informatics

MDPI

Article

Markov-CVAELabeller: A Deep Learning Approach for the Labelling of Fault Data

Christian Velasco-Gallego * and Nieves Cubo-Mateo

Grupo de Investigación ARIES, Universidad Nebrija, Calle de Santa Cruz de Marcenado, 27, 28015 Madrid, Spain; ncubo@nebrija.es

* Correspondence: cvelasco@nebrija.es

Abstract: The lack of fault data is still a major concern in the area of smart maintenance, as these data are required to perform an adequate diagnostics and prognostics of the system. In some instances, fault data are adequately collected, even though the fault labels are missing. Accordingly, the development of methodologies that generate these missing fault labels is required. In this study, Markov-CVAELabeller is introduced in an attempt to address the lack of fault label challenge. Markov-CVAELabeller comprises three main phases: (1) image encoding through the application of the first-order Markov chain, (2) latent space representation through the consideration of a convolutional variational autoencoder (CVAE), and (3) clustering analysis through the implementation of *k*-means. Additionally, to evaluate the accuracy of the method, a convolutional neural network (CNN) is considered as part of the fault classification task. A case study is also presented to highlight the performance of the method. Specifically, a hydraulic test rig is considered to assess its condition as part of the fault diagnosis framework. Results indicate the promising applications that this type of methods can facilitate, as the average accuracy presented in this study was 97%.

Keywords: fault diagnosis; image encoding; time series analysis; deep learning; variational autoencoder; convolutional neural network; first-order Markov chain; artificial intelligence; health management

check for updates

Academic Editor: Long Jin

Received: 14 February 2025
Revised: 16 March 2025
Accepted: 21 March 2025
Published: 25 March 2025

Citation: Velasco-Gallego, C.; Cubo-Mateo, N. Markov-CVAELabeller: A Deep Learning Approach for the Labelling of Fault Data. *Informatics* **2025**, *12*, 35. https://doi.org/10.3390/informatics12020035

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Fault diagnosis can be defined as a process that aims to determine not only whether a fault occurs in a system but also when, where, and what kind of fault occurs and to what extent it occurs [1]. This process typically consists of detection, isolation, and identification of faults. Fault detection focuses on recognising the occurrence of a fault [2], isolation aims to locate the detected fault [3], and identification categorises the detected fault based on its characteristics [4]. With advancements in artificial intelligence, intelligent fault diagnosis has gained significant attention in recent studies due to its capabilities for enhancing decision-making processes in complex systems.

Intelligent fault diagnosis aims to apply machine learning techniques to trace a fault by means of its symptoms [5,6]. When performing fault classification, supervised learning is usually conducted. This type of learning establishes a mapping function for associated input samples into their corresponding target vectors by utilising labelled data [7]. However, a significant challenge when performing fault classification is the limited amount of both fault data [8] and labelled data [9].

Several efforts have been made to address the lack of labelled fault data. For instance, [10] proposed a new cross-domain bearing fault diagnosis method with few samples under different working conditions. Ref. [11] developed a hybrid machine learning model for fault classification in transmission lines using multi-target ensemble classifier with limited data. Thus, although distinct approaches have been developed to establish data-driven methods for fault classification with limited data, there is still a lack of analysis and formalisation of labelling fault data approaches.

Consequently, Markov-CVAELabeller is introduced in an attempt to label faults presented in a dataset so that fault classification can then be implemented. Firstly, the original time series is transformed into an image, as image encoders have demonstrated their capability to explore patterns that could not be perceived in original raw data. Specifically, studies, such as [12,13], have implemented Markov Transition Field (MTF) encoding, which have demonstrated to amplify local feature information from a 2D perspective and characterise signal fluctuations in a more intuitive manner. Furthermore, the implementation of this type of method also led to improved outcomes [14]. For this reason, in this study, the first-order Markov chain transition matrix is explored as image encoder. Then, a convolutional variational autoencoder is introduced to obtain the latent representation of the input images, which will be utilised to perform clustering analysis, and thus label the distinct faults presented in the dataset. To evaluate the capabilities of the method to label fault data and analyse the impact that labels misclassification, a classification task that considers a convolutional neural network is employed.

This paper is structured as follows. Section 2 presents a literature review with regards to the utilisation of variational autoencoders for fault diagnosis. Section 3 describes the proposed methodology, named Markov-CVAELabeller. Section 4 reflects on the results obtained after implementing the proposed methodology through a case study. To finalise, in Section 5, the conclusions are outlined.

2. Literature Review

2.1. Application of VAE in Fault Diagnosis

The application of variational autoencoder (VAE) in the field of fault diagnosis is not new. For instance, an order track variational stacked autoencoder, which was a fault diagnosis model introduced to address the challenge derived from the combined effect of speed fluctuation and noise variation, was presented by [15]. An optimised stacked variational denoising autoencoder was introduced by [16] so that bearing fault diagnosis could be performed. A VAE for the development of a fault diagnosis method of hidden feature label was considered by [17]. A deep order-wavelet convolutional variational autoencoder was developed by [18] so that bearing faults under fluctuating speed conditions could be identified. Ref. [19] introduced a state-supervised variational autoencoder to capture variations in vibration signal data across distinct operating conditions.

One of the most common applications when considering VAE is the detection of anomalies. Anomaly detection, as part of the diagnostic phase, aims to distinguish data patterns that deviate significantly from normal operation behaviour. For instance, a multimode non-Gaussian VAE (MNVAE) to detect anomalies from unknown distribution vibration signals was introduced by [20]. A long short-term memory (LSTM) in tandem with VAE was applied by [21] so that anomaly detection for shielded cable could be employed. Analogously, a DLSTM-based VAE in tandem with image thresholding to perform anomaly detection for fault diagnosis of marine machinery was introduced by [22]. A VAE-based anomaly detector was also proposed by [23]. Specifically, an anomaly detection scheme for multi-level converters based on wavelet packet transform and VAE (WPT-VAE) was introduced.

Another common application when considering VAE is the augmentation of data. Specifically, when the data are imbalanced. VAE can then be employed to generate synthetic instances in the minority classes so that data can be transformed from imbalanced to balanced. Thus, by performing data augmentation, performance capabilities are enhanced when implementing deep learning approaches [24]. For instance, a VAEGAN-RDCNN for the implementation of bearing fault diagnosis was proposed by [25]. Specifically, a variational autoencoder generative adversarial network (GAN) was introduced for imbalanced data augmentation. Analogously, a variational autoencoder for data augmentation was also employed by [26] and thus addresses the data imbalance challenge in the performance of fault diagnosis in power transformers. A deep meta-learning and VAE as a technique was considered by [27]. Another VAE-based data augmentation method was introduced by [28], which was named CVAEGAN-SM. A conditional variational generative adversarial network (CVAE-GAN) was considered by [29] to both perform fault diagnosis of wind turbine bearings and address the challenge of data imbalance. The VAE was used in this instance as the front end of the GAN generator. Finally, a technique was introduced by [30] to address the challenge of learning from limited data. Thus, the authors utilised split latent subspace encoders in the architecture of the proposed VAE and integrated a domain adaptation strategy. Ref. [31] applied the Equilibrium Deep Q-Network-based agent with the Variational Autoencoder with Wasserstein Generative Adversarial Network and Gradient Penalty for rotating machinery fault diagnosis to integrate both data augmentation techniques with reinforcement learning agent for fault diagnosis. Ref. [32] presented a new method for transformer fault diagnosis based on improved deep residual shrinkage network (DSRN) and optimised residual variational autoencoders (ORVAE). DSRN was introduced to enhance the feature extraction capability. ORVAE was applied to address the challenge of insufficient data.

With regard to feature extraction and fusion, a fault diagnosis model comprising a mixture of Gaussians and VAE (Mix-VAE) was developed by [33]. VAE was employed in this instance for feature extraction multi-sensor data fusion. A deep convolutional variable-beta VAE for the fault diagnosis of rotating machinery was considered by [34]. In this instance, VAE was employed specifically to extract discriminative features. Analogously, a multi-modal variational autoencoder (MMVAE) was proposed by [35] to extract features from multiple modalities. A benchmark rolling bearing dataset was considered to validate the extraction capabilities of the proposed model. Ref. [36] implemented an optimised variational autoencoder architecture to fuse the multirepresentational information of the samples, thus establishing an implicit distribution of the fusion features to increase their robustness.

2.2. Application of Data-Driven Methodologies for the Labelling of Fault Data

Several efforts have been made to develop methods that facilitate the generation of fault labels for unlabelled data. For instance, [37] utilised a cluster generator to automatically divide cluster partitions and add pseudo-labels for these. Ref. [38] labelled the operational data by utilising information contained in status and warning datasets. Ref. [39] introduced a novel diagnostic framework based on label-guided contrastive learning and weighted pseudo-labelling strategy to enhance fault diagnosis accuracy. The authors considered a devised hybrid fine-tuning strategy during the development of the framework so that both labelled and unlabelled data participate in fine-tuning via pseudo-labelling, thus enhancing model generalisation. Ref. [40] developed a self-supervised meta-learning fault diagnosis method for rotating machinery based on label updating.

Thus, despite the numerous studies identified where VAEs are employed as part of the fault diagnosis process, there is no evidence that this type of neural networks has been

utilised to label unlabelled fault data for the application of the fault identification task. Furthermore, there is no evidence that a time series imaging approach has been employed in tandem with a convolutional variational autoencoder to extract discriminative features. Also, even though clustering analysis of the resulting latent space has been considered when performing fault diagnosis [41], there is no evidence, to the best of the authors' knowledge, that the consideration of labelling fault data based on the clustering results was presented. Thus, given the preceding limitations and the challenge of labelling unlabelled fault data for the performance of fault identification, the contributions of this study are highlighted hereunder.

- The introduction of a time series imaging approach based on the first-order Markov chain model to extract the main features of the time series being analysed.
- The application of a convolutional variational autoencoder to extract discriminative features from the time series images.
- The consideration of clustering analysis through the employment of *k*-means to label the unlabelled fault data.
- The performance of fault identification by employing supervised learning based on the labelling results obtained from the Markov-CVAELabeller and the consideration of a convolutional neural network.

3. Methodology

Having explored the main contributions of this study, a graphical representation of the proposed methodology is introduced in Figure 1.

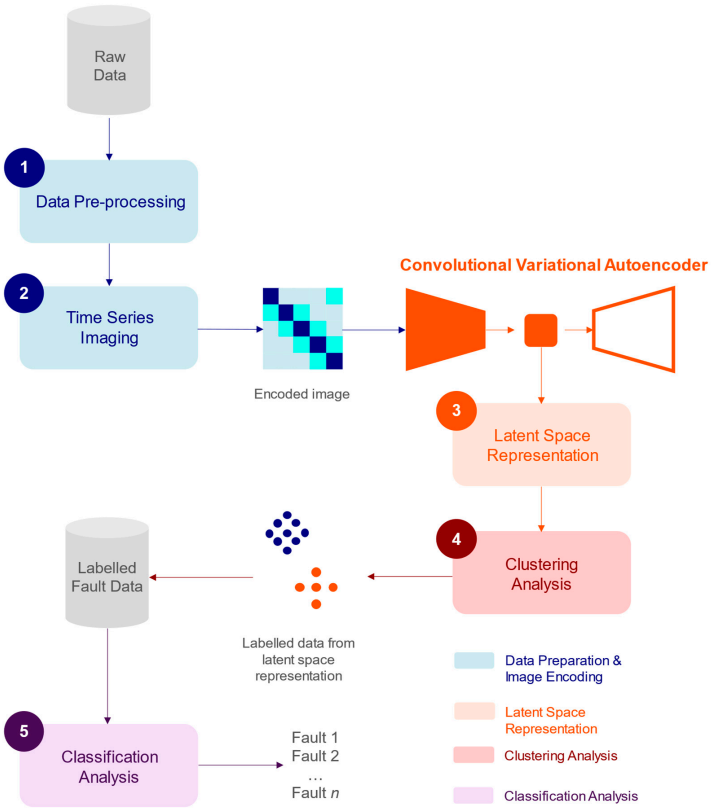


Figure 1. Graphical representation of the proposed methodology.

As it can be perceived, it comprises the following phases:

1. Data Preparation and 2. Image Encoding. These steps aim to address the challenges that data may present. Furthermore, time series are encoded into images through the

application of the first-order Markov chain so that image classifiers for fault classification can be employed.

3. Latent Space Representation. A convolutional variational autoencoder is introduced to obtain the latent space representation of the generated images. Thus, it is expected that discriminative features can be extracted from these time series images.

4. Clustering Analysis. *k*-means is implemented in order to identify the distinct clusters presented in the obtained latent space representation. Thus, the distinct instances can be labelled based on the nature of the fault.

5. Classification Analysis. As the instances are labelled through the application of Markov-CVAELabeller, which refers to the preceding three phases, the fault classification task can then be implemented. In this instance, a convolutional neural network is considered.

3.1. Data Preparation and Image Encoding

To consider the time series data as an image, an image encoding approach is employed. Specifically, the image is generated by considering the transition matrix estimated through the application of the first-order Markov chain. A discrete time stochastic process, $(Z_n)_{n\in\mathbb{N}}$, which takes values in a finite set S, is considered to have the Markov property if the probability distribution of Z_{n+1} at time $n+1$ only hinges on the previous state Z_n, at time n, and not on all the preceding values of Z_k for $k \leq n-1$. Hence,

$$\mathbb{P}(Z_{n+1} = j \mid Z_n = i_n,\ Z_{n-1} = i_{n-1},\ \ldots,\ Z_0 = i_0) = \mathbb{P}(Z_{n+1} = j \mid Z_n = i_n) = p(i,j) \quad (1)$$

where $i_0, i_1, \ldots, i_n, j \in S$. Based on this consideration, the time series is encoded into an image through the consideration of Algorithm 1.

Algorithm 1. Estimation of the transition matrix through the application of the first-order Markov chain.

Input: A collection of occurrences, x, indexed by time.
Number of states, n, that will define the dimensions of the transition matrix $(n \times n)$.

Output: Transition matrix, $\boldsymbol{P}$.

1. Equidistant states are created by following the subsequent steps:
The range of values, r, is defined.
The interval between states, k, is created by following the next equation:

$$k = \frac{r}{n-1}$$

States, s, are generated. Starting from the minimum value, x_{min}, the k is added until the maximum value, x_{max}.

2. Data are assigned to a particular s.

3. The transition matrix $\boldsymbol{P}$ is estimated by considering the distinct transition probabilities, where each (i,j) entry P_{ij} is $p(i,j)$. The probability $p(i,j)$ indicates the probability that the previous state i is followed by the current state j. Thus,

$$\boldsymbol{P} = \left(P_{ij}\right)_{1\leq i,\, j\leq n} = \begin{pmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,n} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n,1} & p_{n,2} & \cdots & p_{n,n} \end{pmatrix}$$

and satisfies

$$0 \ \leq\ P_{ij} \leq 1,\ 1 \ \leq i,\ j \ \leq n,$$
$$\sum_{j=1}^{n} P_{ij} = 1,\ 1 \ \leq i \leq n.$$

3.2. Latent Space Representation

A variational autoencoder (VAE) is a specialised form of the autoencoder (AE) based in Bayesian Inference, which was designed to learn the distribution of data so that it could be represented in a lower-dimensional latent space. Unlike a traditional autoencoder that learns a deterministic latent representation z of the input data x, the encoder $q_\phi(z|x)$ of the VAE approximates the true posterior distribution of the latent variables (see Figure 2). By contrast, the decoder $p_\theta(x|z)$ represents the likelihood of the data generation process, where x is generated from the latent variable z.

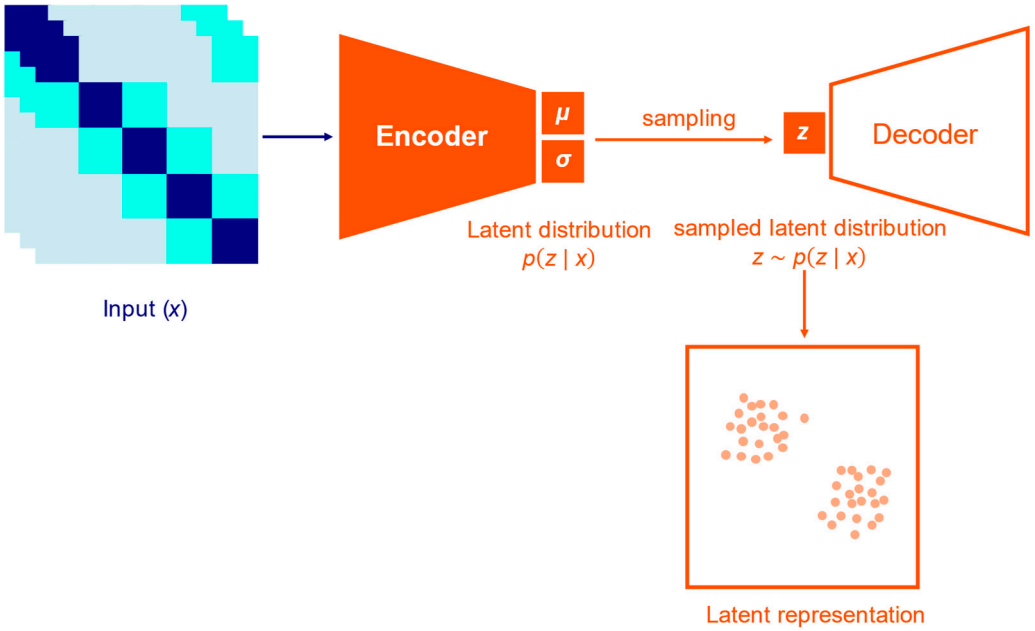


Figure 2. Architecture of the variational autoencoder.

Both the encoder and the decoder of the VAE are modelled as neural networks, which are parametrised by ϕ and θ, respectively. To learn the distribution of the data and how to generate new data samples from the latent space, the VAE is trained by optimising the parameters ϕ and θ by maximising the lower bound of the log-likelihood.

$$\mathcal{L}_{VAE} = D_{KL}(q_\phi(z|x)||\, p_\theta(z)) + E_{q_\phi(z|x)}[\log p_\theta(x|z)] \;\leq \log p(x)$$

Consequently, the encoder compresses the input images, which, in this instance, refer to the transition matrices obtained from the implementation of the first-order Markov chain, into a latent representation, thus capturing essential features while reducing dimensionality. To achieve this, the VAE approximates the latent space to a standard normal distribution $\mathcal{N}(0,\ 1)$ by minimising the Kullback–Leibler (KL) divergence (D_{KL}) between the approximate posterior $q_\phi(z|x)$ and the prior $p_\theta(z)$ of the latent variable [42]. The resulting latent representation is expected to be valuable for capturing the different characteristics of the faults being analysed, enabling their categorisation.

3.3. Clustering Analysis

In order to label the distinct faults based on the clusters presented in the latent space, clustering analysis is performed. Specifically, the k-means algorithm is employed. In this instance, the main aim of the k-means clustering technique is to divide the resulting latent space representation into k clusters, where k is the number of faults to be labelled, by minimising the within-cluster sum of squares. The implementation process can be summarised as follows:

i. **Centroids initialisation.** The centroids of the distinct clusters are initialised at random.
ii. **Cluster assignment.** Each instance is assigned to the nearest cluster. To determine how near an instance is to the cluster, the Euclidean distance between such an instance and the k centroids is determined.
iii. **New centroids computation.** The new centroids are computed based on the instance assignments performed in step ii. To perform this, the mean of all the instances that pertain to a cluster is determined. This is computed for the k clusters.
iv. **Convergence.** Steps ii and iii are implemented until the algorithm converges. The algorithm converges when change is no longer perceived with regards to the cluster assignment.

3.4. Classification Analysis

The generated labels obtained from the application of the preceding steps that constitute the Markov-CVAELabeller model are considered to perform supervised learning and thus apply the fault classification task. To do that, a convolutional neural network (CNN) is considered. This is a type of feedforward artificial neural network that comprises a feature extraction and classification steps.

The main block of the feature extraction stage is the convolutional layer, which consists of a set of filters that convolve with the image and generate a feature map. A pooling layer is also introduced after the implementation of the convolutional layer in order to reduce the dimension of the resulting feature map. To perform the pooling task, the input is sectioned into non-overlapping rectangular subregions. Thus, information from each subregion can be extracted. Even though a loss of information can be perceived, this task assists in both averting overfitting and reducing the computational cost.

Once the features extraction stage has been performed, the classification task is employed. This stage comprises fully connected layers that apply high-level logical operations by considering the features from preceding layers. The output of the final layer is a dimensional vector of equal size from the number of classes being considered.

4. Results

In order to assess the feasibility of applying Markov-CVAELabeller as an automatic label generator for the application of fault diagnosis, a case study is presented in this study. This case study refers to assessing the condition assessment of a hydraulic test rig. The data utilised were experimentally obtained by [43]. The dataset comprises both a primary working and a secondary cooling–filtration circuit. These are connected via the oil tank [44]. For this study, two main faults are assessed, which are the hydraulic accumulator fault and the cooler fault. Thus, in both cases the condition of the cooler and the hydraulic accumulator are close to total failure.

For feature selection, all seventeen parameters have been analysed and their respective performance have been evaluated through the application of the framework. Of all the parameters, the pressure parameter PS1 was considered the most significant. The sampling rate of this parameter is 100 Hz, and its units are bar. In total, 422 instances have been considered (a total of 211 instances for each type of fault). Of all these instances, 70% refer to the training set, and 30% refer to the test set. The process has been applied a total of 10 times to assess the generalisation capabilities of the model.

The following step is to encode the distinct time series into images. To perform that, the first-order Markov chain is applied. A total of 25 states are considered to define the transition matrix, thus being the dimensions of the transition matrix 25×25. A sensitivity analysis was performed to select the number of states, where 20 potential numbers of states were analysed. As observed in Figure 3, the best performance was achieved when the

number of states considered was 25. Analogous performance was perceived for numbers of states greater than 25, although the execution time was longer.

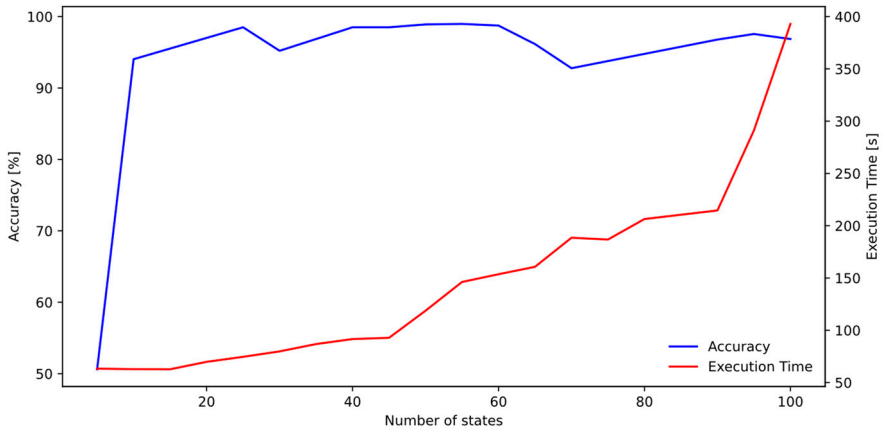


Figure 3. Graphical representation of the results obtained after performing a sensitivity analysis to select the number of states.

With regard to the architecture of the CVAE, the total number of convolutional layers in the encoder is 1 with 32 kernels of size 3×3. The maximum pooling layer is applied after the convolutional layer, and the total of latent dimensions is 2. With respect to the CNN, a total of two convolutional layers are considered with 32 and 64 filters for each layer, respectively, and kernel size of 3×3. The pooling layer is of type maximum with dimensions 2×2. Regarding the fully connected layers, a total of three layers with hidden units 32, 64, and 120 are considered. The activation function is relu. Concerning the configuration of the optimisation and training phase, the Adam optimiser is considered, the number of epochs is 100, and the batch size is set to 5. The main hyperparameters of the CVAE (labelling stage) and CNN (classification stage) are summarised in Tables 1 and 2.

Table 1. Main hyperparameters of the CVAE (labelling stage).

Number of convolutional layers in encoder:	1
Number and size of kernels:	32 of size 3×3
Type of pooling layer:	Max. pooling layer
Latent dimensions:	2

Table 2. Hyperparameters of the CNN (classification stage).

Number of convolutional layers:	2
Number and size of kernels:	32, 64 of size 3×3
Type of pooling layer and dimensions:	Max. pooling layer; 2×2 dimensions
Number of fully connected layers:	3
Hidden units in each fully connected layer:	32, 64, 120
Fully connected layer activation function	relu

Prior to the implementation of the clustering analysis, a comparative study with widely known clustering techniques was conducted. Specifically, the following clustering techniques were considered: (1) *k*-means, (2) Affinity Propagation, (3) Mean Shift, (4) Spectral Clustering, (5) Agglomerative Clustering with Ward Linkage, (6) Density-based Spatial Clustering of Applications with Noise (DBSCAN), (7) Hierarchical Density-based Spatial Clustering of Applications with Noise (HDBSCAN), (8) Ordering Points to Infer Clustering Structure (OPTICS), (9) Balanced Iterative Reducing and Clustering using Hierarchies (BIRCH), and (10) Gaussian Mixture Models. Figure 4 summarises the results of this comparative analysis for one of the analyses. All clustering techniques performed well

in differentiating the two clusters, with the exception of Spectral Clustering, DBSCAN, and HDBSCAN. The remaining models showed similar performance, although *k*-means was selected as it performed slightly better than the other models, was easier to implement, and required less hyperparameter tuning.

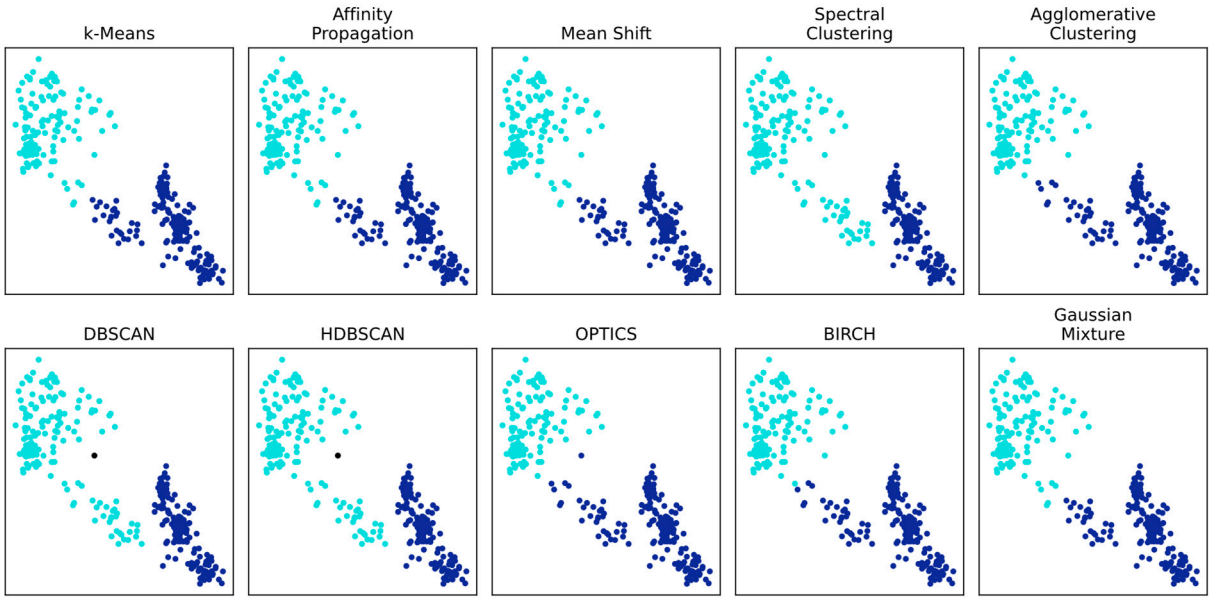


Figure 4. Results of the comparative study of the implemented clustering techniques.

The following paragraphs assess the obtained results. Firstly, the clustering performance of the latent representation is evaluated. Figures 5–8 (left) refer to four examples of resulting latent representations with observed fault labels. For instance, as perceived in Figure 6 (left) two clusters can be observed. The first one is located at the bottom left of the representation and relates to the cooler fault. The second cluster is situated at the top right and refers to the hydraulic accumulator fault. In addition, certain instances are situated between the two clusters, which are associated with the hydraulic accumulator fault. Thus, it can be stated that the cluster that refers to the hydraulic accumulator is more dispersed than the cluster that relates to the cooler fault. Analogous results can be perceived in the other latent representation examples.

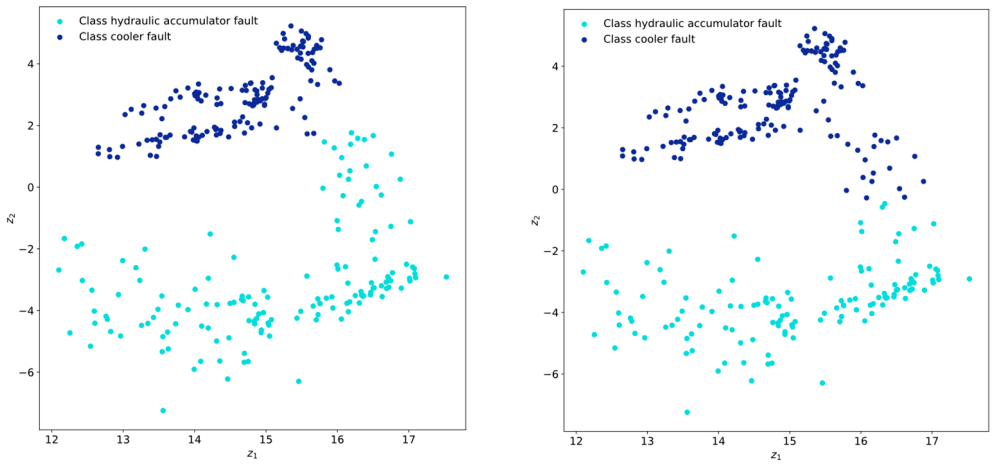


Figure 5. Example 1 of latent representation with observed labels (**left**) and predicted labels (**right**).

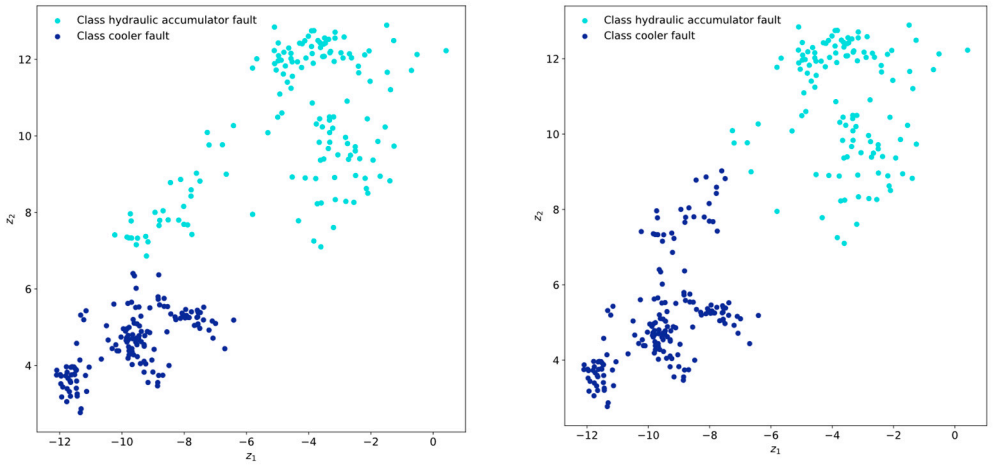


Figure 6. Example 2 of latent representation with observed labels (**left**) and predicted labels (**right**).

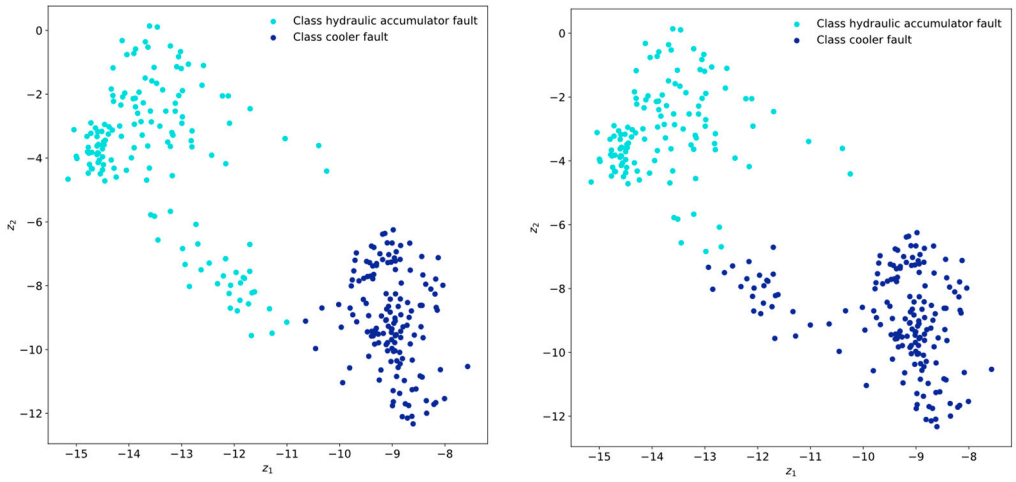


Figure 7. Example 3 of latent representation with observed labels (**left**) and predicted labels (**right**).

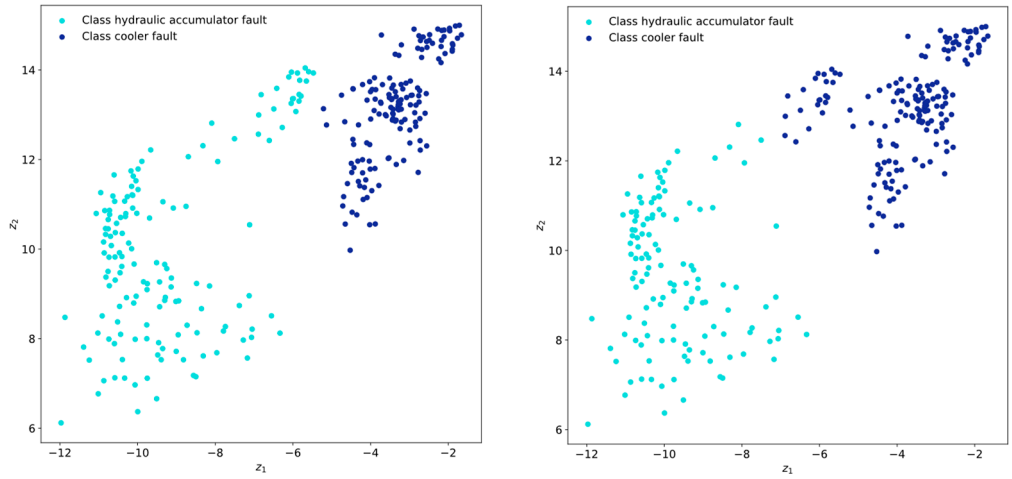


Figure 8. Example 4 of latent representation with observed labels (**left**) and predicted labels (**right**).

The results of the generated labels are presented in Figures 5–8 (right). As it can be observed, the two main clusters are well identified. However, it can be perceived that the labelling in between two clusters that relate to the hydraulic accumulator fault present some labelling discrepancies. Thus, to determine the level of accuracy of the generated labels, a total of three metrics is estimated. These are the accuracy, the precision, and the

recall (please see Table 3). The achieved average accuracy is 92%, which can be considered of a satisfactory result due to the shape of the clusters. Furthermore, if the confusion matrix is also analysed, equal results can be observed than the ones perceived in the graphical representations; certain hydraulic accumulator faults are misclassified as cooler faults (please see Figure 9). These misclassifications may arise because the distinct faults present varying degrees of severity, making it challenging to classify when the severity is high. For instance, it has been considered that the cooler presents a fault when it is close to total failure. However, it is possible that it is experiencing reduced efficiency, which may impact the classification of the fault type.

Table 3. Main results of the labelling performance.

	Markov-CVAELabeller
Accuracy	0.92 ± 0.01
Precision	0.84 ± 0.03
Recall	1.00 ± 0.00

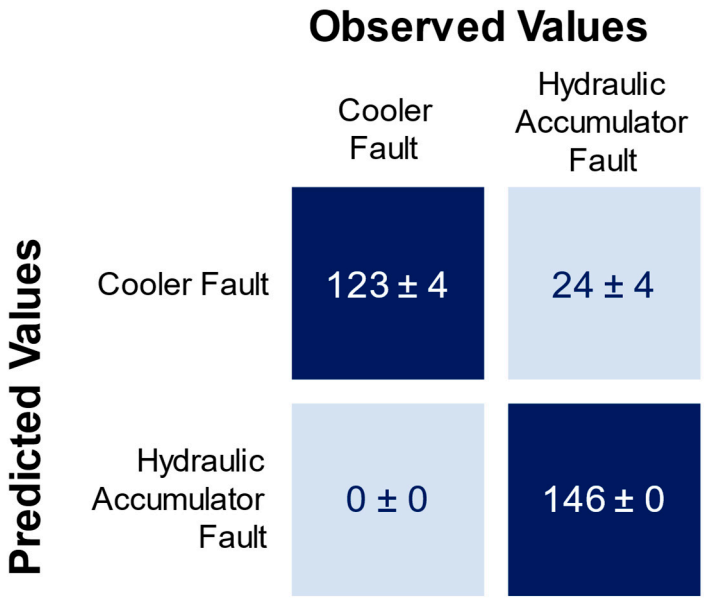


Figure 9. Confusion matrix of the labelling performance.

The misclassification of the hydraulic accumulator fault labels originated a slight decrease in the accuracy performance of the model if compared to the original labels (please see Table 4 and Figure 10). Specifically, the accuracy of the Markov-CVAELabeller-CNN in this study was of 97%, whilst the accuracy of the CNN with observed labels was of 99%, thus experiencing a 2% decrease in the classification performance. Furthermore, a higher variance can be perceived in the results of the CNN if compared to the Markov-CVAELabeller-CNN. This may indicate that certain true labels between the two fault clusters could contribute to the overfitting of the model. Another possibility is that the misclassifications of the pseudo-labels had a favourable impact on the performance consistency of the Markov-CVAELabeller-CNN. However, further research is required to determine the root cause. Thus, despite this decrease, obtaining an accuracy of 97% can be considered promising in the fault classification task when the generation of fault labels is required. As a consequence, the development of this type of method may be a relevant alternative to deal with the lack of fault data.

Table 4. Main results of the classification performance.

	Markov-CVAELabeller-CNN	**CNN**
Accuracy	0.97 ± 0.01	0.99 ± 0.02
Precision	0.84 ± 0.05	1.00 ± 0.00
Recall	1.00 ± 0.00	0.94 ± 0.08

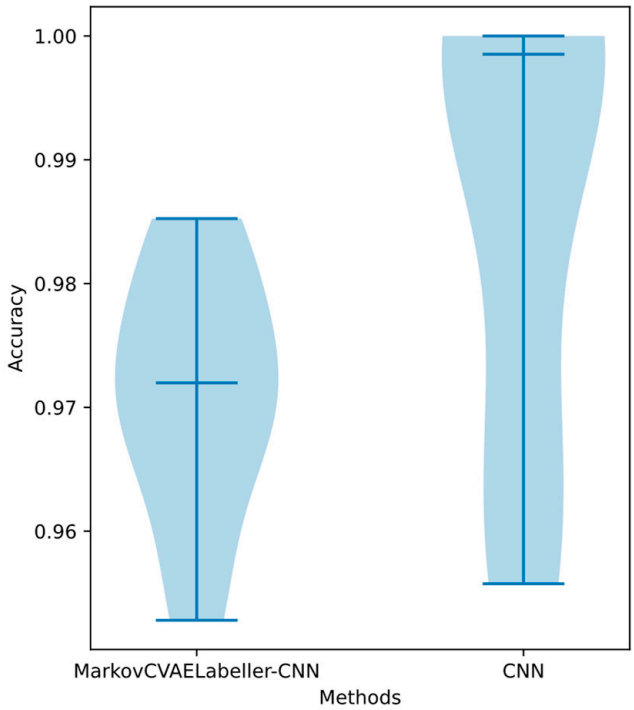


Figure 10. Method comparison in terms of accuracy.

Finally, an additional case study was introduced to assess the generalisation capabilities of the proposed methodology and to determine whether it can handle scenarios where multiple fault classes exist. Consequently, the dataset introduced in [45] was utilised, which relates to an acoustic leakage dataset of gas pipelines. In the first analysis, the following classes were considered: class 0 (0.2 MPa of gas pressure), class 1 (0.4 MPa of gas pressure), and class 2 (0.5 MPa of gas pressure). These three classes also contain high levels of environmental noise and were recorded using microphone 1. Figure 11 shows how the developed model can accurately differentiate between class 0 and classes 1 and 2 due to their significant differences in gas pressure. Classes 1 and 2 can also be clustered adequately, although there is some overlap due to the strong noisy environment and the difference in gas pressure. The accuracy achieved in this instance was 0.98.

The second analysis considered a total of five classes (see Table 5). The remaining classes were not included in this analysis, as no substantial differences were observed between them and the classes currently considered. As shown in Figure 11, similar results can be observed in Figure 12. Class 0 and class 1, which relate to the same gas pressure but different noise environment levels, are very close when compared to the other classes, which correspond to higher gas pressure values. Furthermore, differences between microphones can also be observed when the gas pressure was 0.5 MPa. A possible explanation for this difference could be how these microphones behave depending on environmental factors.

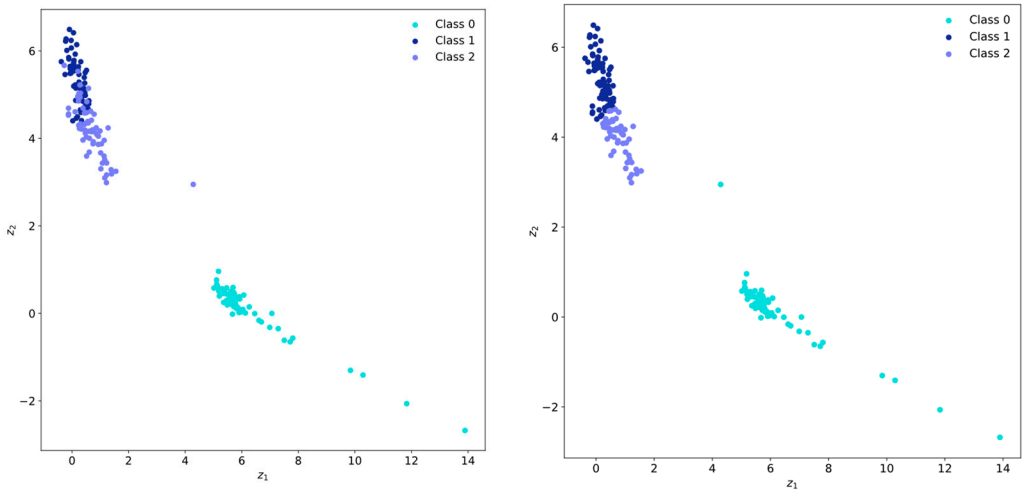


Figure 11. Example of latent representation with observed labels (**left**) and predicted labels (**right**), where number of classes is three.

Table 5. Description of classes utilised in the case study of acoustic leakage dataset of gas pipelines with five classes.

Class	Gas Pressure	Microphone	Environmental Noise
0	0.2 MPa	1	No
1	0.2 MPa	1	Yes
2	0.4 MPa	1	Yes
3	0.5 MPa	1	No
4	0.5 MPa	2	Yes

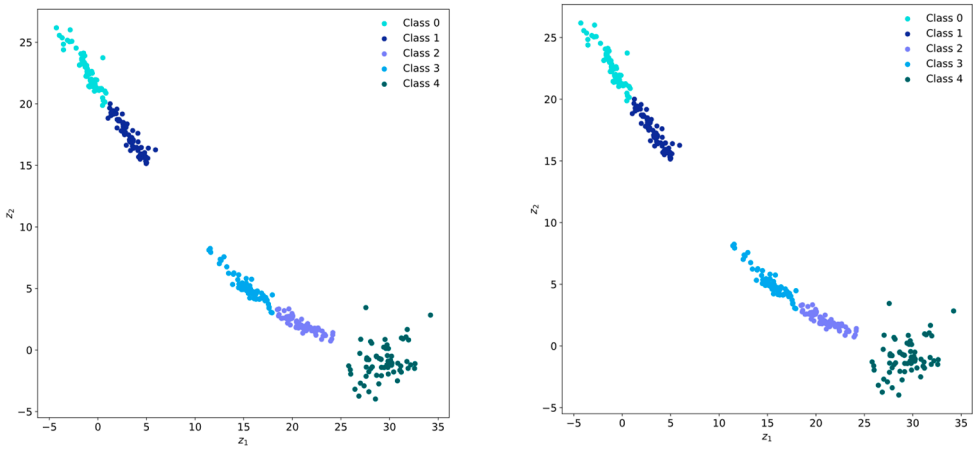


Figure 12. Example of latent representation with observed labels (**left**) and predicted labels (**right**), where number of classes is five.

However, despite these promising results, certain assumptions are required for the adequate performance of the model. These are the following:

- By applying the first-order Markov chain, it is assumed that the time series follows the Markov property. To address such a limitation, other time series imaging methods, such as recurrence plots and Gramian Angular Field approaches, can be studied as part of future work. Additionally, further analysis will be needed for the extension to incipient fault diagnosis in nonlinear systems [46].
- Distinct clusters that refer to the different faults need to be clearly represented in the latent space. Otherwise, the clustering analysis cannot be implemented, and thus

the fault labels cannot be generated. Accordingly, alternatives should be explored in future work to address such a disadvantage. For instance, the application of other deep clustering methods may need to be considered.

- The pseudo-labels obtained from the application of the *k*-means in the latent space may contain errors, which are then utilised for the training of the classification model. Thus, the error may propagate during the performance of the fault diagnosis task, which can yield bias results. Therefore, label refinement and updating strategies need to be considered in future work to minimise mislabels during the labelling stage.

5. Conclusions

Fault diagnosis is of paramount importance to ensure the safety of complex systems. Accordingly, extensive research has been performed in this sense. However, certain challenges are yet to be addressed, for instance, the lack of fault data and the lack of fault labels. For this reason, in this study, the latter aspect is addressed in order to generate labels of those faults that have been collected and that are presented within the dataset.

Thus, the Markov-CVAELabeller is introduced. This method comprises three main phases: (1) image encoding, (2) latent state representation, and (3) clustering analysis. The image encoding phase aims to transform the time series into images through the application of the first-order Markov chain so that image classification can be employed. The second phase aims to obtain a compressed representation of the resulting images through the application of a convolutional variational autoencoder. This compressed representation, also known as latent space representation, is considered for clustering analysis so that the different fault labels can be identified.

To assess the performance of Markov-CVAELabeller, a fault classification task is performed using a convolutional neural network. Additionally, a case study of a hydraulic test rig and a case study of acoustic leakage in gas pipelines are introduced. Results show that the application of Markov-CVAELabeller can be a viable option for labelling fault data prior to performing supervised fault classification tasks. Furthermore, the application of the first-order Markov chain to encode time series data into images enabled the effective representation of faults in the latent space, thus enhancing the model's ability to cluster distinct fault types. The clusters defined in the latent space can then facilitate the generation of pseudo-labels when labelled fault data are scarce. Therefore, the proposed model can serve as an alternative for obtaining fault label data. However, the proposed method presents certain limitations that are expected to be addressed in future research. Examples of these limitations include the assumption that the time series follows a Markov property and that distinct clusters can be identified in the latent space representation. Accordingly, the combination of different time series imaging techniques, such as Recurrence Plots and Gramian Angular Fields, as well as the exploration of other deep clustering approaches, will be considered in future work.

Author Contributions: Conceptualization, C.V.-G. and N.C.-M.; methodology, C.V.-G.; validation, N.C.-M.; formal analysis, C.V.-G.; investigation, C.V.-G.; writing—original draft preparation, C.V.-G. and N.C.-M.; writing—review and editing, C.V.-G. and N.C.-M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data utilised in the first case study were experimentally obtained from ([43,44]). Data can be accessed by utilising the following link: https://archive.ics.uci.edu/dataset/447/condition+monitoring+of+hydraulic+systems (accessed on 16 March 2025). The data utilised in the second case study were obtained from ([45]). Data can be accessed by utilizing the following link: https://github.com/Deep-AI-Application-DAIP/acoustic-leakage-dataset-GPLA-12/tree/main (accessed on 16 March 2025).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Tian, Y.; Zhao, X.; Huang, W. Meta-learning approaches for learning-to-learn in deep learning: A survey. *Neurocomputing* **2022**, *494*, 203–223. [CrossRef]
2. Stefanidou-Voziki, P.; Sapountzoglou, N.; Raison, B.; Dominguez-Garcia, J. A review of fault location and classification methods in distribution grids. *Electr. Power Syst. Res.* **2022**, *209*, 108031. [CrossRef]
3. Wang, J.-G.; Cai, X.-Z.; Yao, Y.; Zhao, C.; Yang, B.-H.; Ma, S.-W.; Wang, S. Statistical process fault isolation using robust nonnegative garrote. *J. Taiwan Inst. Chem. Eng.* **2020**, *107*, 24–34. [CrossRef]
4. Prasad, A.; Edward, J.B.; Ravi, K. A review on fault classification methodologies in power transmission systems: Part—I. *J. Electr. Syst. Inf. Technol.* **2018**, *5*, 48–60. [CrossRef]
5. Galar, D.; Kumar, U. *eMaintenance*; Elsevier: Amsterdam, The Netherlands, 2017; pp. 235–310. [CrossRef]
6. Lei, Y.; Yang, B.; Jiang, X.; Jia, F.; Li, N.; Nandi, A.K. Applications of machine learning to machine fault diagnosis: A review and roadmap. *Mech. Syst. Signal Process.* **2020**, *138*, 106587. [CrossRef]
7. Zhou, X.; Liu, H.; Pourpanah, F.; Zeng, T.; Wang, X. A survey on epistemic (model) uncertainty in supervised learning: Recent advances and applications. *Neurocomputing* **2022**, *489*, 449–465. [CrossRef]
8. Gilanifar, M.; Wang, H.; Cordova, J.; Ozguven, E.E.; Strasser, T.I.; Arghandeh, R. Fault classification in power distribution systems based on limited labeled data using multi-task latent structure learning. *Sustain. Cities Soc.* **2021**, *73*, 103094. [CrossRef]
9. Velasco-Gallego, C.; De Maya, B.N.; Molina, C.M.; Lazakis, I.; Mateo, N.C. Recent advancements in data-driven methodologies for the fault diagnosis and prognosis of marine systems: A systematic review. *Ocean Eng.* **2023**, *284*, 115277. [CrossRef]
10. Dong, X.; Zhang, C.; Liu, H.; Wang, D.; Chen, Y.; Wang, T. A new cross-domain bearing fault diagnosis method with few samples under different working conditions. *J. Manuf. Process.* **2025**, *135*, 359–374. [CrossRef]
11. El Ghaly, A. Hybrid ML Algorithm for Fault Classification in Transmission Lines Using Multi-Target Ensemble Classifier with Limited Data. *Eng* **2025**, *6*, 4. [CrossRef]
12. Dong, S.; Meng, Y.; Yin, S.; Liu, X. Tool wear state recognition study based on an MTF and a vision transformer with a Kolmogorov-Arnold network. *Mech. Syst. Signal Process.* **2025**, *228*, 112473. [CrossRef]
13. Abidi, A.; Ienco, D.; Ben Abbes, A.; Farah, I.R. Combining 2D encoding and convolutional neural network to enhance land cover mapping from Satellite Image Time Series. *Eng. Appl. Artif. Intell.* **2023**, *122*, 106152. [CrossRef]
14. Ma, J.; Wang, H. Anomaly detection in sensor data via encoding time series into images. *J. King Saud Univ. Comput. Inf. Sci.* **2024**, *36*, 102232. [CrossRef]
15. Yu, T.; Li, S.; Lu, J.; Gong, S. Order Track Variational Stacked Autoencoder Fault Diagnosis Model for Complex Working Conditions. In Proceedings of the 2022 Global Reliability and Prognostics and Health Management (PHM-Yantai), Yantai, China, 13–16 October 2022; pp. 1–8.
16. Yan, X.; Xu, Y.; She, D.; Zhang, W. Reliable Fault Diagnosis of Bearings Using an Optimized Stacked Variational Denoising Auto-Encoder. *Entropy* **2021**, *24*, 36. [CrossRef]
17. She, B.; Wang, X. A hidden feature label propagation method based on deep convolution variational autoencoder for fault diagnosis. *Meas. Sci. Technol.* **2022**, *33*, 055107. [CrossRef]
18. Yan, X.; She, D.; Xu, Y. Deep order-wavelet convolutional variational autoencoder for fault identification of rolling bearing under fluctuating speed conditions. *Expert Syst. Appl.* **2022**, *216*, 119479. [CrossRef]
19. Xiao, Y.; Feng, K.; Miao, D.; Zhang, P.; Yang, J. A State-Supervised Model and Novel Anomaly Index for Gas Turbines Blade Fault Detection Under Multi-Operating Conditions. *IEEE Access* **2025**, *13*, 14225–14238. [CrossRef]
20. Luo, Q.; Chen, J.; Zi, Y.; Chang, Y.; Feng, Y. Multi-mode non-Gaussian variational autoencoder network with missing sources for anomaly detection of complex electromechanical equipment. *ISA Trans.* **2022**, *134*, 144–158. [CrossRef]
21. Chang, S.J.; Kwon, G.-Y. Anomaly Detection for Shielded Cable Including Cable Joint Using a Deep Learning Approach. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 4025. [CrossRef]
22. Velasco-Gallego, C.; Lazakis, I. RADIS: A real-time anomaly detection intelligent system for fault diagnosis of marine machinery. *Expert Syst. Appl.* **2022**, *204*, 117634. [CrossRef]

23. Ye, S.; Zhang, F. Unsupervised anomaly detection for multilevel converters based on wavelet transform and variational autoencoders. In Proceedings of the 2022 IEEE Energy Conversion Congress and Exposition (ECCE), Detroit, MI, USA, 9–13 October 2022; pp. 1–6.
24. Temraz, M.; Keane, M.T. Solving the class imbalance problem using a counterfactual method for data augmentation. *Mach. Learn. Appl.* **2022**, *9*, 100375. [CrossRef]
25. Rathore, M.S.; Harsha, S.P. Non-linear Vibration Response Analysis of Rolling Bearing for Data Augmentation and Characterization. *J. Vib. Eng. Technol.* **2022**, *11*, 2109–2131. [CrossRef]
26. Vidal, J.F.; Castro, A.R.G. Diagnosing Faults in Power Transformers with Variational Autoencoder, Genetic Programming, and Neural Network. *IEEE Access* **2023**, *11*, 30529–30545. [CrossRef]
27. Che, C.; Wang, H.; Lin, R.; Ni, X. Deep meta-learning and variational autoencoder for coupling fault diagnosis of rolling bearing under variable working conditions. *Proc. Inst. Mech. Eng. C J. Mech. Eng. Sci.* **2022**, *236*, 9900–9913. [CrossRef]
28. Liu, Y.; Jiang, H.; Wang, Y.; Wu, Z.; Liu, S. A conditional variational autoencoding generative adversarial networks with self-modulation for rolling bearing fault diagnosis. *Measurement* **2022**, *192*, 110888. [CrossRef]
29. Zhang, L.; Zhang, H.; Cai, G. The Multiclass Fault Diagnosis of Wind Turbine Bearing Based on Multisource Signal Fusion and Deep Learning Generative Model. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 8483. [CrossRef]
30. Li, T.; Fung, C.-H.; Wong, H.-T.; Chan, T.-L.; Hu, H. Functional Subspace Variational Autoencoder for Domain-Adaptive Fault Diagnosis. *Mathematics* **2023**, *11*, 2910. [CrossRef]
31. Li, Z.; Jiang, H.; Wang, X. A novel reinforcement learning agent for rotating machinery fault diagnosis with data augmentation. *Reliab. Eng. Syst. Saf.* **2024**, *253*, 110570. [CrossRef]
32. Yao, H.; Xu, Y.; Guo, Q.; Chen, S.; Lu, B.; Huang, Y. Study on transformer fault diagnosisbased on improved deep residual shrinkage network and optimized residual variational autoencoder. *Energy Rep.* **2025**, *13*, 1608–1619. [CrossRef]
33. Wang, C.; Xin, C.; Xu, Z.; Qin, M.; He, M. Mix-VAEs: A novel multisensor information fusion model for intelligent fault diagnosis. *Neurocomputing* **2022**, *492*, 234–244. [CrossRef]
34. Dewangan, G.; Maurya, S. Fault Diagnosis of Machines Using Deep Convolutional Beta-Variational Autoencoder. *IEEE Trans. Artif. Intell.* **2021**, *3*, 287–296. [CrossRef]
35. Xiong, M.; Wu, Y.; Li, C.; Yang, Z. Rolling Bearing Fault Diagnosis Based on Multi-Modal Variational Autoencoders. In Proceedings of the 2022 International Conference on Sensing, Measurement & Data Analytics in the era of Artificial Intelligence (ICSMD), Harbin, China, 30 November–2 December 2022; pp. 1–5.
36. Yang, X.; Cao, X.; Zhao, J.; Zhang, X.; Duan, Y.; Shi, L. Intelligent fault diagnosis method for rotating machinery based on sample multire presentation information fusion under limited labeled samples conditions. *Measurement* **2025**, *250*, 117164. [CrossRef]
37. Zhao, Z.; Jiao, Y.; Xu, Y.; Chen, Z.; Zio, E. A fault diagnosis framework using unlabeled data based on automatic clustering with meta-learning. *Eng. Appl. Artif. Intell.* **2024**, *139*, 109584. [CrossRef]
38. Fazli, A.; Poshtan, J. Wind turbine fault prognosis using SCADA measurements, pre-fault labeling, and KNN classifiers robust against data imbalance. *Measurement* **2024**, *243*, 116202. [CrossRef]
39. Li, X.; Cheng, C.; Peng, Z. Label-guided contrastive learning with weighted pseudo-labeling: A novel mechanical fault diagnosis method with insufficient annotated data. *Reliab. Eng. Syst. Saf.* **2024**, *254*, 110597. [CrossRef]
40. Zhao, Z.; Jiao, Y.; Xu, Y.; Chen, Z.; Zhao, R. Smeta-LU: A self-supervised meta-learning fault diagnosis method for rotating machinery based on label updating. *Adv. Eng. Inform.* **2024**, *62*, 102875. [CrossRef]
41. Ibrahim, R.; Zemouri, R.; Kedjar, B.; Merkhouf, A.; Tahan, A.; Al-Haddad, K.; Lafleur, F. Non-invasive Detection of Rotor Inter-turn Short Circuit of a Hydrogenerator Using AI-Based Variational Autoencoder. *IEEE Trans. Ind. Appl.* **2023**, *60*, 28–37. [CrossRef]
42. Han, P.; Ellefsen, A.L.; Li, G.; Holmeset, F.T.; Zhang, H. Fault Detection With LSTM-Based Variational Autoencoder for Maritime Components. *IEEE Sens. J.* **2021**, *21*, 21903–21912. [CrossRef]
43. Helwig, N.; Pignanelli, E.; Schutze, A. D8. 1–Detecting and Compensating Sensor Faults in a Hydraulic Condition Monitoring System. In Proceedings of the AMA Conferences 2015—SENSOR 2015 and IRS 2015, Nuremberg, Germany, 19–21 May 2015; pp. 641–646.
44. Helwig, N.; Pignanelli, E.; Schütze, A. Condition monitoring of a complex hydraulic system using multivariate statistics. In Proceedings of the 2015 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings, Pisa, Italy, 11–14 May 2015; pp. 210–215. [CrossRef]

45. Li, J.; Yao, L. GPLA-12: An Acoustic Signal Dataset of Gas Pipeline Leakage. *arXiv* **2021**, arXiv:2106.10277.
46. Safaeipour, H.; Forouzanfar, M.; Puig, V.; Birgani, P.T. Incipient fault diagnosis and trend prediction in nonlinear closed-loop systems with Gaussian and non-Gaussian noise. *Comput. Chem. Eng.* **2023**, *177*, 108348. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.