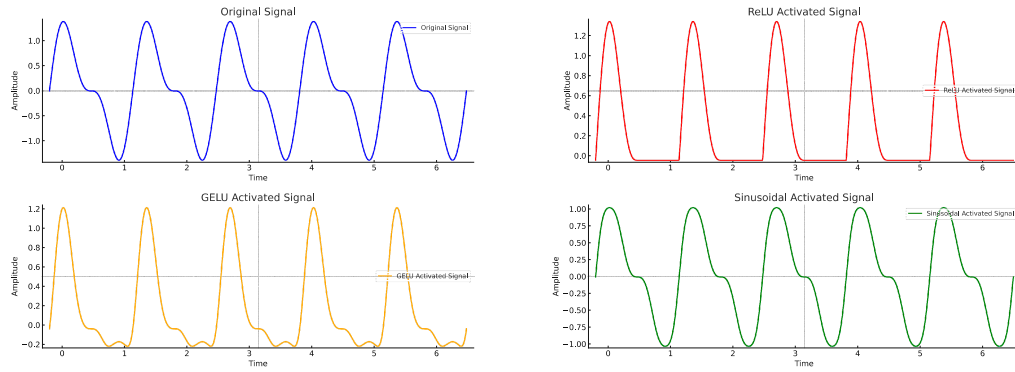
Figure 3. Comparison of activation function effects.

As previously mentioned, the ENF signal exhibits a ro-bust periodicity. In order to bolster the model's capacity for handling ENF data, we have incorporated the sinusoidal function as the activation function. Our objective is to op-timize the preservation of periodic patterns in ENF data through SAF, thereby mitigating potential signal distortion or information loss attributable to alternative activation functions and ultimately enhancing the model's proficiency in processing periodic signals.

3.2. Spectrum Attention Mechanism

A The introduction of the attention mechanism aims to replicate the human capacity for selectively focusing on specific key information while disregarding extraneous data [25].When processing ENF data containing spectral information using deep learning models, we also endeavor to selectively process crucial spectral information while filtering out irrelevant details. Therefore, we have devised a spectral attention module tailored for ENF data based on its unique characteristics and the underlying principles of the attention mechanism. The UniTS model incorporates a CLS tower (depicted in Figure 1) for classification tasks, with the inclusion of the spectral attention module preceding the category embeddings (as illustrated in Figure 4) within the CLS tower.
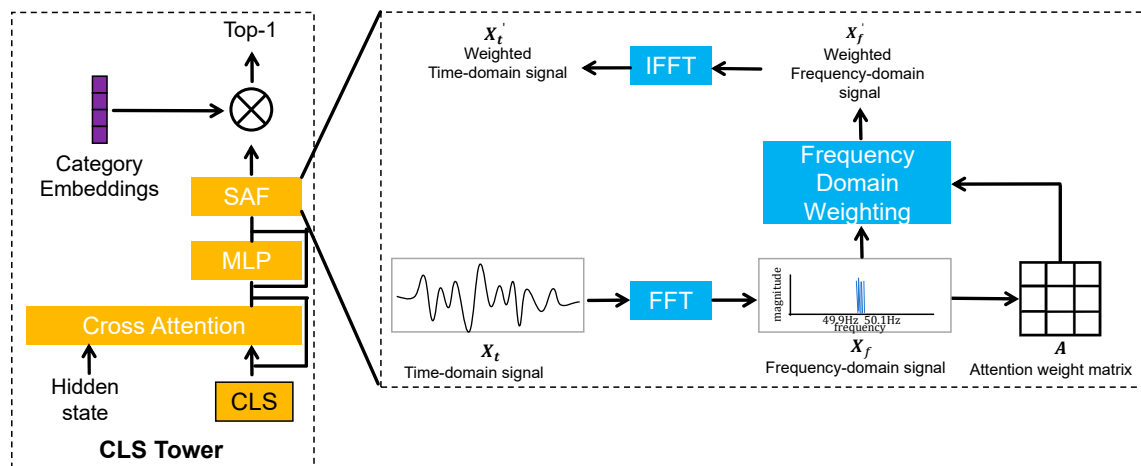
Figure 4. Spectrum Attention Mechanism Improvement Diagram.

In Europe and most of Asia, the nominal value of the ENF is 50 Hz [2]. The allowable fluctuation range typically falls within ±0.1Hz, indicating that the range of 49.9 Hz to 50.1 Hz is considered normal under standard conditions [34]. Therefore, this range becomes the focal point in our analysis of ENF classification in the frequency domain.

The fundamental concept behind our designed spectral attention module is to transform the time-domain signal into the frequency domain, focusing on the frequency range of $49.9\ Hz$ to $50.1\ Hz$. This approach enhances the model's noise robustness by allowing it to effectively capture information

within the critical frequency domain of ENF. Following attention weight calculation, the resulting attention weight matrix is multiplied by weighting and then reconverted into a time-domain signal for category embedding and classification. By concentrating on this specific frequency range, the model can better distinguish the essential features from noise, thereby improving classification accuracy and reliability.

$$A = softmax(W_2 \cdot tanh(W_1 \cdot X + b_1) + b_2 \quad (4)$$

The equation for calculating attention weight is presented in (4). Let X represent the input spectral data, with dimensions *(N, T, F)*, where N denotes the batch size, T represents the time step size, and F indicates the frequency dimension. W_1 and W_2 denote weight matrices, while b_1 and b_2 are bias vectors. The activation function $tanh$ is utilized along with $softmax$ for normalizing the weights. Subsequently, the weighted signal is derived by multiplying the attention weight matrix with the input signal; here A signifies the calculated attention weight matrix and X' denotes the weighted input signal. The module processing flow proceeds as follows:

3.2.1. Frequency Domain Transformation

Conducting FFT on the input time-series data transforms the time-domain signal into the frequency domain, resulting in a spectral plot of the signal. The frequency domain representation offers a more distinct portrayal of the frequency components within the signal, enabling the model to function within the frequency domain, as illustrated in (5). In this equation, X_t denotes the time-domain signal, while X_f signifies the frequency-domain signal.

$$X_f = FFT(X_t) \quad\quad\quad (5)$$

3.2.2. Define Frequency-Domain Attention Calculation

In the frequency range of primary focus (49.9 Hz-50.1 Hz), a linear transformation is applied to map the frequency domain signal to the attention space, and calculate the attention weight for each frequency component. The specific formula is presented in (6), where W_1 and W_2 denote the weight matrices of the linear transformation, b_1 and b_2 represent the bias vectors, tanh serves as the activation function, and softmax is utilized for normalizing the weights. Subsequently, inputting the frequency domain signal X_f yields the frequency domain attention weight matrix A.

$$A = softmax(W_2 \cdot \tanh(W_1 \cdot X_f + b_1) + b_2) \quad\quad 49.9\ Hz \leq x_f \leq 50.1\ Hz \quad\quad (6)$$

3.2.3. Frequency-Domain Weighting

The computed attention weight matrix A is multiplied with the frequency domain signal X_f to obtain the weighted frequency domain signal X_f' . This ensures that the model emphasizes crucial frequency information in the signal. As depicted in (7), X_f' represents the weighted sum of the frequency domain signal.

$$X_f' = A \cdot X_f \quad\quad\quad (7)$$

3.2.4. Inverse Transform

By applying an inverse fast Fourier transform (IFFT) to the features in the frequency domain, these features are transformed back to the time domain. This process enables the extracted frequency domain features to be integrated with the original time series data, thereby restoring the temporal characteristics of the signal. The formula for recovering the temporal characteristics is given by X_t', which represents the time-domain signal obtained through IFFT of the weighted computed frequency domain signal as depicted in (8).

4. Experimental and Results Analysis

Authors should discuss the results and how they can be interpreted from the perspective of previous studies and of the working hypotheses. The findings and their implications should be discussed in the broadest context possible. Future research directions may also be highlighted.

4.1. Data Set and Baseline Settings

4.1.1. Data Set Description and Data Preprocessing

OSF[26] provides power and audio recordings from various global grids, and this study utilizes multiple pub-licly available datasets offered by the OSF. To ensure comparable data volumes across different regions, we se-lected datasets based on their temporal coverage and divided them into two groups, as detailed in Table 2.

Table 2. Original data set.

Time Span	Location of Measurements	Synchronous Area	Time Range of Data Set	Selected Range	Frequency	Resolution	Number of Days
1 year	Baden-Württemberg, Germany	TransnetBW	201107-202003	20180101-20191231	50 *Hz*	1 sec	360
	London, United Kingdom	National Grid	201401-201912	20180101-20191231	50 Hz	1 sec	360
	Zealand, Denmark	Nordic Grid	20180101-20181231	20180101-20181231	50 Hz	1 sec	360
41 days	Karlsruhe, Germany	Continental Europe	20190709 - 20190818	20190709 - 20190818	50 Hz	1 sec	41
	Oldenburg, Germany	Continental Europe	20190710 - 20190807	20190710 - 20190807	50 Hz	1 sec	41
	Istanbul, Turkey	Continental Europe	20190709 – 20190818	20190709 - 20190818	50 Hz	1 sec	41
	Lisbon, Portugal	Continental Europe	20190709 - 20190816	20190709 - 20190816	50 Hz	1 sec	41

To standardize the data format, we preprocessed the da-tasets from seven regions uniformly, resulting in six sets of datasets as detailed in Table 3. The specific processing methods are as follows:

1. Time sequence alignment: Arranging all regional data sets chronologically to preserve contextual time-related information of ENF data;
2. Format alignment: Standardizing all data to actual measured values rather than differences between measured and reference values;
3. Missing value treatment: Replacing all "Nan" values with the median of each region's entire dataset. According to OSF's dataset description, "Nan" values may result from measurement device or calculation errors. Instead of simply replacing missing data with nominal values, which may introduce bias, we opted for replacing all "Nan" values with medians to accurately reflect each dataset's operational characteristics without affecting central trend;
4. Time span grouping: Grouping datasets based on time spans; longer time spans provide more contextual time-related information;
5. Sequence length division: Significant fluctuation pattern differences exist in ENF between daytime and nighttime due to changes in power demand and supply. For instance, electricity demand is typically higher during the day, especially on workdays, leading to a higher frequency. At night, both electricity demand and frequency decrease [32]. Diurnal variation in wind speed also affects ENF; higher wind speeds occur during the day while lower speeds occur at night [33]. To investigate sequence length impact, we selected 1 minute (60 seconds), 1 hour

(3600 seconds), and 1 day (86400 seconds) as dataset sequence lengths—each group divided into sequences of 60, 3600, and 86400 respectively;

6. Training set and test set division: Each group is split into training and test sets at an 8:2 ratio.

Table 3. Preprocessed data set.

Time Span	Region and Label	Sequence Length	Dataset Number
1 year	Baden-Württemberg, Germany-0 London, United Kingdom-1 Zealand, Denmark-2	60	A1
		3600	A2
		86400	A3
41 days	Karlsruhe, Germany-3 Oldenburg,Germany-4 Istanbul,Turkey-5 Lisbon,Portugal-6	60	B1
		3600	B2
		86400	B3

4.1.1. Baseline Setting

When employing machine learning for regional ENF classification, researchers utilize data primarily characterized by regional disparities, diverse data collection methodologies, and variations in nominal frequencies. To ensure fair and accurate comparisons with our study's findings against existing research outcomes necessitated screening prevailing machine learning models' conclusions based on three prerequisites: a minimum of three classification regions; a nominal frequency set at 50 Hz; direct acquisition from power grids.

In response to advancements in deep learning techniques , we endeavored to apply deep learning for regional ENF classification. We selected transformer_huggingface transformer time-series classification model (TH_tsc)[27], LSTM_tsc model[28], along with UniTS base model as benchmark models before training them on our preprocessed dataset. Subsequently modifying solely SAF led us to name it as 'UniTS_SAF', while incorporating spectral attention mechanism improvement module resulted in 'UniTS_SAM'. Additionally introducing both modifications concurrently produced an enhanced version termed 'UniTS-SinSpec'.

The aforementioned benchmark configurations facilitated selection of baseline models meeting comparative criteria detailed in Table 4.

Table 4. Baseline model.

Model Category	Model
machine learning	• power line data based grid identification using signal processing (2016)[29] • location identification using power and audio data based on temporal variation of electric network frequency and its harmonics (2018)[30] • power grid estimation using electric network frequency signals (2019)[31] • on spectrogram analysis in a multiple classifier fusion framework for power grid classification using electric network frequency (2024)[14]
deep learning	• LSTM_TSC(2016)[28] [1] • TH_TSC (2018)[27] [2] • UniTS(2024)[15] • UniTS_SAF [3] • UniTS_SAM [4] • UniTS-SinSpec

10

[1] Transformer-HuggingFace's transformer time-series classification model - TH_tsc[27]. [2] LSTM time-series classification model - LSTM_tsc[28]. [3] UniTS model with only sine activation function - UniTS_SAF. [4] UniTS model with only spectral attention mechanism enhancement module - UniTS_SAM

4.2. Experimental Conditions and Experimental Design

The experimental framework utilized in this study for deep learning models comprises PyTorch 2.2.2 and torchvision 0.17.2. The computational infrastructure consists of two Intel(R) Xeon(R) Gold 6326 CPU @ 2.90GHz processors and four NVIDIA A100-PCIE-40GB graphics cards. The experimental models encompass LSTM_tsc, TH_tsc, UniTS, UniTS_SAF, UniTS_SAM, and UniTS-SinSpec as delineated in Table 4. The datasets employed are A1, A2, A3, B1, B2, and B3 as illustrated in Table 3. Distributed training is executed using PyTorch's distributed data parallel (DDP) tool while hyperparameter optimization (HPO) is conducted utilizing Optuna. The initial parameter configurations for the models are detailed in Table 5.

Table 5. Model initial parameters.

Model	Batch Size	Initial Learning Rate	Model Layers	Hidden Layers	Patch Length &Stride
UniTS	32	0.0001	512	2	16
UniTS _SAF	32	0.0001	512	2	16
UniTS _SAM	32	0.0001	512	2	16
UniTS-SinSpe c	32	0.0001	512	2	16
TH_tsc	32	0.0001	12	12	25
LSTM _tsc	32	0.0001	2	2	none

To assess the efficacy of the enhanced model and investigate the factors influencing its performance, this study designed three experiments, each comprising multiple groups aligned with specific experimental objectives. Each group underwent 20 epochs of training.

- Experiment 1: assessed the performance of the UniTS baseline model and examined the impact of temporal span and sequence length on model training. The LSTM_tsc, TH_tsc, and UniTS models were trained using preprocessed datasets A1, A2, A3, B1, B2, and B3 to establish a total of six comparative experiments.
- Experiment 2: validated the effectiveness of the improved UniTS-SinSpec model through three ablation experiments. The first group utilized UniTS_SAF, the second group used UniTS_SAM, and the third group employed UniTS-SinSpec. The dataset used in this experiment was consistent with that used in Experiment 1 where optimal training results were achieved by the UniTS baseline model.
- Experiment 3: Optuna was employed to search for hyperparameters for optimizing the performance of the UniTS-SinSpec model. The original parameters are detailed in Table 5. Similar to Experiment 1's dataset usage comparison against benchmark models including UniTS_SAF and UniTS_SAM.

To assess the efficacy of the enhanced model and investigate the factors influencing its performance, this study designed three experiments, each comprising multiple groups aligned with specific experimental objectives. Each group underwent 20 epochs of training.

- Experiment 1: assessed the performance of the UniTS baseline model and examined the impact of temporal span and sequence length on model training. The LSTM_tsc, TH_tsc, and UniTS

models were trained using preprocessed datasets A1, A2, A3, B1, B2, and B3 to establish a total of six comparative experiments.

- Experiment 2: validated the effectiveness of the improved UniTS-SinSpec model through three ablation experiments. The first group utilized UniTS_SAF, the second group used UniTS_SAM, and the third group employed UniTS-SinSpec. The dataset used in this experiment was consistent with that used in Experiment 1 where optimal training results were achieved by the UniTS baseline model.
- Experiment 3: Optuna was employed to search for hyperparameters for optimizing the performance of the UniTS-SinSpec model. The original parameters are detailed in Table 5. Similar to Experiment 1's dataset usage comparison against benchmark models including UniTS_SAF and UniTS_SAM.

4.3. *Experimental Results and Analysis*

4.3.1. Experiment 1

The findings from Experiment 1 are presented in Table 6, demonstrating that the UniTS model exhibits significantly superior performance across all available datasets compared to other models. Additionally, models trained on datasets with larger effective sample sizes (A) generally display enhanced performance relative to those trained on datasets with smaller effective sample sizes (B). It is noteworthy that training with a shorter sequence length (60) yields poor results due to limited contextual information. Conversely, increasing the sequence length from 60 to 3600 leads to improved model performance, as the longer sequence provides more comprehensive context for enhancing classification accuracy. However, both the UniTS and TH_tsc models cannot be effectively trained on extremely long sequences (86400) due to their complex architecture and high computational resource requirements, resulting in an error message indicating large tensor size or insufficient data. In contrast, the LSTM_tsc model's relatively simple recurrent structure and lower computational resource demands enable it to continue training with extremely long sequences; however, its performance is subpar.

Table 6. Results of experiment 1.

Data Set Number	Number of Valid Samples	Sequence Length	UniTS	TH_tsc	LSTM_tsc
A1	1575357	60	88.95%	57.86%	26.6%
A2	188928	3600	91.49%	58.29%	32.54%
A3	1046	86400	none	none	6.25%
B1	188921	60	77.03%	48.93%	24.02%
B2	22656	3600	79.61%	50.21%	29.37%
B3	124	86400	none	none	5.64%

4.3.2. Experiment 2

The results of Experiment 1 demonstrate that the model's overall training effect on dataset A2 is superior, leading to the selection of dataset A2 for Experiment 2. The findings from Experiment 2 are presented in Table 7, where three ablation experiments were conducted. These included the UniTS_SAF model with only sinusoidal activation function, the UniTS_SAM model with solely spectral attention mechanism, and the UniTS-SinSpec model combining both. The experimental outcomes reveal that the average validation accuracy of the UniTS-SinSpec model reaches 96.24%. Furthermore, ablation experiment results indicate significant improvements in both UniTS_SAF and UniTS_SAM compared to the original UniTS model. This underscores a substantial enhancement in performance due to improvement efforts, with the combined UniTS-SinSpec model exhibiting optimal performance. However, it is noteworthy that this enhancement has led to an increase in training time.

Table 7. Results of experiment 2.

Index	UniTS	UniTS_SAF	UniTS_SAM	UniTS-SinSpec
Time-consuming Average	33.28s	44.01s	45.674s	46.549s
Verification Accuracy	91.49%	94.43%	94.11%	96.24%

4.3.3. Experiment 3

After conducting HPO with Optuna, we found that the batch size of 16, model depth of 256, hidden layer depth of 1 layer, patch size and stride of 300, resulted in the best performance for the UniTS-SinSpec model. The final results of Experiment 3 are shown in Table 8.

The results demonstrate that the UniTS-SinSpec model, optimized through HPO, achieves an average accuracy rate of 97.47% in classifying ENF data, surpassing significantly both the baseline model and other enhanced models. Notably, the optimized UniTS-SinSpec model exhibits a nearly 6 percentage point improvement in accuracy compared to the baseline UniTS model.

It is noteworthy that the optimal patch length & stride step of 300, corresponding to inputting ENF data sequentially at 5-minute intervals, yields the model's highest classification performance and an average reduced running time of approximately 9.0597 seconds. We posit that due to its one-dimensional nature, ENF data exhibits lower complexity, thus a reduction in model complexity and an appropriate patch length & stride step can effectively enhance both model performance and efficiency.

Table 8. Model performance comparison.

Model Category	Model Name	Average Accuracy
machine learning	• power line data based grid identification using signal processing (2016)[29]	88.23%
	• location identification using power and audio data based on temporal variation of electric network frequency and its harmonics (2018)[30]	88.67%
	• power grid estimation using electric network frequency signals (2019)[31]	88.23%
	• on spectrogram analysis in a multiple classifier fusion framework for power grid classification using electric network frequency (2024)[14]	96%
deep learning	• LSTM_tsc	32.54%
	• TH_tsc	57.86%
	• UniTS	91.49%
	• UniTS_SAF	94.43%
	• UniTS_SAM	94.11%
	• UniTS-SinSpec	96.24%
	• UniTS-SinSpec (HPO via Optuna)	97.47%

4.4. Discussion and Conclusion

The experimental results demonstrate that the enhanced UniTS-SinSpec model exhibits significantly improved classification accuracy and speed compared to the baseline model and other enhanced models. Specifically, Experiment 1 indicates that the UniTS model outperforms the traditional LSTM_tsc and TH_tsc models on the selected dataset. Through an analysis of the impact of different time spans and sequence lengths on model performance, it was observed that longer time series can offer more contextual information, with one hour being identified as most suitable, leading

to a significant enhancement in classification performance. However, excessively long sequence length may result in insufficient data or large tensors, thereby affecting the training effectiveness of the model. Experiment 2 reveals that both the sinusoidal activation function and spectral attention module notably enhance model performance, with the UniTS-SinSpec model incorporating both demonstrating superior performance, thus validating the efficacy of this enhanced approach. Nevertheless, it is noted that this improved model requires a longer training duration, highlighting a trade-off between performance improvement and computational resources which needs consideration in practical applications. Finally, Experiment 3 optimizes hyperparameters for UniTS-SinSpec using Optuna; resulting in an optimized average accuracy of 97.47% for ENF data classification task—significantly surpassing both baseline and other enhanced models. It was discovered that lower model complexity actually contributes to improved classification performance for ENF data; furthermore, when inputting ENF data in sequences of 5 minutes (300), optimal classification performance is achieved with an average runtime of only 9.0597 second.

5. Conclusion and Prospects

This paper presents an enhanced model, UniTS-SinSpec, designed for ENF region classification based on the UniTS model. The model is trained on preprocessed data from a publicly available dataset and then optimized through hyperparameter search. The key contributions of this study are as follows:

- We propose an effective method for ENF region classification that fully leverages frequency domain information and periodic features. The efficacy of the enhancement is validated through ablation experiments, and further improvements in model performance are achieved by optimizing hyperparameters. The final UniTS-SinSpec model attains an average classification accuracy of 97.47%.
- Discussing and proposing data preprocessing methods: We advocate for a standardized dataset for ENF domain classification, encompassing temporal sequence uniformity, data format consistency, handling of missing values, time span regularity, and sequence length consistency.
- We propose the following suggestions for tuning model parameters: It is our belief that the model should not be overly complex for the task of ENF region classification, given that the ENF data is one-dimensional and increased model complexity may lead to decreased performance. Notably, extending the time span of the dataset can enhance training effectiveness, and appropriately adjusting the length of the time series in the dataset can optimize model performance.

Finally, we acknowledge that there are still certain limitations and areas for improvement in the study. Firstly, the experimental dataset utilized in this study primarily originates from OSF, potentially leading to insufficient data coverage and diversity. Future studies could consider incorporating additional datasets from diverse regions and time periods to validate the model's generalizability and robustness. Secondly, all data used in this study are grid measurement data. Subsequent research endeavors could explore classifying ENF data obtained through multimedia or electromagnetic radiation to augment the practical applicability of the model. Furthermore, future investigations will delve into exploring potential applications of the UniTS-SinSpec model in preventing and combating crimes such as telecommunications fraud, terrorism, and child pornography by identifying specific ENF features and patterns associated with criminal activities.

References

1. D. P. Nicolalde Rodriguez, J. A. Apolinario, and L. W. P. Biscainho, "Audio Authenticity: Detecting ENF Discontinuity With High Precision Phase Analysis," IEEE Transactions on Information Forensics and Security, vol. 5, no. 3, pp. 534-543, 2010. doi:10.1109/TIFS.2010.2051270.

2. National Technical Information Service, "NTIS Homepage," [Online]. Available: https://www.ntis.gov/index.xhtml#.

3. M. Kajstura, A. Trawinska, and J. Hebenstreit, "Application of the Electrical Network Frequency (ENF) Criterion: A case of a digital recording," Forensic Science International, vol. 155, no. 2-3, pp. 165-171, 2005.

4. M. Huijbregtse and Z. Geradts, "Using the ENF criterion for determining the time of recording of short digital audio recordings," Forensic Science International, vol. 175, no. 2-3, pp. 148-157, 2008.

5. R. Adi Hajj-Ahmad, Z. Geradts, and M. Wu, "Instantaneous Frequency Estimation and Localization for ENF Signals," in Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), pp. 2877-2881, 2012.

6. C. Grigoras, "Applications of ENF criterion in forensic audio, video, computer and telecommunication analysis," Forensic Science International, vol. 167, no. 2-3, pp. 136-145, 2006.

7. Yao Liu, Zhong Yuan, Paul N. Markham, Roger Conners, and Yuguang Liu, "Application of Power System Frequency for Digital Audio Authentication," IEEE Transactions on Power Delivery, vol. 27, no. 4, pp. 1820-1828, 2012.

8. R. Garg, A. Hajj-Ahmad, M. Wu, "Geo-Location Estimation from Electrical Network Frequency Signals," in Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), pp. 2862-2866, 2013.

9. W. H. Chuang, R. Garg, M. Wu, "Anti-Forensics and Countermeasures of Electrical Network Frequency Analysis," IEEE Transactions on Information Forensics and Security, vol. 8, no. 12, pp. 2073-2088, 2013.

10. S. Chakma and S. A. Fattah, "Location Identification Using Power and Audio Data Based on Temporal Variation of Electric Network Frequency and Its Harmonics," in Proceedings of 2018 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE), pp. 1-4, 2018. doi:10.1109/WIECON-ECE.2018.8783110.

11. W. Bang and J. W. Yoon, "Power Grid Estimation Using Electric Network Frequency Signals," Security and Communication Networks, vol. 2019, pp. 1-11, 2019. doi:10.1155/2019/1982168.

12. R. Garg, A. Hajj-Ahmad, M. Wu, "Feasibility Study on Intra-Grid Location Estimation Using Power ENF Signals," arXiv preprint arXiv:2105.00668, 2021. doi:10.48550/arXiv.2105.00668.

13. I. Khairy, "ENF based classification and extraction," IEEE Signal Processing Cup 2016, [Online]. Available: https://github.com/ibrahimkhairy/ENF_based_classification_and_extraction. [Accessed: Jun. 27, 2024].

14. G. Tzolopoulos, C. Korgialas, and C. Kotropoulos, "On Spectrogram Analysis in a Multiple Classifier Fusion Framework for Power Grid Classification Using Electric Network Frequency," ArXiv abs/2403.18402, 2024.

15. Shuang Gao, Tarik Koker, Oliver Queen, Thomas Hartvigsen, Theodoros Tsiligkaridis, and Marinka Zitnik, "UNITS: A Unified Multi-Task Time Series Model," 2024.

16. G. U. Yule, "On a Method of Investigating Periodicities in Distributed Series, with special reference to Wolfer's Sunspot Numbers," Phil. Trans. R. Soc. London A, vol. 226, pp. 267-298, 1927. doi:10.2307/91170.

17. Y. Kim, "Time Series Analysis for Macroeconomics and Finance," 2024.

18. R. Jumar, H. Maaß, B. Schäfer, L. R. Gorjão, and V. Hagenmeyer, "Database of Power Grid Frequency Measurements," arXiv preprint arXiv:2006.01771, 2020.

19. T. Kruse, L. Vanfretti, and F. Silva, "Data-Driven Trajectory Prediction of Grid Power Frequency Based on Neural Models," Electronics, vol. 10, no. 2, p. 151, 2021. doi:10.3390/electronics10020151.

20. D. Hendrycks and K. Gimpel, "Gaussian Error Linear Units (GELUs)," arXiv preprint arXiv:1606.08415, 2016. doi:10.48550/arXiv.1606.08415.

21. V. Nair and G. E. Hinton, "Rectified Linear Units Improve Restricted Boltzmann Machines," in Proceedings of the 27th International Conference on Machine Learning (ICML-10), pp. 807-814, 2010.

22. G. Parascandolo and T. Virtanen, "Taming the Waves: Sine as Activation Function in Deep Neural Networks," 2016.

23. M. S. Gashler and S. C. Ashmore, "Training Deep Fourier Neural Networks to Fit Time-Series Data," arXiv preprint arXiv:1405.2262, 2014. doi:10.48550/arXiv.1405.2262.

24. Vincent Sitzmann, Julien N. P. Martel, Alexander W. Bergman, David B. Lindell, and Gordon Wetzstein, "Implicit Neural Representations with Periodic Activation Functions (SIREN)," Cross Validated, ar5iv, 2020.

25. S. Zhou and Y. Pan, "Spectrum Attention Mechanism for Time Series Classification," in Proceedings of 2021 Chinese Control and Decision Conference (CCDC), pp. 1-6, 2021. doi:10.1109/DDCLS52934.2021.9455697.

26. "Power Grid Frequency," [Online]. Available: https://power-grid-frequency.org/. [Accessed: June 27, 2024].

27. Hugging Face, "Time Series Transformer Classification," [Online]. Available: https://huggingface.co/keras-io/timeseries_transformer_classification/tree/main. [Accessed: June 27, 2024].

28. R. Romijnders, "LSTM Time Series Classification," [Online]. Available: https://github.com/RobRomijnders/LSTM_tsc. [Accessed: June 27, 2024].

29. S. Chakma, D. Chowdhury, M. Sarkar, "Power line data based grid identification using signal processing," in Proceedings of 2016 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE), pp. 1-4, 2016. doi:10.1109/WIECON-ECE.2016.8009106.

30. S. Chakma and S. A. Fattah, "Location Identification Using Power and Audio Data Based on Temporal Variation of Electric Network Frequency and Its Harmonics," in Proceedings of 2018 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE), pp. 1-4, 2018. doi:10.1109/WIECON-ECE.2018.8783110.

31. W. Bang and J. W. Yoon, "Power Grid Estimation Using Electric Network Frequency Signals," Security and Communication Networks, vol. 2019, pp. 1-11, 2019. doi:10.1155/2019/1982168.

32. B. Schäfer, D. Witthaut, M. Timme, and V. Latora, "Non-Gaussian power grid frequency fluctuations characterized by Lévy-stable laws and superstatistics," Nature Energy, vol. 2, p. 17058, 2017. doi:10.1038/s41560-017-0058-z.

33. R. Goss, F. Mulder, and B. Howard, "Implications of diurnal and seasonal variations in renewable energy generation for large scale energy storage," Journal of Renewable and Sustainable Energy, vol. 12, no. 4, p. 045501, 2020. doi:10.1063/5.0012921.

34. P. Tumino, "Frequency Control in a Power System," EE Power, Oct. 15, 2020. [Online]. Available: https://eepower.com/technical-articles/frequency-control-in-a-power-system/. [Accessed: Jul. 25, 2024].