RESEARCH ARTICLE

Detection and Classification of Fault Types in Distribution Lines by Applying Contrastive Learning to GAN Encoded Time-Series of Pulse Reflectometry Signals

JAVIER GRANADO FORNÁS[ID]1, **ELÍAS HERRERO JARABA**[ID]2, **(Member, IEEE),**
ANDRÉS LLOMBART ESTOPIÑAN[ID]1, **(Member, IEEE),**
AND JOSE SALDANA[ID]1, **(Senior Member, IEEE)**
1CIRCE Technology Center, 50018 Zaragoza, Spain
2Department of Electronic Engineering and Communications, University of Zaragoza, 50018 Zaragoza, Spain

Corresponding author: Javier Granado Fornás (jgranado@fcirce.es)

ABSTRACT This study proposes a new method for detecting and classifying faults in distribution lines. The physical principle of classification is based on time-domain pulse reflectometry (TDR). These high-frequency pulses are injected into the line, propagate through all of its bifurcations, and are reflected back to the injection point. According to the impedances encountered along the way, these signals carry information regarding the state of the line. In the present work, an initial signal database was obtained using the TDR technique, simulating a real distribution line using (PSCAD^{TM}). By transforming these signals into images and reducing their dimensionality, these signals are processed using convolutional neural networks (CNN). In particular, in this study, contrastive learning in Siamese networks was used for the classification of different types of faults (ToF). In addition, to avoid the problem of overfitting owing to the scarcity of examples, generative adversarial neural networks (GAN) have been used to synthesise new examples, enlarging the initial database. The combination of Siamese neural networks and GAN allows the classification of this type of signal using only synthesised examples to train and validate and only the original examples to test the network. This solves the problem of the lack of original examples in this type of signal of natural phenomena which are difficult to obtain and simulate.

INDEX TERMS Artificial Neural Networks (ANNs), deep learning, siamese networks, generative adversarial neural networks (GAN's), fault classification, fault detection, transmission lines.

I. INTRODUCTION

The automatic detection and classification of short circuits (faults) in distribution lines (especially in low-voltage installations) is a hot research topic with significant challenges ahead. Several techniques are available for the detection and classification of the type of fault (hereinafter referred to as ToF) [1], [2]. One is the analysis of the event signal (high-frequency spike) that appears when a fault occurs somewhere

in the network [3], [4], [5]. Owing to the catastrophic and chaotic nature of these events (faults), the acquisition, simulation, and analysis of these types of signals represent a very difficult problem.

Another way to detect and classify the ToF is to analyse the response to the injection of high-frequency pulses into the power grid. The different responses of the grid to these injected signals, owing to the different types, makes it possible to classify these faults.

This study presents an analysis of these signals using neural networks. The objective was to solve the main problem

The associate editor coordinating the review of this manuscript and approving it for publication was Fahmi Khalifa[ID].

of a lack of original examples. Finding a method to train a neural network model to classify faults with good accuracy and proper generalisation.

To understand this problem, it is necessary to know that faults in distributed lines [6], [7] depend on a physical phenomenon which, as mentioned above, produces a catastrophic event in the electrical network (short circuit of one or several phases to earth). This phenomenon has a chaotic and random nature (falling trees, lightning, fire, etc.). Therefore, it is a major handicap to obtain easily available real signals. By contrast, the training of a neural network [8] requires a large number of examples (typically tens of thousands). Neural networks learn to discover a particular pattern within a set of examples through a training process. In this process, the weights of the network were adjusted. Each time a new example is taught, the output of the network tries to "adjust" a little bit more and even closer to the real output (label). This fact makes it necessary to have a large number of examples that allow the network to generalise the desired pattern well. The objective is avoiding to "learn" just some few examples, which is known as *overfitting* [9].

On the other hand, the technique frequently used to deal with this phenomenon is to model mathematically the electrical network [10]. Using this model, we can produce a database of fault signals for certain scenarios and locations. However, even so, it is always a very complex task to obtain a database with typically tens of thousands of fault signals that will allow us to train a neural network with sufficient reliability; for instance, there may be a wide range of cases that produces each ToF. This implies that the electrical network may be in a very diverse state, even for the same ToF. Therefore, it is difficult to simulate all possible random conditions that would give rise to the same ToF.

However, a more plausible solution is to start from an existing database which includes several simulated faults. This database, although insufficient (hundreds of examples), is representative of the problem that needs to be solved. More details of this database are provided in subsequent sections.

It is at this point that our work starts by solving the problem of the limited amount of training data using a generative adversarial network (GAN) [11].

A GAN consists of two neural networks that are trained simultaneously. One of them, called the *Generator*, can generate an example from a random input (noise) of a certain dimension. The generated example is used as an input for another network, called the *Discriminator*. The *Discriminator*, on the other hand, receives these so-called *Fake* examples, along with others, from the original database. The latter, therefore, will be correctly labelled *Real* examples. The *Generator* attempts to generate more realistic examples, and the *Discriminator* attempts to identify whether an example is *Real* or *Fake*. With this "zero-sum" game, both networks are trained, and this process ends when the *Discriminator* is unable to tell whether the example generated by the *Generator* is *Real* or *Fake*.

From that point on, the *Generator* can be used to generate examples of this type. These have the particularity of having the main information that characterises them as belonging to that type. However, each example is different from the others because it is generated from a random signal (noise).

A final problem we will face is that although these fake images are very similar to the originals, they will not have the same distribution in common. This leads to the fact that a "standard" network, even if it is sufficiently powerful, will not be able to generalise well between different distributions.

This fact has generally been addressed in other studies by mixing the data from both distributions to enable the network to learn both distributions. In our case, on the contrary, we will tackle the problem using Siamese networks [12], [13], thus exploiting their best feature: adapting well to different distributions.

II. ENVIRONMENT

The new challenges facing modern society have resulted in an increasingly important electrification of the energy system with a growing role in distribution networks. It is therefore necessary to prepare these infrastructures so that new challenges will be faced in the future, including the installation of multiple distributed generation resources.

The detection and classification of faults in electricity grids [14], as well as their localisation, are considered essential requirements to achieve the objective of implementing the smart grid concept in electricity grids. Unfortunately, the locating task is a truly complex challenge, but it is also often an arduous task that involves the movement of specialised personnel with ground or aerial means across the entire affected section. This is the main reason because the automatic location of fault systems is very worthwhile and is the essence of this article.

Currently, a widely used option is to minimise the impact of outages by introducing self-healing mechanisms [15]. In any case, the use of this type of mechanism implies that they are strongly meshed (radial and redundant topology) and are able to isolate a faulty section until the problem is solved without the need for the entire network to be out of service. These faults are usually due to short circuits between one of the phases and the earth or between different phases. It should be borne in mind that as all the sections of a network are connected in parallel, a short circuit is transferred to the entire network that is "connected" to that section. Once the protections of the corresponding section closest to the fault are opened, leaving that section isolated, it must be possible to detect the fault and its type as well as to locate the distance to it as accurately and quickly as possible.

Alternatively, in some lines (high-voltage and medium-voltage distribution lines), there are fault passage detectors [16]. These detectors continuously monitor the voltage and current of the network. When a fault occurs, the device detects it and can supply this signal as an input for the fault locator system. However, these types of detectors are not usually deployed in low-voltage lines because of their high

cost; therefore, the detection and classification of faults is undoubtedly very useful information as a preliminary step to locating them.

In this article, our main objective is to address the task of automatically detecting/classifying faults, which is a preliminary work to solve the bigger problem of localisation, while being aware that both are parts of the same solution and have to work together.

We use neural networks to obtain promising results from the time series obtained from certain captured signals. The physical principle used to obtain these signals is time-domain reflectometry (TDR) [17], [18]. It consists of performing periodic pulse injections on the line in each of the phases of the network (R-S-T) so that the electrical response of the network is updated frequently (which will be explained in detail below). When a fault occurs, the returned signal carries implicit information regarding the ToF that has occurred. This information is used by a trained neural network to classify faults that have occurred and classify them properly.

When TDR is employed, the pulses injected into each of the phases travel throughout the network and, at each impedance change (node, junction, etc.), part of the pulse continues its journey, and another part is reflected. This reflection of the pulses makes it possible to record at the injection point all the signals that "return" after "bouncing" through the different bifurcations of the network. These pulses carry embedded information regarding the network status. When the network is operating normally, the information from the signals is different from the information they carry when a fault has occurred, and this is precisely the information that the neural network is able to extract to determine the ToF that has occurred.

III. STATE OF THE ART

As discussed in the previous section, the detection/ classification process and the localisation process are part of a more global process which would be the localisation and interpretation of faults in electricity distribution networks. In other words, they represent two different phases of the same process.

Having said that, we wanted to give a more general approach to our state of the art, and in the following, we will not only review the state-of-the-art classification methods, but also provide a summary of the localisation methods (outside the scope of this publication).

A. METHODS FOR FAULT CLASSIFICATION

In this section, we show state-of-the-art fault detection/classification. According to [19] and, [20] it comprises three major groups.

- Prominent techniques
- Hybrid techniques
- Modern techniques

Prominent techniques are most commonly used for fault classification and are further divided into three subgroups. One of the techniques that falls under this group is numerical tools based on wavelet transform (WT).

In this technique, the key is to choose a "*mother wavelet*" and then perform checks using versions of this WT. The WT can separate the signal into frequencies that can then be analysed with the help of multi resolution analysis (MRA). For example, it is possible to use discrete wavelet transform (DWT) to classify faults once the fault currents are known at a given location [21].

On the other hand, but also in the same group, artificial neural networks (ANN) are usually used to tackle with this problem. These types of systems can be trained by showing labelled examples to learn the features common to that class. These trained systems can then classify the new example shown to them.

There are some experiences of this technique implemented on MATLAB software to notice the fault on transmission lines. The output of the Simulink model was used to train the ANN to identify faults in transmission systems [22].

Also, under this group, efforts have been made to address this problem by means of Fuzzy Logic. The fuzzy-logic technique uses an easy relationship between the input and output variables. Therefore, the fuzzy logic technique performs simple control to deal with numerous issues, especially when the numerical model is not well known or is difficult to solve.

There are strategies for investigating overhead line failures based on a fuzzy system. For example, by comparing the s-transform and wavelet transform [23], it can be concluded that the fuzzy decision tree (DT) based on the s-transform provides accurate fault classification.

Hybrid techniques attempt to compensate for the shortcomings of separate methods. Neuro-fuzzy techniques have also been found in this group. These neural systems have attempted to adapt to changing situations. The combination of fuzzy inference systems that link human learning and performance approaches with certain improvement strategies has led to this technique.

One type of combination is the Stockwell transform (ST) and multilayer perceptron neural network (MLP-NN) / FFNN [24], [25] tested in a simulated IEEE 13-node test feeder. IEEE 13-node is a very small circuit model used to test common features of distribution analysis software. In the proposed technique, the three-phase current waveforms are measured from different points and then processed using ST to extract statistical features. The features are later fed into the MLP-NN / FFNN system to detect and classify the faults.

Another type of system that combines techniques is wavelet and ANN techniques. Wavelet and ANN techniques combine the features of a wavelet approach and an artificial neural network to obtain better results in fault classification. Some work has been carried out on fault classification using current signals used for thyristor-controlled series-compensated transmission systems by integrating the DWT and ANN algorithms [26].

On the other hand, the combination of wavelet and fuzzy logic is used to decompose the output signals of the simulated power grid. These output signals were then used as inputs to the fuzzy-logic block. The fuzzy logic block has the particular rules used for this example that result in the type of fault.

In [27], a method that uses the mother wavelet Daubechies4 (db4) and a combination of DWT and fuzzy logic to classify faults was developed.

Finally, wavelet and neuro-fuzzy techniques [28] can be mixed for fault detection and classification using wavelet MRA coefficients only. The fault location for a series-compensated transmission system using WT and ANFIS was developed in [29].

Modern techniques include support vector machines. This is a technique for learning separation functions in classification tasks (pattern recognition). This technique was based on statistical learning. Therefore, the input vectors were mapped nonlinearly into a high-dimensional feature space. This has been effectively applied to many classification problems, for example, by combining the SVM and wavelet techniques used to detect and classify fault types [30].

Another group is the genetic algorithm. Genetic algorithms (GA) work with variable encoding. (GA) uses a population of points at a time, in contrast to the single-point approach of traditional optimisation methods. It has been proposed that some methods contain a pre-processing unit that depends on both the DWT and GA, in which the DWT has been used to extract characteristic features from the input current signal [31].

In addition, we find Euclidean distance-based methods for this group. The Euclidean distance between successive current samples can be used for power-line fault detection and to identify the faulty phase [32].

Focusing on neural networks, which is the technique used by us, we have found some additional articles that have special relevance for the work developed in this paper:

Some studies have used neural networks to process power line voltage and current signals to classify faults [33], [34], [35], [36], [37]. The voltages and currents were digitised to form the input to the neural network. Also, there are some works to making detection and location of aged cable sections in underground lines [38]. In this case, the transfer function of the cable is used as the input of the CNN.

Some studies have also been conducted on fault detection, attempting to extrapolate the voltage values from several simulated power grids. These values were used to train a neural network to detect faults in a real electrical network [39].

We have found some studies in which convolutional neural networks (CNN) were used for processing time series. The fault signals are time-series signals, so this kind of work is very relevant for us. In particular, EEGNet, a compact CNN for the classification and interpretation of EEG-based signals (electroencephalography EEG)), has been used to classify this type of signal [40].

Another interesting work is the processing of the time signals of partial discharge phenomena in power grids using image transformations (scalogram) [41]), in which the nature of the signals is very similar to our own. The transformation of this type of signal into images allows it to be processed by convolutional neural networks (CNN).

This type of neural network provides excellent results in image processing.

Another type of transform of time signals to images presented in the literature is the Gramian angular fields (GAF) transform [42].

We studied this transformation in the present study to use them with our signals.

The state of the art of deep learning includes publications that discuss the use of a special type of network called a generative adversarial network (GAN). Using these GAN, synthesised examples can be generated from a limited database. In our case, we have the problem of the scarcity of examples, so we can probably use this type of neural network to expand our database of signals of the transients produced in the electrical grid in fault [43], [44].

As has been shown, in the literature, there are several publications proposing the use of Neural Networks for the detection of faults and other phenomena of a similar nature in electrical distribution networks.

The possibility of being able to detect and classify the ToF by injecting periodic pulses into the network is currently state-of-the-art and can be a great help as a preliminary step in locating faults in automatic systems.

B. METHODS FOR FAULT LOCATION
Two state-of-the-art techniques are available for locating faults in power lines.
- Impedance measurement
- Methods based on travelling waves

1) IMPEDANCE MEASUREMENT
The first technique involves measuring the impedance of the line at the fundamental frequency from a point on the network. The point at which it occurs can be calculated by observing how it varies when a failure occurs.

Fault location in double circuit power networks is presented in [45]. In this paper, the new method considers the mutual effect of double circuit lines It has been tested over IEEE 13-node test feeder.

However, although this method works well for transport grids because it is generally simple with no bifurcations [46], it does not yield good results in distribution lines. In these lines, there are numerous bifurcations, and it is not possible to equip each section with a detector because of the high cost.

However, the measurement of these networks from a single point results in a large error because of the lack of proportionality between the impedance value and the distance caused by the multiple bifurcations of the network.

2) METHODS BASED ON TRAVELLING WAVES
Methods based on travelling waves are based on the propagation of waves through conductors. These methods are based
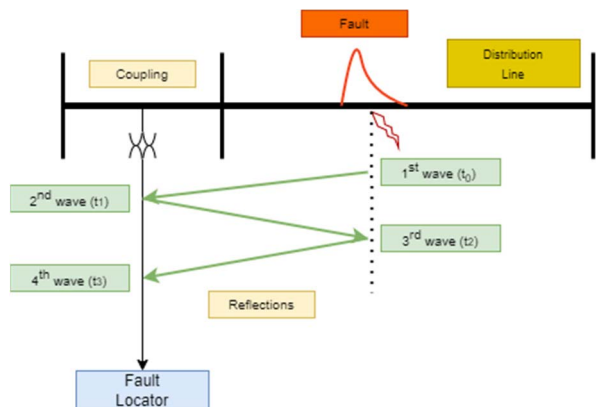
FIGURE 1. Fault location using the travelling wave of the spike event principle.

FIGURE 2. Pre-fault signal (a), Fault signal (b), comparison of both signals (c).

on the well-known telegrapher equations. One of the techniques within this group consists of measuring the bounce of a high-frequency signal that appears when a failure event occurs [47]. Short circuits caused by faults produces a high frequency transient that travels through the network (Fig 1).

This technique has provided good results in transportation lines where the use of two devices, one at each end of the section to be monitored, is widespread.

However, in distribution lines with many bifurcations, it is not possible to instal a device at each node because of the high cost.

Another difficulty lies in the time synchronisation that must exist between the equipment for the calculation of the distance and the prior knowledge of each section of the lines.

Other travelling wave methods are based on the injection of a signal and the analysis of the electrical response of the system. Following the telegrapher's equations, these signals undergo reflections at each location impedance change, as well as attenuations and distortions, until they return to the starting point. The occurrence of a fault implies a change in the characteristic impedance such that localisation is possible, at least in theory.

This technique is known as time-domain reflectometry (TDR). In distribution lines, the biggest challenge that TDR must overcome is resolving the multiple reflections caused by shunts, which complicates the analysis of the electrical response of the system. Thus, [48] it is proposed to inject through a healthy and faulted phase, and then perform a modal transformation to decouple the three-phase signals (using the Karrenbauer transformation matrix). The problem with this methodology is that it requires a healthy phase, which is not always available (three-phase faults), and that the injected pulse must have too high the amplitude (high voltage), which requires the network to be de-energised.

Techniques have also been proposed to measure the travel time of transients generated by the fault itself and by an injected signal. Later, the comparison of both allows the localisation of the fault to be calculated [49]. The method is accurate at the simulation level; however, it requires a high
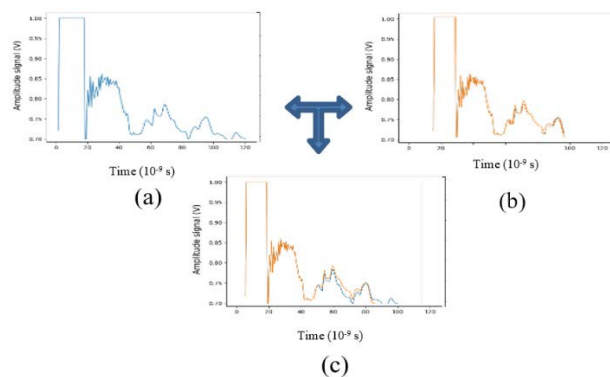
sampling rate (the use of the wavelet transform is proposed). In addition, it is not possible to locate all the faults because it is required that the occurrence of the fault temporally coincides with the high part of the sine wave cycle in order to have the detection capability.

Another proposed technique based on TDR consists of performing a periodic injection on the line in a pre-fault state so that the electrical response of the network is updated frequently. When a fault occurred, the injection was repeated, and the responses were compared (Fig. 2). Thus, an attempt can be made to locate the fault from the first point of divergence between the two signals [17], [47], [50], [51]. In the figure, it can be observed that the signals begin to diverge at approximately 50×10^{-9}s.

The pre-fault signal (line response to pulse injection when the fault has not yet occurred) and the fault signal (line response to pulse injection when the fault has already occurred) are almost the same until the pulse reaches the point where the fault is located. It is this phenomenon the one that is exploited in the TDR technique to locate the fault and, in our case, to detect and classify the fault.

Regardless of the type of fault, the pre-fault and fault signals begin to differ from the moment they reach the point where the fault has occurred

3) USE OF FAULT PASSAGE RELAYS TO DETECT THE FAULT

In medium-voltage lines (distribution lines/15 kV to 20 kV), there are devices in transformer substations called fault passage relays. These devices monitor the voltage and current in each phase and detect when a fault occurs in any phase. Once the fault passage relays warning that a fault has occurred, the locating system can operate with certainty that the line is indeed faulted. Protection systems are alerted, and they can open the section of the line where the fault is located until the problem is resolved (fault-cleared).

However, in low-voltage distribution lines (230 V ac), these devices usually do not exist, and as there are large numbers of bifurcations, the installation of many devices would be unfeasible.
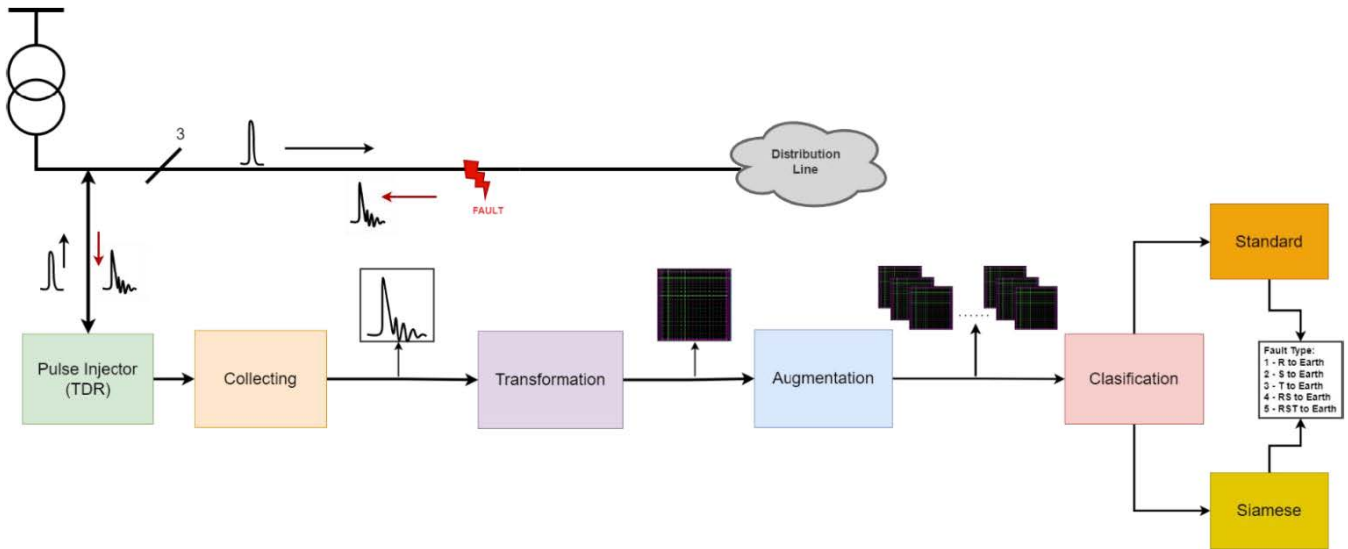
FIGURE 3. Methodology block diagram.

C. STRUCTURE OF THE PAPER

This paper aims to demonstrate the use of the pulse injection technique (TDR) for fault detection and classification as a preliminary step in fault localisation.

For this purpose, neural networks are used, which are trained to classify the signals obtained as a response of the electrical grids to the injected signals.

IV. METHODOLOGY

This section describes the proposed methodology based on the injection of signals into the distribution line and the analysis and classification of the received bounces. The methodology, including the pulse injection and subsequent transformations, is illustrated in Fig. 3. The following subsections explain the different steps of the methodology in detail.

To obtain a set of examples with which the neural network can be trained, the electrical grid of interest is modelled, as well as the injector with which the pulses are produced in each of the phases.

In this study, a real electrical grid was modelled using PSCAD^TM software. Faults at different locations as well as the responses to the injected signals were simulated. Despite being able to simulate a large number of examples (tens or hundreds), obtaining a large database to train a neural network typically requires tens of thousands of examples. A large number of examples allow the generalisation of the problem and prevent the network from learning examples (known as overfitting).

Based on the above, the contribution of the present paper is threefold:

1) To demonstrate that there is sufficient information embedded in the signals reflected by the network by injected pulses (TDR), and that it can be extracted to detect and classify them by type.
2) To demonstrate that with the applied methodology of GAN, it is possible to generate high-quality examples

of synthesized signals and solve the problem of the scarcity of original examples of this type of signal.
3) To propose a methodology for the union of GAN and Siamese networks (a type of neural network that is explained in detail below). With this methodology, the accuracy improves dramatically when training with only synthetic examples.

A. COLLECTING PHASE: TDR SIGNALS

The first process consists of creating a database with signals obtained using the software (PSCAD^TM). To do this, a real electrical network was modelled beforehand, and a series of five fault types was simulated at different points in the distribution line:

- Fault type 1: Short circuit between R to Earth.
- Fault type 2: Short circuit between S to Earth.
- Fault type 3: Short circuit between T to Earth.
- Fault type 4: Short circuit between R-S to Earth.
- Fault type 5: Short circuit between R-S-T to Earth.

The procedure followed, based on the time-domain reflectometry (TDR) technique, consists of injecting short-period pulses (∼10 ns) separated every few seconds (e.g. every 5 s). It should be noted that these pulses were injected into each of the three phases (R-S-T in a three-phase system).

According to the physics of transmission lines, these pulses travel through the network and are reflected at every bifurcation; therefore, some of them are bounced back. The magnitude of these reflections depended on the impedance of the line at each bifurcation.

As shown in Fig. 4, the pulses were injected into each phase with a time lag to allow the signal to be extinguished and therefore, to prevent the induction of the pulses into other phases.

Injection and reading of the response signal were performed at the same physical point on the line. In such a way that as soon as the pulse is injected, the response signal
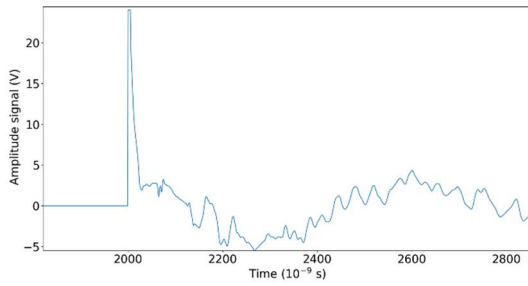
FIGURE 4. Detail of the injected pulse and the grid response during the very first moments in one phase.

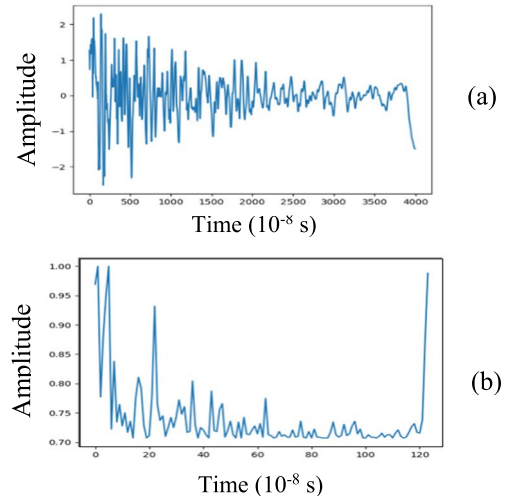**FIGURE 5.** Comparison of (a) an original time-series signal with (b) a 128-dimension GASF matrix recovered as time-series signal.

received by the line starts to be digitised. When the pulse encounters a point on the line at fault (short circuit), the impedance value at that point changes. The returning signal carries information about the impedance change at that point.

This technique is suitable for high-, medium-, and low-voltage distribution lines. However, its use is most relevant for low-voltage lines, as fault passage relays are not normally **installed** because of high installation costs.

In this study, the injected signals were sampled in a simulator at 100 Msps (Mega samples per second). For each of the three injected signals (R-S-T), approximately 40 μs were digitised, after which the signal was practically extinguished. Thus, the total sampling time for the three phases was 40 μs $\times$ 3 = 120 μs. At the indicated sampling rate, 12,000 digitised values (parameters) were obtained, which could be used as input to the neural network.

B. TRANSFORMATION PHASE: FROM TIME SIGNALS TO IMAGES

Currently, in the state of the art, there are several examples of time-series problems that are treated very satisfactorily by Convolutional Neural Networks through their prior transformation from time series to images [41], [42], [52].

Therefore, a standard technique is to convert a time series into images for better and easier data processing. For instance, in [52], the Gramian Angular field (GAF) transform was used to transform the time series into images.

In the GAF technique, we represent the time series in a polar co-ordinate system instead of typical Cartesian coordinates. Here, the amplitude of the signal is encoded as the angular cosine, and the timestamp is the radius, as shown in Equation (1). This information is gathered in the form of a matrix in which the relationships between different time instants can be identified. In this matrix, each element is the cosine of the summation of angles (GASF) (2) or the sine of the difference of the angles (GADF) (3).

$$\begin{cases} \varphi = arccos(\tilde{x}_i), & -1 \le \tilde{x}_i \le 1, \tilde{x}_i \in \tilde{X} \\ r = \dfrac{t_i}{N}, & t_i \in \mathbb{N} \end{cases} \tag{1}$$

$$GASF = \left[cos\left(\varphi_i + \varphi_j\right)\right] \tag{2}$$

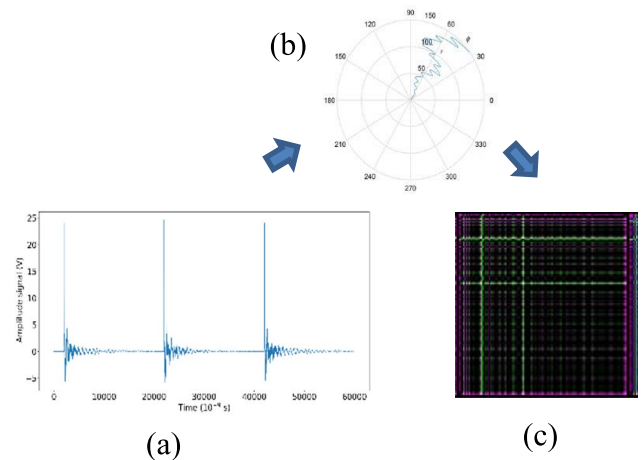$$GADF = \left[sin\left(\varphi_i - \varphi_j\right)\right] \tag{3}$$

FIGURE 6. Encoding of the fault temporal signal (a) by transform into a polar coordinate system (b) by (1) and finally calculate its GASF image (c) with (2) and (3). (Numbers in brackets refer to equations).

Applying this technique to each phase (R-S-T), we digitised 4,000 samples. In total, we have 12,000 samples (4,000 $\times$ 3), which leads to high dimensionality in the output matrix of the GAF transformation. To reduce this dimensionality, [17] an algorithm called *Piecewise Aggregate Approximation* (PAA) was proposed in [38]. First, we use the PAA algorithm to reduce the dimensionality to 128 (the choice of this parameter is justified below in the "Results" section). After applying PAA, GASF is applied to obtain a 128 $\times$ 128 $\times$ 3 channel matrix, where each channel corresponds to the signal of each of the three phases of the distribution line (R-S-T). As demonstrated in [52], the GASF transformation has an advantage over the GADF transformation in that the temporal signal can be reconstructed from the image (Fig. 5). This property will allow us to compare the original signals with the "reduced" ones (Fig. 6).

Once the signals have been adapted in such a way, the new format can be used in subsequent processing phases.
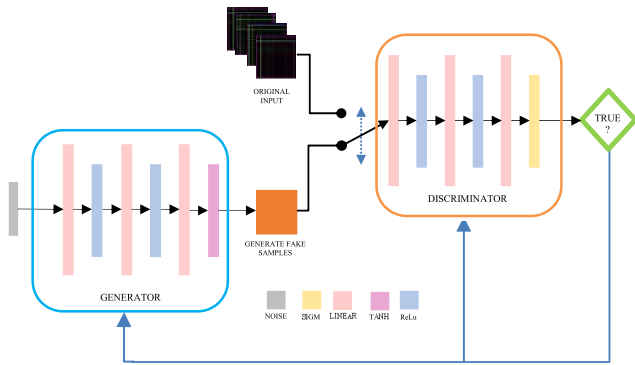
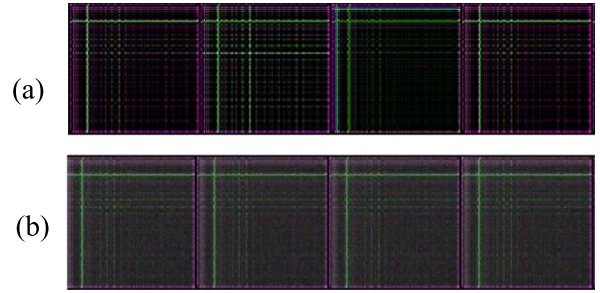**FIGURE 7.** GAN generator and discriminator layers structure.

FIGURE 8. Appearance of the original GASF signals (a) and synthesized by the GAN (b) once the GAN has been trained. It can be seen that they are very similar even to the naked eye.

TABLE 1. % examples used.

TRAINING	VALIDATION	TESTING
100%	100%	100%
SYNTHETICS	SYNTHETICS	ORIGINAL

C. DATASET AUGMENTATION PHASE: GAN

At this stage, our database was converted to a GASF image database. However, owing to the problem discussed above and the difficulty of obtaining more examples quickly and easily, the sample set had a limited size of 200 signals/images that included five types of faults. Thus, only 40 samples were available for each class.

Because our next phase will be automatic classification, a very small number of samples will be a real handicap, and it is well known that any learning algorithm highly depends on a large training dataset.

This issue is common in many cases where data extraction is extremely difficult, and in some cases, becomes impossible. As previously mentioned, simulating tens of thousands of examples is unfeasible because of the simulation time required. Another related problem is the need to manually adjust the parameters to generate diverse fault types. A wide range of cases can occur due to the stochastic nature of the phenomenon for each ToF. Therefore, we used GAN to generate synthetic examples, thereby increasing the initial database.

As we have seen, dimensionality reduction leads to a substantial reduction in training time and allows for easier implementation of GAN. We have tested different types of GAN: Convolutional GAN (CGAN) [53], Conditional GAN (cGAN) [54], even Conditional Deep Convolutional GAN (cDCGAN) [55]. We did not obtain satisfactory results due to the well-known problem of convergence of this type of Neural Networks known as "*mode collapse*." This problem causes the GAN to generate examples of a limited number of types.

Finally, we generated examples with a simpler GAN of linear layers, training each type separately with an independent GAN. By simplifying the GAN, we achieved convergence and obtained a database of 10,000 examples (2,000 examples for each ToF). Although we have achieved this at the cost of having a large number of parameters (104,306,000), we have prioritized the ease of training over the GAN optimization.

The GAN consists of a generator (D) and discriminator (D), as shown in Fig. 7. The generator has a random signal (noise) as input, which is used by the generator to produce new output images of the required dimension (*Generated sample*) ($128 \times 128 \times 3$).

On the other hand, the Discriminator is presented with original images (*Real samples*) and images generated by the Generator (*Generated samples*).

The Discriminator has to try to label them as "*True*" or "*False*". The successive training of the generator and discriminator ends once the discriminator is not able to distinguish the generated examples from the real ones. At this point, the generator can be used to produce as many examples as necessary (Fig. 8).

D. CLASSIFICATION PHASE: FAULT TYPES

Finally, at this stage, we have a database of "synthetic" GASF data (*Generated samples*) produced with the help of GAN, and a set of "original" data (*Real samples*) obtained from the simulation.

Usually, the training, validation, and testing data are split into a mixture of *Generated* and *Real samples* to achieve measurable results and improvements. For example, in, [44] the authors proposed using mixtures of no more than 50% of *Generated samples*.

However, we have a database of very few examples (200), and we cannot train the neural network with 200 original examples + 200 synthetic examples (400 total), as they are still too few examples to train a neural network.

To avoid this situation, we have proposed a different training solution, in which the distribution of training, validation, and testing data is made up as shown in Table 1. The training and validation sets are composed only of *Generated* data, and the test set is built only with *Real samples*.

As a first check, we verified the results obtained using the well-tested networks. From the simplest to the most complex, LeNet, AlexNet, and ResNet18 were used in this work. With these classifiers, the results were very poor, as expected and as previously reported.
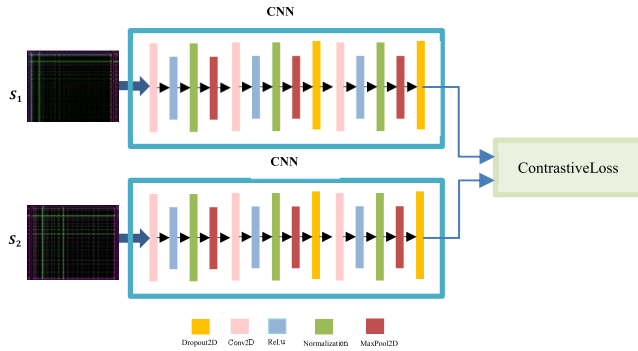
FIGURE 9. Architecture of siamese convolutional networks which are joined by the loss function in (4).

All these tests lead to the conclusion that the model cannot be generalized. The model only learns the distribution of the training examples (*Generated*) but fails to work with *Real samples*.

Then, as a suitable alternative to try to obtain better results, we aimed to verify whether we could use another approach based on neural networks that could be trained only with synthetic data and had a good performance when tested with original data. For this purpose, we used Siamese Convolutional Networks [13]

As shown in Fig. 9, the Siamese are twin nets that share the same weights. In this case, it can be seen that even if there are two convolutional networks, overall, we have half the parameters with respect to the GAN (46,854,530). The main feature is that they can be trained to learn a space in which the features of different classes are very close to each other. This is also our scenario, in which the different fault GASF images appear similar to each other. This was achieved by exposing the network to a pair of similar and dissimilar observations.

The network minimizes the Euclidean distance between similar pairs and maximizes the distance between dissimilar pairs (L = *contrastive loss*).

In (4), we can see that s_1 and s_2 are two samples (GASF images), y is a Boolean denoting whether the two samples belong to the same class, α and β are two constants, and m is the margin.

$$L(s_1, s_2, y) = \alpha(1-y)D^2w + \beta y max(0, m-D_w)^2 \qquad (4)$$

$$D_w = \|f(s_1; w_1) - f(s_2; w_2)\|_2 \qquad (5)$$

where (5) is the Euclidean distance computed in the embedded feature space. It can be seen that if y is 1 (different classes), the left-hand term disappears, and we attempt to maximize the distance between examples. On the other hand, if y is 0 (same class), the right-hand term disappears, and we attempt to minimize the distance between examples.

V. RESULTS

This section describes the methodology used in our experiments. An introduction is made to the configuration of the examples in the database, specifically how they have been selected for the different phases of the training (training,

validation, and testing). Then, the results obtained for each case are presented.

A. STRUCTURE OF THE EXPERIMENTS

To carry out the experiments in this work, we relied on a database of 200 examples of *Real samples* (Fig. 6 (a)), converted to images using the GASF transformation (Fig. 6 (b)), and 10.000 GASF *Generated samples* (2.000 for each of the five types under study).

In the first group of experiments, we focused on demonstrating the potential losses due to the GASF transformation and the subsequent suitability of GAN images. This will provide us with significant conclusions to determine whether these steps affect the quality of our results.

In the second group of experiments, where existing state-of-the-art neural networks (LeNet, AlexNet, and ResNet18) were trained, the database of synthetic examples was split into training (80%) and validation (20%). We then tested the database of original examples (200 examples, 40 examples of each ToF: $40 \times 5 = 200$).

Finally, when dealing with the Siamese network in the last experiment, a set of pairs randomly selected from the *Generated samples* were used to train the network. Two important facts should be mentioned at this point:1. This set was chosen to prevent imbalance among the types. 2. The validation and training phases were achieved with synthetic data (*Generated*), and tests were carried out with the original data (*Real*).

B. STUDY OF THE GASF TRANSFORMATION AND ITS RELATED LOSSES

As previously mentioned, the first step in processing is to transform the time signals into images using the GASF transform. In this step, it is necessary to decide which dimensionality reduction should be applied when performing this transformation. As explained in [56], this dimensionality reduction is obtained by applying *Piecewise Aggregate Approximation* (PAA).

Therefore, PAA is a source of potential loss and distortion. Therefore, a study in the frequency domain can demonstrate how this reduction distorted the original signals. Therefore, to proceed, we compute the FFT of the recovered signals already reduced according to the following procedure:

- A *Real sample* is initially transformed into GASF images of different dimensionalities (4000, 1000, 512, and 128, respectively).
- Each GASF *Generated sample* was back processed to obtain new versions of the same signal.
- Finally, we computed the FFT for the last signals (see Fig. 10).

From the study of these FFTs, it can be deduced that this dimensionality reduction is equivalent to a low-pass filter, as it is observed that as the reduction increases, the higher frequencies disappear and the lower frequencies are preserved.

For other types of studies, where the objective is to extract other types of more complex information (distance to the
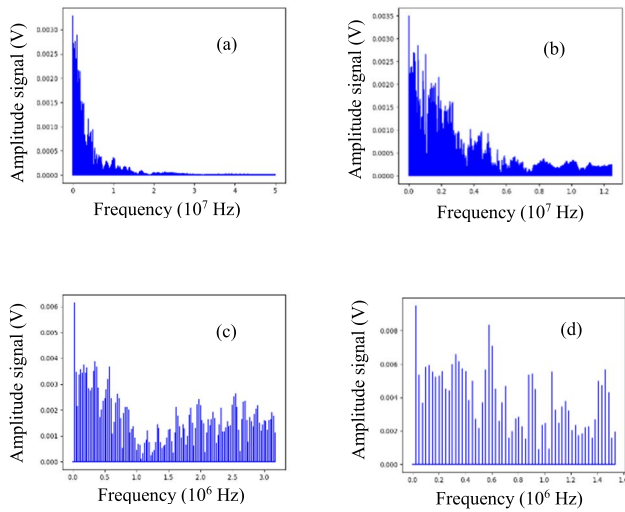
FIGURE 10. FFT of GASF images recovered of the same signal. (a) Dimensionality = 4,000 (Original) / (b) Dimensionality = 1,000 / (c) Dimensionality = 512 / (d) Dimensionality = 128. It can be observed that the higher frequencies disappear, while the lower frequencies are preserved when the dimensionality reduction increases.

fault, for instance), these high frequencies can become relevant. In our case and in view of the subsequent results, the assumed reduction (128×128) is adequate and allows us to save time in the training/validation phase.

C. ANALYSIS OF POTENTIAL MEASUREMENT ERRORS

Simulations based on real power distribution networks can lead to many difficulties in obtaining an adequate amount of data. This is why the central motif of this article is to obtain valuable fake data (generated samples) that can be used to train a neural network and obtain optimal results.

But at the same time, simulations can give us some uncertainty about the performance of the classifier in the face of possible failures of the measurement devices that can occur in a real environment. Indeed, this is a significant factor that distinguishes a simulated network from a real one. That is why this aspect has a specific section in this article.

In this sense, we are going to focus our interest on the study of three types of measurement errors:

1) RANDOM ERROR

Configured as a white noise signal, which has the next properties a) the signal is not statistically correlated between two different times, and b) its power spectral density (PSD) is constant throughout its spectrum. This type of error provides us with information on the behavior of the classifier over the entire frequency spectrum.

The results of the tests with this type of error are presented in Fig. 11, where errors of different signal-to-noise ratios have been used.

Several training epochs with different accuracy rates have been chosen, and each of them has been processed with noisy signals of 1 dB, 10 dB, 20 dB, 30 dB and 40 dB.
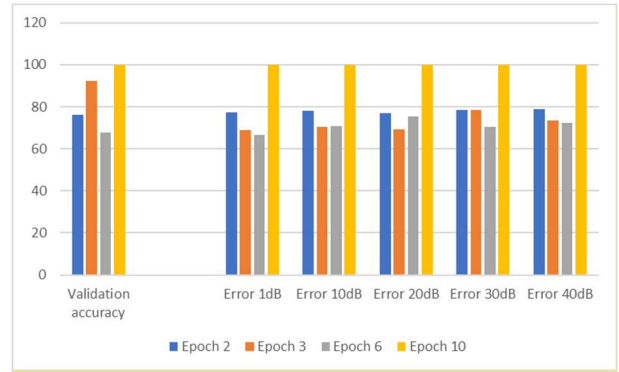

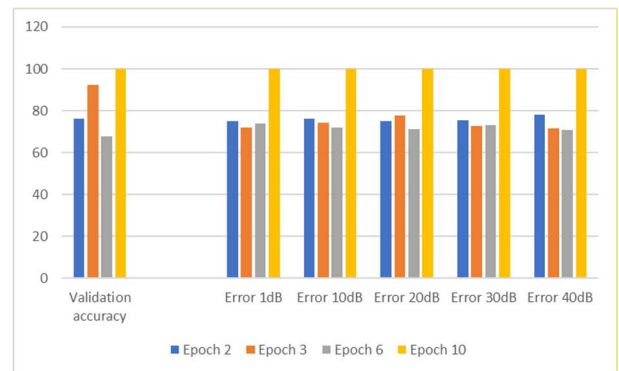**FIGURE 11.** Impact of measurement error (white noise) on results.

FIGURE 12. Impact of measurement error (bias) on results.

The conclusions that can be drawn from the figure are, on the one hand, the stability of the network with similar accuracy rates when the performance of the network in the corresponding epoch is partial or different from 100%. Except for epoch 3, where there is a generalized drop in network performance in the presence of noise. On the other hand, in epochs with a 100% hit rate, similar 100% accuracy rates are obtained from the network regardless of the level of noise.

2) BIAS ERROR

It provides an estimation of the accuracy of the measurement system and represents the systematic error that can occur in the system.

Fig. 12 shows the results with respect to this type of error. The experiment is set up identically to the random error case, and the results are also very similar. A drop in performance at epoch 2, and identical behavior when 100% accuracy is obtained in the validation can be observed.

3) AMPLIFICATION ERROR

In the latter case, we consider the error that may be introduced by the instrumentation, due to its limited bandwidth.

The classifier's tolerance to this type of error has a much simpler justification, as this would be a subset of the errors considered in the first two cases.

TABLE 2. Original vs synthetic data correlation.

Correlation	Type 1	Type 2	Type 3	Type 4	Type 5
max	0.93	0.97	0.95	0.95	0.93
min	0.53	0.72	0.54	0.35	0.41

TABLE 3. Original vs original data correlation.

Correlation	Type 1	Type 2	Type 3	Type 4	Type 5
max	1	1	1	1	1
min	0.31	0.61	0.38	0.24	0.46
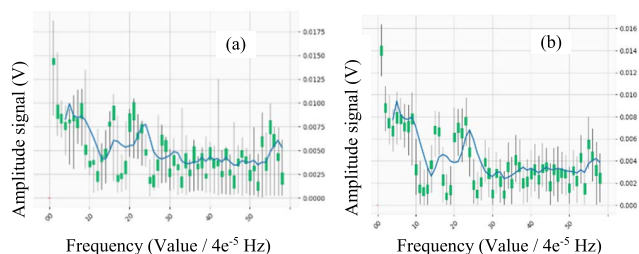
FIGURE 13. Original samples fault type 3 FFT statistics (a) and synthetics samples fault type 3 FFT statistics (b). Blue line is the moving average using 5 samples.
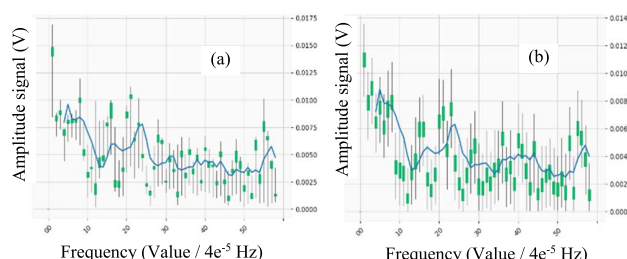
FIGURE 14. Original samples fault type 5 FFT statistics (a) and synthetics samples fault type 5 FFT statistics (b). Blue line is the moving average using 5 samples.

D. CHECKING THE SUITABILITY OF THE IMAGES OBTAINED WITH THE GAN

As mentioned above, we trained a GAN to synthesize examples and increase the database size after transforming the temporal signals into images. This would be useless if those images were not of sufficient quality and did not appear to be very similar. To do this, we select the *Real* signals and compare them with the *Generated* signals, converted back from GASF images, to check whether they are suitable. Therefore, a comparison will be made with signals and not with GASF images for a better understanding of the negative effects.

In Fig. 11 and 12, we can see the comparison between each harmonic of all the signals, both the *Real* and *Generated* signals for each ToF. In these figures, the amplitude variation of each harmonic is represented by the black lines. Similarly, the green flag represents the 50th percentile.

Looking at the graphs and moving average (blue line), it can be seen that both types of signals have a very similar frequency distribution pattern, which means that the GAN has consistently synthesized in terms of frequency spectrum vs. amplitudes.

Another aspect to highlight, in case of *Generated* signals, the percentile variation remains constant throughout the frequency spectrum. Meanwhile, the *Real* ones show different variabilities in each of the harmonics.

Table 2 lists the maximum and minimum values shown in Figure 13. It can be stated that, in general, both signals are strongly correlated.

However, there are some *Real* signals that show a low correlation (blue color) with the rest of the Generated signals.

Fig. 15 (a) shows them with a frequency distribution pattern that is different from the rest. In these cases, the GAN has not learned to generalize this type of signal, which may be a limitation of GAN. Moreover, the rest of the examples with similar patterns were well synthesized by GAN and showed strong correlations.

Finally, the reader can appreciate the degree of correlation between the *Real* signals (see Fig. 14 and Table 3). The signals have a different pattern, since they are originally poorly correlated with all the others.

E. TRAINING STANDARD NEURAL NETWORKS MODELS

Our first attempt to classify was to train the standard models LeNet [57], AlexNet [58], and ResNet18 [59] with our new extended database to verify the results. Fig. 16 shows the progression of the complete training, validation, and testing processes for the AlexNet model. We continue applying the same strategy to ensure a fair comparison (training and validation phase using newly *Generated* data and testing results with the *Real* dataset).

The model can obtain values close to 100% after a few epochs, both in training and validation (with *Generated* data). However, in the test stage, with *Real* examples, the accuracy is poor, which means that the model cannot be generalized well. This is in line with what exists in the literature regarding the need to train using a limited number of syn-
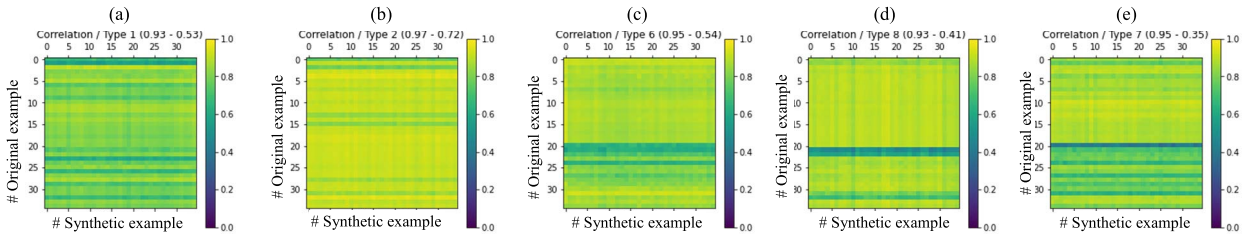
FIGURE 15. FFT correlation synthetic vs original examples for each ToF: type 1 (a), type 2 (b), type 3 (c), type 4 (d), type 5 (e).
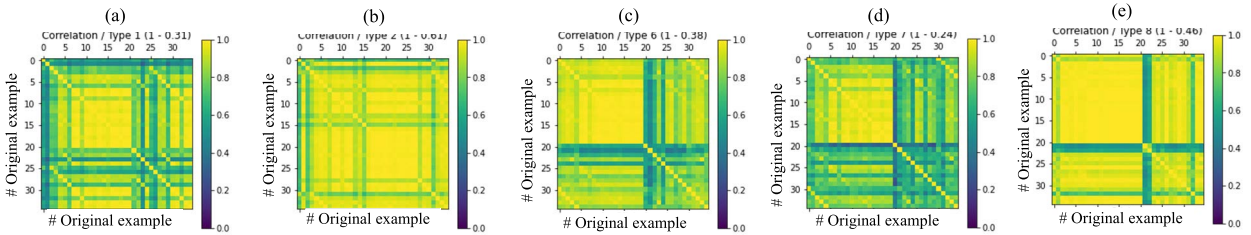

FIGURE 16. FFT correlation original vs original examples for each ToF: type 1 (a), type 2 (b), type 3 (c), type 4 (d), type 5 (e).
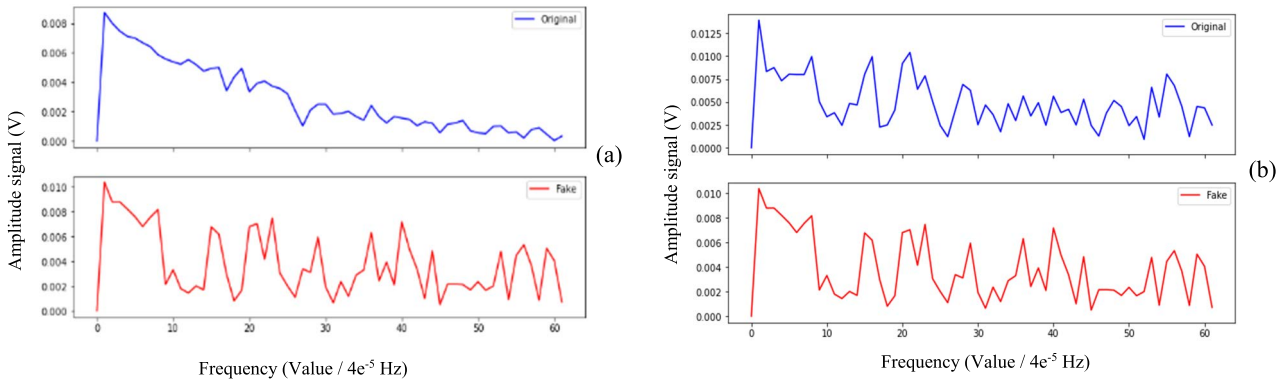

FIGURE 17. Comparison of FFT between original and synthetic signals: Comparison between original signal number 22 (blue) and synthetic signal number 34 (red), both belonging to fault type 5 (a). Comparison between the original signal number 10 (blue) and synthetic signal number 34 (red), both belonging to fault type 5 (b).

TABLE 4. Accuracy (%).

	TRAINING	VALIDATION	TESTING
LeNet	99.98%	100%	27.36%
AlexNet	99.98%	100%	33.83%
ResNet18	99.98%	100%	35.82%

thetic examples [44]. Table 4 lists the training results of the LeNet, AlexNet, and ResNet18 models. Similar results were obtained for the three models.

F. TRAINING THE SIAMESE NEURAL NETWORK

Siamese networks are well known for their effectiveness in *one-shot* or *few-shots* learning strategies [28], being able to adapt to new distributions quickly.

Applying the same learning strategy as in the previous sections, in Fig. 17 we can observe the evolution of the training loss for each epoch, as well as the loss and accuracy in training and validation phase for different learning rates. For smaller values of the latter parameter, less aggressive curves are achieved.
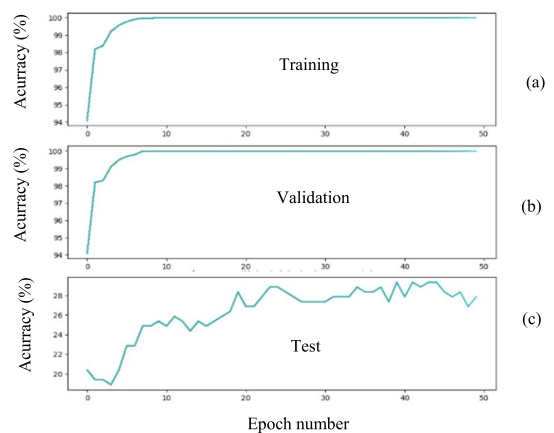

FIGURE 18. Accuracy of the training phase (a) with known synthetic data, validation phase (b) with unknown synthetic data, and test phase (c) with unknown original data of the AlexNet model. The training and validation phases were performed using 100% synthetic examples, and the test phase was performed using 100% original examples.

We would also like to highlight the performance with the test data. The Siamese network (SN) is able to generalize between both distributions (synthetic and original), and the
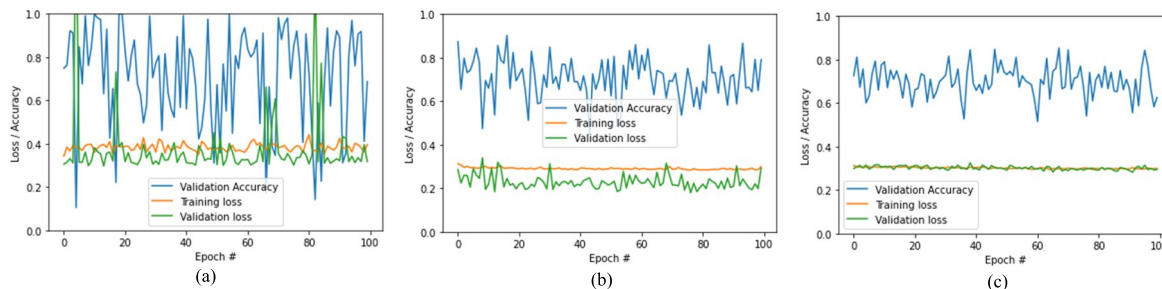
FIGURE 19. Validation accuracy and Training / Validation Loss phases for different learning rates, 1e-4 (a), 1e-5 (b), 1e-6 (c), and for each Epoch.
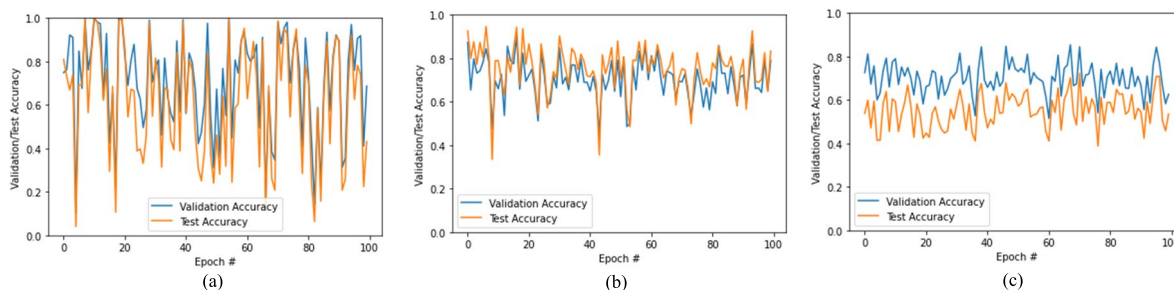

FIGURE 20. Validation and Test accuracy phases for different learning rates, 1e-4 (a), 1e-5 (b), 1e-6 (c), and for each Epoch.

results are even better if the correct learning rate is finally chosen (in this case, $1e^{-5}$). In Fig. 18, we can see this comparison between the validation and test accuracies for each epoch as a function of the learning rate.

In conclusion, it can be said that the network converges to validation accuracy values of approximately 70% and test accuracy values of approximately 80% for the correct learning rate. This represents an improvement of more than two times with respect to the standard networks trained in Section D. This result shows that contrastive learning helped us to effectively classify original signals. In this way, the SN has been able to classify original examples that it has never seen before, but also belongs to a different distribution.

Apart from this result, it can be seen that with high learning rates ($1e^{-4}$) as shown in Fig.17 (a), it is worth noting that even with this high variability, and owing to the ability of Siamese networks, it is able to return 100% accuracy rates both in the validation and test phases.

As mentioned before, this behavior can be explained because this type of network can be used for *one-shot-learning*, so the network does not require much training.

Thus, the model can correctly classify never-seen examples belonging to another distribution. With a sufficient number of synthetic examples, it took a few epochs to correctly learn to separate and classify unseen originals.

VI. CONCLUSION

We propose a solution to the problem of fault detection and classification with a few examples: the generation of synthetic examples by means of GAN networks, com-

bined with the application of contrastive loss (Siamese networks).

In relation to the expected objectives set in this study:

1. It has been demonstrated that fault-type information is present in the signals reflected by the network as a result of injected pulses (TDR). Very good results were achieved by performing dimensionality reduction by applying the *Piecewise Aggregate Approximation* (PAA) technique. This reveals that the information of the ToF is in the low frequencies of the signal and can be dispensed with, at least in this study, higher frequencies.

2. It has been possible synthesized examples of each of the five types of faults used in this study were generated. We went from having a database of 200 examples (original examples) to a database of 10,000 synthesized examples (2,000 of each type).

3. It has been possible to train a network (Siamese with *Contrastive Loss* output) only with synthesized examples and test it with the original examples, yielding very high accura**cy**.

These synthesized examples have been divided into two groups, one for training and one for validation.

These two groups will belong to the same data distribution, so the generalization could not be tested yet. To certify that the system does not only generalize to the training distribution a test has been performed on the original 200 data, which belong to a different distribution.

This new method facilitates the training of neural networks with this type of signal, for which few examples are available owing to their nature.

VII. FUTURE WORK

Our work is directly linked to a research institution that is strongly committed to TDR technology, which leads us to continue advancing in this field. First, we are trying to understand the limits to which we can take neural networks in the understanding of this type of phenomenon; second, we are excited by the results obtained, and we are encouraged to discover new particularities that can help solve different problems in this area.

However, our short-term points of interest are as follows:

1. To explore whether our method is valid also for high impedances and for simultaneous fault detection.

2. To explore whether our method may be affected by the presence of distributed generations in the distribution networks

3. To explore whether the neural network can generalize the learned fault types to other real distribution lines, without retraining or at least with a minimal learning phase. The versatility of the process must be such that it is functional in diverse environment

4. To explore the possibility of extracting more complex information, such as fault location, by adapting the technique described in this paper. The challenge here is to move from a classification problem to a linear regression problem by generating synthetic examples.

5. Finally, we believe that we have kept open the door that others have opened with one-short learning methods. This philosophy has a huge potential to be pursued, and new contributions can generate the common know-how that the research community needs.

REFERENCES

[1] D. P. Mishra and P. Ray, "Fault detection, location and classification of a transmission line," *Neural Comput. Appl.*, vol. 30, no. 5, pp. 1377–1424, Sep. 2018, doi: 10.1007/s00521-017-3295-y.

[2] M. Mirzaei, M. Z. A. A. Kadir, E. Moazami, and H. Hizam, "Review of fault location methods for distribution power system," *Austral. J. Basic Appl. Sci.*, vol. 3, no. 3, pp. 2670–2676, 2009.

[3] R. J. Hamidi and H. Livani, "Traveling-wave-based fault-location algorithm for hybrid multiterminal circuits," *IEEE Trans. Power Del.*, vol. 32, no. 1, pp. 135–144, Feb. 2017, doi: 10.1109/TPWRD.2016.2589265.

[4] S. Wei, G. Yanfeng, and L. Yan, "Traveling-wave-based fault location algorithm for star-connected hybrid multi-terminal HVDC system," in *Proc. IEEE Conf. Energy Internet Energy Syst. Integr.*, Nov. 2017, pp. 1–5, doi: 10.1109/EI2.2017.8245645.

[5] J. Teng, X. Zhan, L. Xie, X. Zeng, Y. Liu, and L. Huang, "A novel location method for distribution hybrid lines," in *Proc. IEEE Conf. Energy Internet Energy Syst. Integr. (EI2)*, Nov. 2017, pp. 1–5, doi: 10.1109/EI2.2017.8245425.

[6] L. Wang, "The fault causes of overhead lines in distribution network," in *Proc. MATEC Web Conf.*, vol. 61, Jun. 2016, p. 02017, doi: 10.1051/MATECCONF/20166102017.

[7] S. Das, S. Santoso, and S. N. Ananthan, "Fault location on transmission and distribution lines," in *Fault Location on Transmission and Distribution Lines*, Jan. 2022, doi: 10.1002/9781119121480.

[8] S.-C. Huang and T.-H. Le, "Training neural network," in *Principles and Labs for Deep Learning*. New York, NY, USA: Academic, Jun. 2021, pp. 117–146, doi: 10.1016/B978-0-323-90198-7.00007-0.

[9] O. A. M. López, A. M. López, and J. Crossa, "Overfitting, model tuning, and evaluation of prediction performance," in *Multivariate Statistical Machine Learning Methods for Genomic Prediction*. Cham, Switzerland: Springer, 2022, pp. 109–139, doi: 10.1007/978-3-030-89010-0.

[10] K. Berns, A. Köpper, and B. Schürmann, "Electrical networks," in *Technical Foundations of Embedded Systems* (Lecture Notes in Electrical Engineering), vol. 732, pp. 19–43, 2021, doi: 10.1007/978-3-030-65157-2_3.

[11] D. Zhang, M. Ma, and L. Xia, "A comprehensive review on GANs for time-series signals," *Neural Comput. Appl.*, vol. 34, no. 5, pp. 3551–3571, Mar. 2022, doi: 10.1007/S00521-022-06888-0.

[12] G. Koch, R. Zemel, and R. Salakhutdinov, "Siamese neural networks for one-shot image recognition," in *Proc. ICML Deep Learn. Workshop*, 2015, pp. 1–30.

[13] S. Dey, A. Dutta, J. Ignacio Toledo, S. K. Ghosh, J. Llados, and U. Pal, "SigNet: Convolutional Siamese network for writer independent offline signature verification," 2017, *arXiv:1707.02131*.

[14] H. Ungrad, W. Winkler, and A. Wiszniewski, "The main criteria for detecting faults," in *Protection Techniques in Electrical Energy Systems*. Boca Raton, FL, USA: CRC Press, Aug. 2020, pp. 28–38, doi: 10.1201/9781003067504.

[15] J. Jiao, G. Lai, L. Zhao, J. Lu, Q. Li, X. Xu, Y. Jiang, Y. He, C. Ouyang, F. Pan, H. Li, and J. Zheng, "Self-healing mechanism of lithium in lithium metal," *Adv. Sci.*, vol. 9, no. 12, Apr. 2022, Art. no. 2105574, doi: 10.1002/ADVS.202105574.

[16] M. Brehm, L. Trigo, and D. Slomovitz, "Semiautomatic testing system for high voltage detectors," in *Proc. IEEE 9th Power, Instrum. Meas. Meeting (EPIM)*, Nov. 2018, pp. 1–3, doi: 10.1109/EPIM.2018.8756432.

[17] J. Oasa, M. Yamanaka, S. Higashiyama, Y. Inaoka, T. Hisakado, O. Wada, T. Matsushima, T. Hirayama, and K. Yamaoka, "Verification of fault location by TDR measurement on an actual line including multiple ground-mounted equipment," in *Proc. 26th Int. Conf. Exhib. Electr. Distrib.*, Sep. 2021, pp. 1274–1278.

[18] R. Lacoste, "The darker side: Time domain reflectometry," in *Circuit Cellar*, no. 225. Apr. 2009, ch. 3, pp. 33–47, doi: 10.1016/B978-1-85617-762-7.00003-4.

[19] A. Prasad, J. B. Edward, and K. Ravi, "A review on fault classification methodologies in power transmission systems: Part—I," *J. Electr. Syst. Inf. Technol.*, vol. 5, no. 1, pp. 48–60, May 2018, doi: 10.1016/J.JESIT.2017.01.004.

[20] A. Prasad, J. Belwin Edward, and K. Ravi, "A review on fault classification methodologies in power transmission systems: Part—II," *J. Electr. Syst. Inf. Technol.*, vol. 5, no. 1, pp. 61–67, May 2018, doi: 10.1016/J.JESIT.2016.10.003.

[21] B. Nayak. (2014). *Impact Factor: 1.852 Classification of Transmission Line Faults Using Wavelet Transformer*. Accessed: Jun. 1, 2022. [Online]. Available: https://www.ijesrt.com

[22] S. Kesharwani, D. Kumar, and M. Scholar, "Detection of power quality disturbances using wavelet transform," *Int. J. Sci., Eng. Technol. Res.*, vol. 3, no. 5, pp. 1–6, 2014. Accessed: Jun. 2, 2022. [Online]. Available: https://www.researchgate.net/publication/301695115

[23] S. R. Samantaray, "A systematic fuzzy rule based approach for fault classification in transmission lines," *Appl. Soft Comput.*, vol. 13, no. 2, pp. 928–938, Feb. 2013, doi: 10.1016/J.ASOC.2012.09.010.

[24] A. Aljohani, A. Aljurbua, M. Shafiullah, and M. A. Abido, "Smart fault detection and classification for distribution grid hybridizing ST and MLP-NN," in *Proc. 15th Int. Multi-Conf. Syst., Signals Devices (SSD)*, Mar. 2018, pp. 94–98, doi: 10.1109/SSD.2018.8570582.

[25] M. Shafiullah and M. A. Abido, "S-transform based FFNN approach for distribution grids fault detection and classification," *IEEE Access*, vol. 6, pp. 8080–8088, 2018, doi: 10.1109/ACCESS.2018.2809045.

[26] W. J. Cheong, "A novel fault location technique based on current signals only for thyristor controlled series compensated transmission lines using wavelet analysis and self organising map neural networks," in *Proc. 8th IEE Int. Conf. Develop. Power Syst. Protection*, 2004, pp. 224–227, doi: 10.1049/CP:20040104.

[27] A. Ngaopitakkul, C. Apisit, S. Bunjongjit, and C. Pothisarn, "Identifying types of simultaneous fault in transmission line using discrete wavelet transform and fuzzy logic algorithm," *Int. J. Innov. Comput. Inf. Control*, vol. 9, no. 7, pp. 2701–2712, 2013.

[28] N. Hooda and M. Malik, "Review on neuro-fuzzy system," *SSRN Electron. J.*, vol. 2022, pp. 1–11, Mar. 2022, doi: 10.2139/SSRN.4033495.

[29] E. S. M. T. Eldin, "Fault location for a series compensated transmission line based on wavelet transform and an adaptive neuro-fuzzy inference system," in *Proc. Electr. Power Quality Supply Rel. Conf.*, Jun. 2010, pp. 229–235, doi: 10.1109/PQ.2010.5549994.

[30] M. Singh, B. K. Panigrahi, and R. P. Maheshwari, "Transmission line fault detection and classification," in *Proc. Int. Conf. Emerg. Trends Electr. Comput. Technol.*, Mar. 2011, pp. 15–22, doi: 10.1109/ICETECT.2011.5760084.

[31] J. Upendar, C. P. Gupta, and G. K. Singh, "Discrete wavelet transform and probabilistic neural network based algorithm for classification of fault on transmission systems," in *Proc. IEEE Conf. Exhib. Control, Commun. Autom.*, vol. 1, Dec. 2008, pp. 206–211, doi: 10.1109/INDCON.2008.4768827.

[32] C. D. Prasad and D. J. V. Prasad, "Fault detection and phase selection using Euclidean distance based function for transmission line protection," in *Proc. Int. Conf. Adv. Electr. Eng. (ICAEE)*, Jan. 2014, pp. 1–4, doi: 10.1109/ICAEE.2014.6838516.

[33] P. Sharma, D. Saini, and A. Saxena, "Fault detection and classification in transmission line using wavelet transform and ANN," *Bull. Electr. Eng. Informat.*, vol. 5, no. 3, pp. 284–295, Sep. 2016, doi: 10.11591/eei.v5i3.537.

[34] D. Mnyanghwalo, H. Kundaeli, E. Kalinga, and N. Hamisi, "Deep learning approaches for fault detection and classifications in the electrical secondary distribution network: Methods comparison and recurrent neural network accuracy comparison," *Cogent Eng.*, vol. 7, no. 1, Jan. 2020, Art. no. 1857500, doi: 10.1080/23311916.2020.1857500.

[35] J. J. Q. Yu, Y. Hou, A. Y. S. Lam, and V. O. K. Li, "Intelligent fault detection scheme for microgrids with wavelet-based deep neural networks," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 1694–1703, Mar. 2019, doi: 10.1109/TSG.2017.2776310.

[36] D. Baskar and P. Selvam, "Machine learning framework for power system fault detection and classification," *Int. J. Sci. Technol. Res.*, vol. 9, p. 2, Jan. 2020. [Online]. Available: https://www.ijstr.org

[37] F. Rudin, G.-J. Li, and K. Wang, "An algorithm for power system fault analysis based on convolutional deep learning neural networks," *Int. J. Res. Educ. Sci. Methods.*, vol. 5, no. 9, pp. 11–18, 2017.

[38] Y. Wu, P. Zhang, and G. Lu, "Detection and location of aged cable segment in underground power distribution system using deep learning approach," *IEEE Trans. Ind. Informat.*, vol. 17, no. 11, pp. 7379–7389, Nov. 2021, doi: 10.1109/TII.2021.3056993.

[39] N. Sapountzoglou, J. Lago, B. De Schutter, and B. Raison, "A generalizable and sensor-independent deep learning method for fault detection and location in low-voltage distribution grids," *Appl. Energy*, vol. 276, Oct. 2020, Art. no. 115299, doi: 10.1016/j.apenergy.2020.115299.

[40] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces," *J. Neural Eng.*, vol. 15, no. 5, Oct. 2018, Art. no. 056013, doi: 10.1088/1741-2552/aace8c.

[41] S. Barrios, D. Buldain, M. P. Comech, I. Gilbert, and I. Orue, "Partial discharge classification using deep learning methods—Survey of recent progress," *Energies*, vol. 12, no. 13, p. 2485, Jun. 2019, doi: 10.3390/en12132485.

[42] Z. Wang and T. Oates, "Encoding time series as images for visual inspection and classification using tiled convolutional neural networks," in *Proc. Workshops 29th AAAI Conf. Artif. Intell.*, 2015, pp. 1–7. [Online]. Available: https://www.aaai.org

[43] K. Yang, W. Gao, R. Fan, T. Yin, and J. Lian, "Synthetic high impedance fault data through deep convolutional generated adversarial network," in *Proc. IEEE Green Technol. Conf. (GreenTech)*, Apr. 2021, pp. 339–343, doi: 10.1109/GREENTECH48523.2021.00061.

[44] X. Wang, H. Huang, Y. Hu, and Y. Yang, "Partial discharge pattern recognition with data augmentation based on generative adversarial networks," in *Proc. Condition Monitor. Diagnosis (CMD)*, Sep. 2018, pp. 1–4, doi: 10.1109/CMD.2018.8535718.

[45] R. Dashti, S. Salehizadeh, H. Shaker, and M. Tahavori, "Fault location in double circuit medium power distribution networks using an impedance-based method," *Appl. Sci.*, vol. 8, no. 7, p. 1034, Jun. 2018, doi: 10.3390/app8071034.

[46] L. Peretto, R. Tinarelli, A. Bauer, and S. Pugliese. (2011). *Fault Location in Underground Power Networks: A Case Study*. Accessed: Mar. 30, 2022. [Online]. Available: https://ieeexplore-ieee-org.cuarzo.unizar.es:9443/document/5759198/

[47] N. E. Halabi, "Localizadores de faltas para redes de distribución eléctrica," Fundación CIRCE, Universidad de Zaragoza, Zaragoza, Spain, Tech. Rep., 2012.

[48] C.-C. Zhou, Q. Shu, and X.-Y. Han, "A single-phase earth fault location scheme for distribution feeder on the basis of the difference of zero mode traveling waves," *Int. Trans. Electr. Energy Syst.*, vol. 27, no. 5, p. e2298, May 2017, doi: 10.1002/ETEP.2298.

[49] W. Chonglin, W. Yangyang, L. Rui, and S. Gang. (2010). *Fault Location for Single-Phase-to-Earth Faults Based on Transient Traveling Wave Method and Artificial Pulse Signal Injection Method*. Accessed: Mar. 30, 2022. [Online]. Available: https://ieeexplore-ieee-org.cuarzo.unizar.es:9443/document/5629889/

[50] M. Abad, U. Zaragoza, N. El, H. Saudi, A.-S. Arabia, and M. G.-G. Circe-Spain, "New fault location method for up-todate and upcoming distribution networks," in *Proc. 23rd Int. Conf. Electr. Distrib.*, 2015, pp. 15–18.

[51] J. Ballestín-Fuertes, D. Cervero, H. Bludszuweit, R. Martínez, and J. A. S. Castro, "Fault location in low-voltage distribution networks based on reflectometry—A case study," *Renew. Energy Power Quality J.*, vol. 18, pp. 735–740, Jun. 2020, doi: 10.24084/REPQJ18.483.

[52] Z. Wang and T. Oates, "Imaging time-series to improve classification and imputation," 2015, *arXiv:1506.00327*.

[53] A. Patil and Venkatesh, "DCGAN: Deep convolutional GAN with attention module for remote view classification," in *Proc. Int. Conf. Forensics, Anal., Big Data, Secur. (FABS)*, Dec. 2021, pp. 1–10, doi: 10.1109/FABS52071.2021.9702655.

[54] M. Li, J. Lin, Y. Ding, Z. Liu, J.-Y. Zhu, and S. Han, "GAN compression: Efficient architectures for interactive conditional GANs," *IEEE Trans. Pattern Anal. Mach. Intell.*, early access, Nov. 9, 2021, doi: 10.1109/TPAMI.2021.3126742.

[55] F. Zhu, M. He, and Z. Zheng, "Data augmentation using improved cDCGAN for plant vigor rating," *Comput. Electron. Agricult.*, vol. 175, Aug. 2020, Art. no. 105603, doi: 10.1016/J.COMPAG.2020.105603.

[56] E. J. Keogh and M. J. Pazzani, "Scaling up dynamic time warping for datamining applications," in *Proc. 6th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2000, pp. 285–289.

[57] V. Verdhan, "Image classification using LeNet," in *Computer Vision Using Deep Learning*, 2021, pp. 67–101, doi: 10.1007/978-1-4842-6616-8_3.

[58] V. Verdhan, "VGGNet and AlexNet networks," in *Computer Vision Using Deep Learning*. Irvine, CA, USA: Univ. of California, Irvine, 2021, pp. 103–139, doi: 10.1007/978-1-4842-6616-8_4.

[59] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

JAVIER GRANADO FORNÁS received the B.Sc. degree in industrial engineering (specialized in industrial electronics) and the M.Sc. degree in electronics engineering (intelligent environments specialization) from the University of Zaragoza, Spain, in 1994 and 2014, respectively. He is currently pursuing the Ph.D. degree in deep learning around classification and localization of faults in distributions lines. Since 2009, he has been working as a Senior Researcher with the Electronics Systems Group, CIRCE Technology Center. His main research interests include electronic designs control projects around deep learning and algorithms for fault detection.

ELÍAS HERRERO JARABA (Member, IEEE) received the Ph.D. degree in engineering from the University of Zaragoza.

In 1999, he worked in the automotive field at the Production Department, Opel Spain, for four years. Since June 2002, he has been a Professor at the University of Zaragoza. In the meantime, he was a Coordinator of the smart vehicle initiative at the Aragon Institute for Engineering Research (I3A). He currently holds a European patent and has led to more than six projects funded in public competitions, more than 20 projects with companies, 12 indexed publications, and more than 50 contributions to international conferences. It is worth mentioning the 3 million awards for his research in the CvLAB Research Group. His teaching work has focused on electronics, from its basics to power electronics, including embedded systems. Within the university, he has developed his research in the area of computer vision. Since 2012, he has focused his research interests in the field of neural networks, and more recently, in deep learning.

ANDRÉS LLOMBART ESTOPIÑAN (Member, IEEE) received the Industrial Engineering degree from the University of Zaragoza, in 1994.

From December 1994 to May 2001, he was an Associate Professor with the Department of Electrical Engineering, University of Zaragoza, where he performed the following teaching activities, such as electric circuit theory, industrial actuation, wind energy, and renewable energy integration. He was also a Lecturer at the Department of Electrical Engineering, University of Zaragoza, from June 2003 to May 2018. From March 2007 to March 2009, he was the Sub-Director of Institutional Relations at the Superior Polytechnic Centre, University of Zaragoza. In CIRCE Technology Center, he was in charge of Innovation and Promotion Unit, created by himself on May 2009. He has been the General Director of the CIRCE Technology Center, since April 2016, and the Former Executive Director, since January 2011. In November 2011, he was designated by the Science and Innovation Ministry as an Expert in Energy Area Committee of the 7th Funding Program of the European Union, being in charge of the coordination of electricity grids topics. He participated in more than 35 research and development +i projects; in 16 of them, he was the primary researcher. He has author of ten articles in indexed journals and more than 50 contributions in international congresses. His active participation in forums, associations, and platforms linked to activity lines. He has eight patents, of which seven are being exploited. He researched the impact reduction of power electronic source grids using passive filters in the kilowatt range. His research gave him a Ph.D. title (specialized in electrical engineering) with the University of Zaragoza, in 2000.

JOSE SALDANA (Senior Member, IEEE) was born in San Sebastián, Spain, in 1974. He received the B.S. and M.S. degrees in telecommunications engineering and the Ph.D. degree in information technologies from the University of Zaragoza, Spain, in 1998, 2008, and 2011, respectively. He is currently a Senior Researcher at the CIRCE Technology Center. He has participated in research projects related to the digitalization of power systems, including ICT performance and security, and the use of wireless communications in industrial environments. He has published over 60 articles in peer-reviewed journals and conferences and RFC7962 in the Internet Engineering Task Force (IETF) on alternative network deployments. His research interests include wired and wireless networks with tight delay constraints, including multimedia services and digital communications in electric substations. He serves on the Editorial Board for IEEE Access (Associate Editor) and the *KSII Transactions on Internet and Information Systems* (Area Editor). He has also served on the Steering Committee and TPC for many conferences, such as the IEEE Consumer Communications and Networking, the ACM Multimedia Systems Conference, IEEE ICC, and the IEEE Globecom.

● ● ●