

Titre: Développement d'une méthode pour la détection d'anomalies dans des signaux de capteurs : application dans des systèmes aéronautiques
Title:

Auteur: Luis Miguel Moreno Haro
Author:

Date: 2024

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Moreno Haro, L. M. (2024). Développement d'une méthode pour la détection d'anomalies dans des signaux de capteurs : application dans des systèmes aéronautiques [Mémoire de maîtrise, Polytechnique Montréal]. PolyPublie.
Citation: <https://publications.polymtl.ca/61636/>

 **Document en libre accès dans PolyPublie**
Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/61636/>
PolyPublie URL:

Directeurs de recherche: Bruno Agard, & Souheil-Antoine Tahan
Advisors:

Programme: Maîtrise recherche en génie industriel
Program:

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

**Développement d'une méthode pour la détection d'anomalies dans des signaux
de capteurs : application dans des systèmes aéronautiques**

LUIS MIGUEL MORENO HARO

Département de mathématiques et de génie industriel

Mémoire présenté en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*
Génie industriel

Décembre 2024

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Ce mémoire intitulé :

**Développement d'une méthode pour la détection d'anomalies dans des signaux
de capteurs : application dans des systèmes aéronautiques**

présenté par **Luis Miguel MORENO HARO**

en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*
a été dûment accepté par le jury d'examen constitué de :

Martin TRÉPANIÉ, président

Bruno AGARD, membre et directeur de recherche

Antoine TAHAN, membre et codirecteur de recherche

Lucas HOF, membre

DÉDICACE

*Pour mes parents, Alicia et Lino,
qui m'embrassent toujours avec leur amour...*

REMERCIEMENTS

Je tiens à remercier mon directeur Bruno Agard et mon codirecteur Antoine Tahan de m'avoir donné l'opportunité de travailler avec eux sur ce projet. Je suis reconnaissant envers mes collègues du Laboratoire en Intelligence de Données (LID). Je tiens également à exprimer ma gratitude aux personnels chez le partenaire industriel du projet pour leur grande disponibilité et pour leur gentillesse tout au long de ce projet. Je suis également profondément reconnaissant envers MITACS pour le soutien financier qui a rendu possible la réalisation de ce projet. De plus, je tiens à remercier mes parents, mon frère et ma soeur, mes amis pour leur soutien inestimable pendant cette étape de ma vie.

RÉSUMÉ

Ce projet vise à développer une méthode de détection des anomalies dans les séries temporelles de mesures obtenues par des capteurs implémentés dans des systèmes aéronautiques. En collaboration avec un partenaire industriel, nous avons utilisé les données d'un ensemble de capteurs d'un système commercial et industriel. Les données collectées étaient présentées sous forme de séries chronologiques multivariées.

Après une revue de la littérature, nous avons identifié différents types de défaillances des capteurs, tels que le biais, l'étincelle, le gel, le bruit et la saturation. À partir de ces éléments, nous avons développé une méthodologie en quatre étapes :

1. **Sélection de la région d'intérêt** : Identification des zones de variations significatives dans les données.
2. **Calcul de descripteurs statistiques** : Calcul de la moyenne, variance, obliquité, kurtosis, et entropie pour caractériser les séries temporelles. Ces descripteurs ont été choisis après avoir réalisé des tests et déterminé leur impact sur la détection des différents types de défaillances constatées dans les séries temporelles.
3. **Conversion en images en niveaux de gris** : Conversion des données des descripteurs statistiques normalisés en images .
4. **Formation et validation du modèle d'apprentissage automatique** : Calcul de l'erreur de reconstruction et analyse dans l'espace latent à l'aide d'un modèle CVAE (Variational Autoencoder avec filtres CNN).

La méthode proposée a montré son efficacité pour la détection d'anomalies dans les données de mesure des capteurs. En transformant les données de séries temporelles en images caractéristiques, notre modèle CVAE a permis une séparation claire entre les événements normaux et anormaux. Les résultats, basés sur les courbes caractéristiques de l'opérateur du récepteur (ROC) avec une aire sous la courbe (AUC) de 1.00 pour l'erreur de reconstruction, ainsi que sur les distances euclidiennes dans l'espace latent, montrent que notre approche est robuste et précise. Cependant, il est important de noter que ce résultat de 1,00 est le fruit d'une identification réussie des capteurs présentant des défaillances spécifiques, notamment celles liées au gel. En intégrant des données provenant d'autres capteurs, nous anticipons une diminution de l'AUC, car le modèle n'a pas encore appris à reconnaître ces nouveaux cas. En d'autres termes, dans un scénario idéal, le modèle devrait être formé sur l'ensemble des types de défaillances dans chacun des capteurs de tout le système pour maintenir une aire sous la

courbe d'une valeur très élevée. Cette méthode non seulement fournit un cadre solide pour la détection des défaillances des capteurs, mais elle offre également une visualisation complète des types de défaillances dans l'espace latent, contribuant ainsi à la création de systèmes de capteurs aéronautiques plus résilients.

ABSTRACT

This project aims to develop a method for failure detection in time series, specifically tailored for monitoring aeronautical sensors. We utilized sensors from a commercial and industrial system. The collected data is presented as a multivariate time series.

After reviewing the literature, we categorized anomaly detection techniques into statistical methods, machine learning approaches, and combinations of both. We also identified various types of sensor failures, including bias, spark, frozen, noise, and saturation. Based on these approaches, we developed a methodology consisting of four steps:

1. **Detection of the Region of Interest** : Identification of areas with significant variations in the data.
2. **Calculation of Statistical Descriptors** : Utilization of mean, variance, skewness, kurtosis, and entropy to characterize the time series. These features were chosen after carrying out tests and determining their impact on the detection of the different types of defects observed in the time series.
3. **Image Grayscale Creation** : Conversion of data into images using the statistical features and normalization of data between -1 and 1 to ensure a uniform scale.
4. **Application of a Machine Learning Model** : Calculation of reconstruction error and visualization in the latent space using a CVAE (Variational Autoencoder with CNN filters).

The proposed method demonstrated its effectiveness in detecting anomalies in sensor monitoring. By transforming time series data into feature-based images, our CVAE model achieved clear separation between normal and abnormal events. Results, based on Receiver Operator Characteristic (ROC) curves with an area under the curve (AUC) of 1.00 for reconstruction error, and Euclidean distances in the latent space, indicate that our approach is both robust and precise. However, it is important to note that this result of 1.00 is the result of successfully identifying sensors with specific failures, especially those related to frozen failures. By integrating data from other sensors, we expect a decrease in AUC, because the model has not yet learned to recognize these new cases. In other words, in an ideal scenario, the model should be trained on all failure types to maintain a very high area under the curve value. This method not only provides a solid framework for sensor failure detection but also offers a comprehensive visualization of fault types in the latent space, contributing to the development of more resilient aeronautical sensor systems.

TABLE DES MATIÈRES

DÉDICACE	iii
REMERCIEMENTS	iv
RÉSUMÉ	v
ABSTRACT	vii
TABLE DES MATIÈRES	viii
LISTE DES TABLEAUX	x
LISTE DES FIGURES	xi
LISTE DES SIGLES ET ABRÉVIATIONS	xiii
CHAPITRE 1 INTRODUCTION	1
CHAPITRE 2 ÉTAT DE L'ART	3
2.1 Définition d'un capteur	3
2.2 Types de défaillances d'un capteur	3
2.3 Techniques de détection des défaillances de capteurs	8
2.4 Synthèse des méthodes selon les défaillances détectées	14
CHAPITRE 3 PROBLÉMATIQUE	16
CHAPITRE 4 MÉTHODOLOGIE PROPOSÉE	18
4.1 Analyse statistique des données	18
4.1.1 Nomenclature	18
4.1.2 Descripteurs statistiques	19
4.1.3 Matrices des descripteurs statistiques	20
4.1.4 Corrélation de Pearson	22
4.2 Méthodologie proposée	22
4.3 Étapes proposées	27
4.4 Indices de santé	39
CHAPITRE 5 RÉSULTATS ET DISCUSSIONS	41

5.1	Résultats	41
5.1.1	Images basées sur les descripteurs à partir des données des capteurs .	41
5.1.2	Architecture et formation CVAE	42
5.1.3	Résultats : Erreur de reconstruction - image composite à 5 descripteurs	43
5.1.4	Résultats : Erreur de reconstruction - image composite à 6 descripteurs	43
5.1.5	Résultats : Espace latent - image composite à 5 descripteurs	44
5.1.6	Résultats : Espace latent - image composite à 6 descripteurs	47
5.1.7	Évaluation de l'approche de détection proposée	48
5.2	Déformation temporelle dynamique : méthodologie indépendante	50
CHAPITRE 6	CONCLUSION	54
RÉFÉRENCES	56

LISTE DES TABLEAUX

Tableau 2.1	Relation entre la méthode et le domaine avec le type de défaillance des capteurs	14
-------------	--	----

LISTE DES FIGURES

Figure 1.1	Fonctionnement schématique d'un capteur	1
Figure 2.1	Signal d'un phénomène physique (inconnu) et signal mesuré par un capteur sain.	4
Figure 2.2	Défaillance de type biais du capteur	5
Figure 2.3	Défaillance de type saturation du capteur	5
Figure 2.4	Défaillance du signal de "type gel capteur"	6
Figure 2.5	Défaillance de bruit du capteur	7
Figure 2.6	Défaillance de type étincelle du capteur	8
Figure 4.1	Descripteurs statistiques à partir de la série chronologique i de l'événement j	19
Figure 4.2	Construction d'une matrice de moyennes et sa visualisation à partir d'une base de données d'un système industriel réel	21
Figure 4.3	Schéma de reconstruction et erreur de reconstruction e du modèle CVAE proposé	24
Figure 4.4	Organigramme illustrant les étapes de la méthodologie proposée pour la détection des défaillances de capteurs dans les systèmes industriels	27
Figure 4.5	Exemple de détection de la région d'intérêt d'un événement $j = 1$ avec des séries temporelles de i capteurs	28
Figure 4.6	Signal d'un capteur i utilisé pour effectuer le processus de détection de la région d'intérêt	29
Figure 4.7	Relation entre les capteurs d'un événement et l'image avec pixels de corrélation	30
Figure 4.8	Structure de l'auto-encodeur variationnel	33
Figure 4.9	Capteur anonyme en noir dans le système avec injection de défaillance synthétique de type étincelle en rouge	36
Figure 5.1	Images en niveaux de gris : représentation visuelle des caractéristiques des données pour différentes conditions d'état	41
Figure 5.2	Erreur de reconstruction en utilisant des images à 5 descripteurs	43
Figure 5.3	Erreur de reconstruction en utilisant des images à 6 descripteurs	44
Figure 5.4	Espace latent utilisant des images à 5 descripteurs	45
Figure 5.5	Distance entre une image anormale et le centre du groupe normal (graphique rouge) et une image anormale et le centre du groupe gelé (graphique orange). Utilisation d'images à 5 descripteurs	46

Figure 5.6	Espace latent utilisant des images à 6 descripteurs	47
Figure 5.7	Distance entre une image anormale et le centre du groupe normal (graphique rouge) et une image anormale et le centre du groupe gelé (graphique orange). Utilisation d'images à 6 descripteurs	48
Figure 5.8	Courbe ROC de la méthodologie proposée	49
Figure 5.9	Organigramme présentant les étapes de la méthodologie proposée en utilisant déformation temporelle dynamique (DTW) pour la détection des défaillances de capteurs dans les systèmes industriels	51
Figure 5.10	Histogramme présentant les résultats de la méthodologie, avec les échantillons de type 'étincelle' en gris et les échantillons de type 'sain' en bleu	52
Figure 5.11	Histogramme présentant les résultats de la méthodologie, avec les échantillons de type 'saturation' en rose et les échantillons de type 'sain' en bleu	53

LISTE DES SIGLES ET ABRÉVIATIONS

AE	Auto-Encoder
AUC	Area Under the Curve
CNN	Convolutional Neural Network
CVAE	Convolutional Neural Network Variational Auto-Encoder
CPS	Cyber-physical system
DB2	Daubechies 2
ER	Error Reconstruction
GAF	Gramian Angular Field
KNN	K-Nearest Neighbor
LSTM	Long Short-Term Memory
LS	Latent Space
MSE	Mean Squared Error
ML	Machine Learning
NaN	Not a Number
PCA	Principal Component Analysis
PNN	Probabilistic Neural Network
RNN	Recurrent Neural Network
ROC	Receiver Operating Characteristic
RPROP	Retro Propagation
SVM	Support Vector Machine
VAE	Variational Auto-Encoder

CHAPITRE 1 INTRODUCTION

La détection automatique d'anomalies est une approche qui est utilisée dans divers domaines de l'ingénierie. Elle vise à éviter des coûts liés à des défaillances manquées, des fausses alarmes ou des temps d'arrêt imprévus d'un système. Ainsi, la surveillance (en mode continue ou en mode intermitente) des systèmes est une technique dont le but est d'identifier des comportements inhabituels.

La détection à l'aide des capteurs peut donc amener un gain de productivité et une amélioration de la sûreté de fonctionnement du système en question. La Figure 1.1 montre le fonctionnement schématique d'un capteur dont le but est de mesurer un phénomène physique et de générer un signal interprétable qui peut être analogique ou digital.

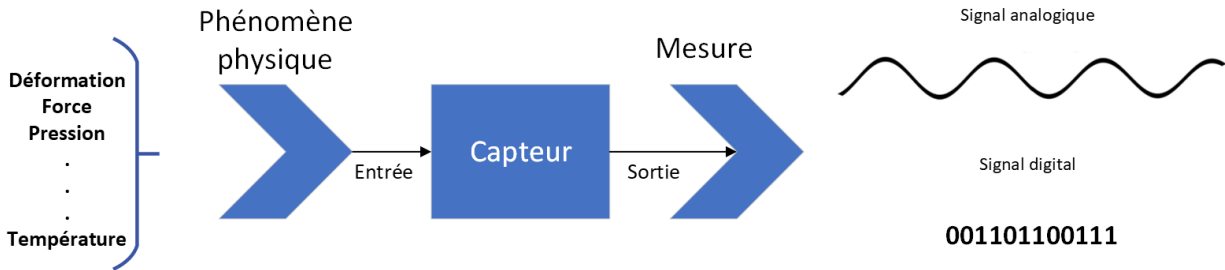


FIGURE 1.1 Fonctionnement schématique d'un capteur

De manière générale, la mesure résultante d'un capteur varie dans le temps et est typiquement modélisée par une série temporelle. Des techniques d'analyse des données chronologiques ou fréquentielles peuvent être appliquées pour identifier des comportements inhabituels ou anormaux dans ces mesures.

Les progrès technologiques dans le domaine des capteurs ont conduit à une capacité croissante à collecter des informations de plus en plus détaillées et précises. Ainsi, les capteurs, ou les systèmes de capteurs, jouent un rôle important pour statuer de l'état d'un système. Dans cette perspective, les capteurs doivent impérativement fournir des informations fiables.

Le principal problème est que les capteurs ne fonctionnent pas toujours correctement, ils peuvent donner des mesures anormales. La détection des anomalies des capteurs est donc vitale pour garantir un fonctionnement approprié d'un système. On définit une défaillance comme le fait, pour un système (dans notre cas, les capteurs), de ne pas assurer pleinement son rôle ou sa fonction. Différents types de défaillances peuvent affecter les capteurs : biais, saturation, erreur de gain, gel du signal, bruit excessif, dérive et étincelle.

Cette étude vise à développer des méthodes pour la détection des anomalies des capteurs dans les systèmes aéronautiques.

Le document est structuré en six chapitres. Le chapitre 1 introduit le sujet. Dans le chapitre 2, on présente la problématique, la définition d'un capteur, les types de défaillances étudiés et les techniques de détection des défaillances de capteurs. Le chapitre 3 approfondit la problématique, les données disponibles et les objectifs spécifiques du projet. Le chapitre 4 décrit la méthodologie proposée. Le chapitre 5 présente les principaux résultats et le chapitre 6 conclut le travail.

CHAPITRE 2 ÉTAT DE L'ART

Ce chapitre présente la définition d'un capteur, accompagnée d'une classification des types de défaillances. Nous y présentons des techniques de détection de défaillance dans les capteurs telles que recensées dans la littérature scientifique. Finalement, un tableau résumé présente les types de techniques selon le type de défaillances détectées.

2.1 Définition d'un capteur

Un capteur est un dispositif spécialisé qui mesure une entrée physique dans son environnement et la transforme en données exploitables qui sont interprétées par un système informatique [Ali et al., 2023]. Cependant, ce signal brut peut nécessiter un traitement supplémentaire avant d'être exploitable par un système informatique. Ce traitement est effectué par le circuit de conditionnement, qui fait partie d'un système de mesure complet.

Les capteurs sont des composants essentiels des systèmes d'acquisition de données. Cependant, leur fiabilité peut être compromise suite à une défaillance ou en raison de leur environnement de travail. Par conséquent, l'identification précoce des défaillances des capteurs est essentielle afin de mettre en place des mesures correctives [Li et al., 2020].

2.2 Types de défaillances d'un capteur

Une défaillance de capteur est définie comme une perte partielle ou totale de fonctionnalité [Moreno Haro et al., 2023]. La défaillance entraîne donc une mesure inattendue dans le signal observé. Les défaillances des capteurs se produisent pour diverses raisons, telles que des défauts de fabrication, l'usure, une calibration incorrecte ou une mauvaise manipulation [Balaban et al., 2009].

Soit $r(t)$ le signal réel d'un phénomène physique (inconnu) et $m(t)$ le signal tel que mesuré par un capteur au temps t , on souhaite avoir [Balaban et al., 2009] :

$$m(t) = r(t) + \epsilon \quad (2.1)$$

Où ϵ est le bruit de mesures.

Dans un scénario idéal, le signal réel $r(t)$ est exactement le signal mesuré $m(t)$ [Balaban et al., 2009]. Également, quand ϵ est relativement petit par rapport à $r(t)$, cela indique que

le capteur est dans un état de fonctionnement sain. Cette situation est illustrée dans la figure 2.1, où le signal réel est représenté en vert et le signal mesuré en rouge (courbes superposées).

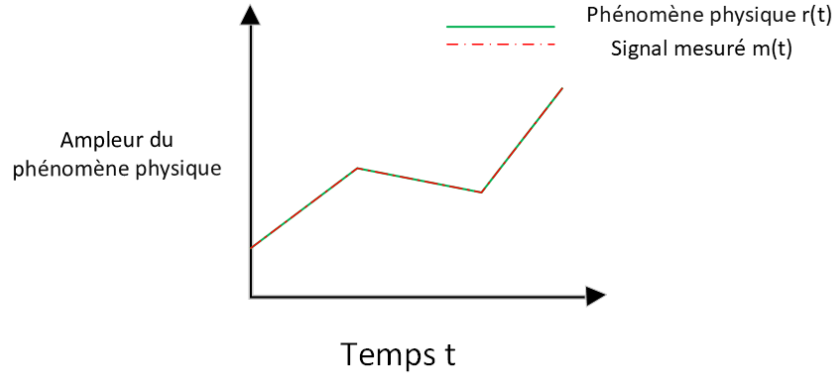


FIGURE 2.1 Signal d'un phénomène physique (inconnu) et signal mesuré par un capteur sain.

2.2.1 Absence de signal

Une d'absence de données se produit lorsque les données détectées ne contiennent pas de valeurs de séries chronologiques. L'équation de ce type de défaillance de capteur est la suivante [Noshad et al., 2019] :

$$m(t) = \emptyset \quad (2.2)$$

Où \emptyset est un ensemble vide ou avec des valeurs Not a Number (NaN).

Bien sûr, la présence du signal ne signifie pas toujours que le capteur est en bon état. Il existe d'autres états qui sont considérés comme des défaillances comme le biais, la saturation, le gel du capteur, le bruit, et aussi l'étincelle.

2.2.2 Défaillance de type biais

Une défaillance type biais est définie comme un décalage (erreur systématique) constant du signal mesuré du capteur [Balaban et al., 2009]. La figure 2.2 montre cette défaillance.

L'équation 2.3 représente ce type de défaillance de capteur [Balaban et al., 2009] :

$$m(t) = r(t) + \theta_0 \quad (2.3)$$

Où θ_0 est une valeur constante dans le temps.

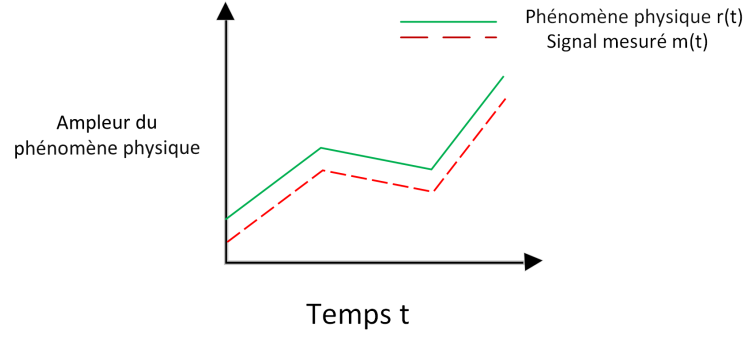


FIGURE 2.2 Défaillance de type biais du capteur

Un biais peut se produire en raison d'un étalonnage incorrect ou d'une altération [Balaban et al., 2009].

2.2.3 Défaillance de type saturation

La saturation se produit généralement lorsqu'un capteur mesure une valeur supérieure à son étendue de mesure [Dang and Suh, 2014]. Cette défaillance est montrée sur la Figure 2.3.

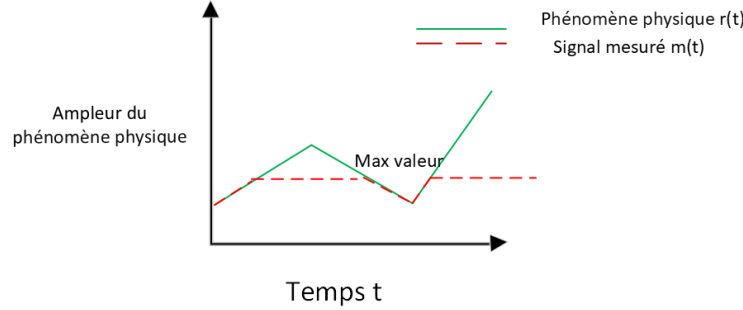


FIGURE 2.3 Défaillance de type saturation du capteur

L'équation 2.4 représente la défaillance dans la Figure 2.3 :

$$m(t) = \begin{cases} \text{Max valeur,} & \text{si } r(t) > \text{Max valeur} \\ r(t), & \text{si } \text{Min valeur} \leq r(t) \leq \text{Max valeur} \\ \text{Min valeur,} & \text{si } r(t) < \text{Min valeur} \end{cases} \quad (2.4)$$

Où «Max valeur» et «Min valeur» sont les limites supérieure et inférieure autorisées par le capteur avant que la saturation se produise [Dang and Suh, 2014]. Dans certains cas, une

saturation peut survenir si la puissance ou l'intensité du signal atteignant le capteur est trop élevée [Caruso et al., 2003].

2.2.4 Défaillance du signal de "type gel capteur"

Le "type gel capteur" fait référence à un type de défaillance du capteur lorsque la variation de la série de données mesurée par le capteur est constante sur une plage de temps. [Zidi et al., 2018]. L'équation générale de ce type de défaillance de capteur est montrée en 2.5 :

$$m(t) = \begin{cases} r(t), & \text{si } t < t_1 \\ \text{Valeur constante}, & \text{si } t_1 \leq t \leq t_2 \\ r(t), & \text{si } t > t_2 \end{cases} \quad (2.5)$$

Où "Valeur constante" représente la valeur "gelée" du signal mesuré pendant la période $t_1 \leq t \leq t_2$, t_1 est le moment où commence la période du gel, t_2 est le moment où se termine la période du gel.

La Figure 2.4 montre un exemple de signal avec ce type de défaillance.

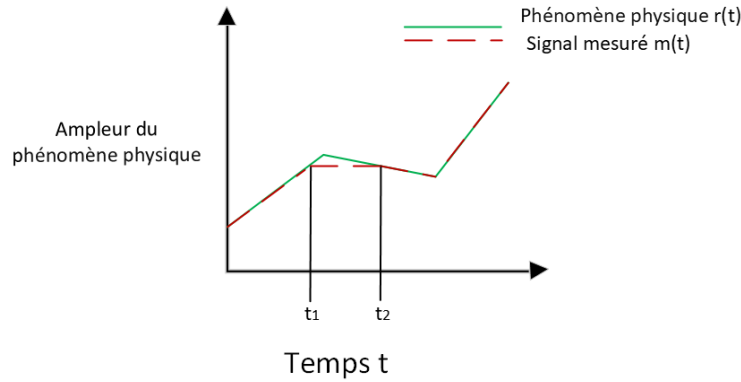


FIGURE 2.4 Défaillance du signal de "type gel capteur"

2.2.5 Défaillance de type bruit excessif

L'environnement dans lequel les capteurs fonctionnent peut être complexe et sensible à diverses sources de bruit. Le bruit interne provient principalement des caractéristiques des composants du capteur et du circuit, comme le bruit généré par les amplificateurs. Le bruit externe, quant à lui, provient d'interférences d'origine humaine ou environnementale situées en dehors du circuit du capteur. Bien que la présence de bruit soit courante dans les données

des capteurs, un niveau de bruit anormalement élevé peut entraîner des problèmes dans les signaux des capteurs. Le bruit excessif limite en effet les performances que l'on peut obtenir d'un dispositif donné [Bordoni and D'Amico, 1990]. La figure 2.5 est la représentation de cet type de défaillance.

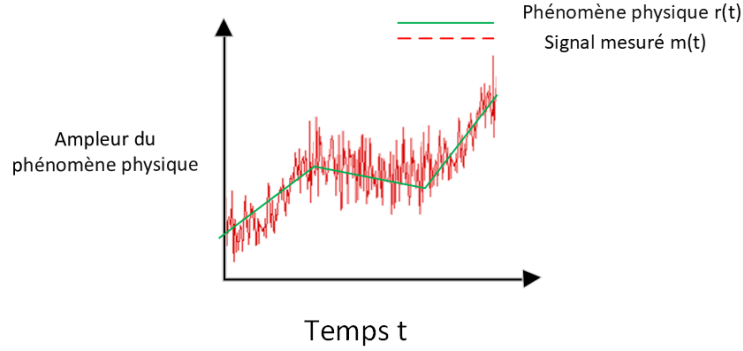


FIGURE 2.5 Défaillance de bruit du capteur

L'équation 2.6 montre la défaillance de bruit excessive du capteur :

$$m(t) = r(t) + \epsilon \quad (2.6)$$

avec ϵ est le bruit de mesure. Dans ce cas on considère que ϵ est du même ordre de grandeur que $r(t)$.

2.2.6 Défaillance de type étincelle

La défaillance de type "étincelle" dans un signal de capteur peut être modélisée comme une perturbation soudaine et brève dans la mesure du capteur. Cette perturbation est représentée mathématiquement en utilisant la fonction delta de Dirac, symbolisée par $\delta(t)$. La delta de Dirac est une fonction impulsionnelle qui est nulle partout sauf en un seul point.

La défaillance de type "étincelle" est donc modélisée par une telle impulsion dans le signal de mesure $m(t)$, au moment précis t_0 , et est généralement associée à une amplitude particulière (A) [Li and Li, 2013]. Cette modélisation est exprimée par l'équation 2.7.

$$m(t) = r(t) + A \cdot \delta(t - t_0) \quad (2.7)$$

Où :

- A représente l'amplitude de l'impulsion "étincelle".
- $\delta(t)$ est la fonction delta de Dirac.
- t_0 est le moment où se produit l'impulsion "étincelle".

Cette équation permet de modéliser une perturbation instantanée dans le signal du capteur, ce qui peut être utilisé pour décrire une défaillance de type "étincelle" dans la mesure du capteur. La figure 2.6 montre cette défaillance.

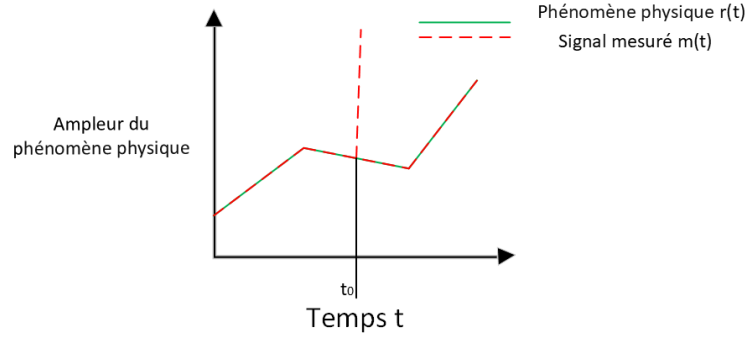


FIGURE 2.6 Défaillance de type étincelle du capteur

2.3 Techniques de détection des défaillances de capteurs

Dans cette section on présente les techniques de détection de défaillances de capteurs. Elle provient en grande partie de [Moreno Haro et al., 2023]. [Cook et al., 2020], ont proposé une classification des approches et des méthodes pour la détection défaillances.

2.3.1 Détection de défaillance type biais

Pour la détection de défaillances type biais [Zhao et al., 2016] proposent une méthode qui utilise l'Analyse en Composantes Principales (ACP). Dans la phase de modélisation, les sorties des capteurs sont collectées et soumises à une normalisation (moyenne nulle et variance unitaire). Ensuite, la technique de ACP est appliquée pour décomposer les données normalisées en composantes principales, obtenant ainsi les matrices de scores, ainsi qu'une matrice résiduelle contenant les erreurs non expliquées. Le nombre de composantes principales à retenir est déterminé en fonction de leur contribution cumulative (ex. jusqu'à 90% de la variance observée). Enfin, des statistiques de surveillance, en l'occurrence la projection dans l'espace d'erreurs, sont calculées pour détecter les défaillances de capteurs lorsque ces statistiques dépassent certaines limites.

2.3.2 Détection de défaillance de type "bruit"

[Kim et al., 2019b] ont proposé une modèle basée sur des auto-encodeurs variationnels (VAE) pour l'identification des défaillances dues à un excès de bruit dans les signaux. La méthodologie proposée pour détecter ces défaillances comprend deux étapes principales : le pré-entraînement du modèle VAE et l'identification des défaillances dans les signaux. Lors de la phase de pré-entraînement, le VAE est formé en utilisant des ensembles de signaux normaux dans le but d'apprendre des caractéristiques implicites de ces signaux et de pouvoir générer de manière réaliste des ensembles de signaux normaux dans le même espace. Une fois que le VAE est pré-entraîné, il est utilisé dans l'identification des défaillances des signaux. Le modèle reçoit des ensembles de signaux présentant des défaillances en tant qu'entrées et compare les sorties générées avec les entrées d'origine. L'objectif est de détecter les défaillances de type bruit à la fois entre l'entrée et la sortie. Pour valider l'applicabilité de cette méthode d'identification des défaillances, des séquences de défaillances ont été appliquées à 1 000 données unitaires sélectionnées au hasard avec des bruits artificiellement injectés. Si l'erreur de reconstruction d'un signal spécifique est supérieure à la valeur de multiplication de l'erreur de reconstruction moyenne de ce signal pendant le processus de pré-entraînement du modèle VAE et le multiplicateur prédéfini, alors le signal est identifié comme signal défectueux. Sinon, le signal est considéré comme signal normal.

2.3.3 Détection de défaillance de type saturation

La détection de saturation est essentielle pour identifier le moment où les signaux des capteurs ne fournissent plus de données fiables en raison de valeurs extrêmes. Selon [Reddy et al., 2020], dans les tests des auteurs, la saturation de l'application se produit lorsque les amplitudes du signal atteignent soit zéro, soit la tension de référence maximale du capteur. Dans leur système proposé, la détection de saturation est obtenue en surveillant l'amplitude du signal et en la comparant à des seuils prédéfinis, garantissant que le signal reste dans la plage de fonctionnement du capteur. Par exemple, dans un système de quantification 10 bits, les signaux sont considérés comme saturés si l'amplitude se situe entre 0 et 3 (saturation proche de zéro) ou 1020 et 1023 (saturation proche du maximum). Ces seuils aident à capturer non seulement la saturation complète, mais également les petites fluctuations proches des limites de la résolution du capteur, ce qui peut indiquer des problèmes potentiels dans l'acquisition du signal. Lorsque plusieurs échantillons consécutifs sont détectés dans ces plages, le système identifie le signal comme saturé et le rejette du traitement ultérieur. Pour améliorer encore la fiabilité du système, l'algorithme de détection peut classer les segments saturés en fonction de leur durée (par exemple, saturation de courte, moyenne ou longue durée), ce qui permet

une analyse plus détaillée de l'impact de la saturation sur le signal global. Cette méthode permet de garantir que les données erronées dues à la saturation du capteur n'affectent pas l'analyse globale et que les parties non saturées du signal peuvent toujours être traitées sans distorsion.

2.3.4 Détection de divers défaillances par un seul modèle

[Balaban et al., 2009] ont développé un modèle de détection de défaillances de capteurs en utilisant un réseau neuronal artificiel (ANN). Ce modèle vise à détecter plusieurs types de défaillances de capteurs, notamment le biais, et l'absence de signal. Les données de capteurs sont soumises à ces différents types de défaillances, puis les caractéristiques extraites de ces données sont utilisées comme entrées pour le réseau neuronal. Le réseau est composé de trois couches, y compris une couche d'entrée, une couche cachée et une couche de sortie. Les caractéristiques d'entrée sont standardisées et servent à former le réseau à l'aide de l'algorithme RPROP (La rétropropagation résiliente, est un outil d'apprentissage pour l'apprentissage supervisé dans les réseaux de neurones artificiels). Le réseau est ensuite testé pour évaluer sa capacité à détecter et à distinguer avec succès les défaillances de capteurs.

[Zhang et al., 2014] proposent un algorithme de détection de défaillances d'aiguillage basé sur un réseau de neurones probabiliste (PNN). L'aiguillage joue un rôle crucial dans le système ferroviaire, étant un équipement essentiel pour planifier l'itinéraire des trains et effectuer les changements de voie. L'algorithme présenté synthétise les courbes de courant d'action caractéristiques des défaillances d'aiguillage, utilisant le réseau neuronal PNN pour créer et évaluer des jeux de données cartographiant lesdites courbes. Les résultats expérimentaux révèlent que l'algorithme de détection des défaillances d'aiguillage basé sur le réseau de neurones PNN offre une précision plus élevée tout en nécessitant moins d'ajustements de paramètres, ce qui le rend facile à mettre en œuvre. De plus, il peut identifier une variété de problèmes potentiels, tels que des défaillances de type biais, de type gel, et aussi étincelles inattendus de courant et d'autres irrégularités dans les courbes de courant d'action, contribuant ainsi à la sécurité et à l'efficacité du système ferroviaire.

[Abid et al., 2015] proposent une technique de détection de défaillances sur les données des capteurs WSN utilisant l'algorithme KNN (K-Nearest Neighbor) en se basent sur la distance euclidienne comme critère. Dans cette technique, les données se présentent sous la forme d'une série chronologique de valeurs de données triées par ordre décroissant et les valeurs sont agrégées en groupes de K-Voisins. Pour chaque valeur, les distances euclidiennes entre son successeur et son prédécesseur sont calculées, et un indicateur est ajouté à la plus petite distance. La comparaison est faite entre les différentes valeurs de distance et

les valeurs qui produisent la plus grande distance sont éliminées (si elles sont identifiées comme des défaillances). Les auteurs se concentrent sur les défaillances de données atypiques, principalement les défaillances biais.

[Ehlenbröcker et al., 2016] examinent la détection des défaillances des capteurs à l'aide d'un algorithme qui évalue la cohérence entre les valeurs des capteurs, en utilisant la structure d'un algorithme de fusion de capteurs basé sur une classification non supervisée à plusieurs niveaux. Il se concentre sur la détection des défaillances des capteurs dans les systèmes de fusion d'informations multisensorielles, en appliquant une méthode pour calculer la fiabilité des capteurs et détecter les défaillances grâce à la cohérence des valeurs des capteurs. Les types de données utilisés incluent des signaux de capteurs et des caractéristiques extraites de ces signaux, traités par fuzzification pour les transformer en un espace unitaire sans dimension. La méthodologie de détection des défauts est basée sur le calcul dynamique et statique de la fiabilité des capteurs et sur la détection des incohérences entre les observations des capteurs. Les types de défauts détectés englobent des défauts complets de capteurs difficiles à identifier. Le domaine d'application est vaste, incluant la surveillance conditionnelle et la détection de défauts dans les processus industriels. La méthode utilisée combine des approches de la théorie des possibilités, de la fuzzification et des algorithmes d'apprentissage automatique pour la modélisation et la détection des défaillances de capteurs.

[Feng et al., 2019] présentent un algorithme de détection et de réconciliation des données de défauts de capteur (SFD&DR) utilisant quatre techniques de modélisation : filtre de Kalman robuste pour les valeurs aberrantes, moindres carrés partiels pondérés localement, méthode basée sur le sous-espace prédictif et moindres carrés récurrents basés sur le noyau avec un critère de dépendance linéaire approximatif. Cet algorithme analyse les résidus à l'aide de réseaux de neurones artificiels pour classer différentes conditions de signal du capteur. Les données utilisées proviennent d'expériences cliniques avec des systèmes de pancréas artificiel chez des personnes diabétiques, analysant 896 heures de données de surveillance continue de la glycémie dans 17 expériences. Cette approche peut détecter et diagnostiquer les défaillances des capteurs et réconcilier les signaux erronés des capteurs avec les valeurs estimées par le modèle avec une plus grande précision.

[Darvishi et al., 2020] présentent une approche non seulement pour la détection, mais aussi pour l'isolement et l'adaptation des défaillances des capteurs, cruciale pour l'intégration des données en temps réel dans les jumeaux numériques au sein de l'Industrie 4.0. L'architecture du modèle est basée sur l'apprentissage automatique, qui s'appuie sur des estimateurs de réseaux neuronaux et un classificateur pour valider les capteurs. Les auteurs ont réparti les capteurs en deux groupes : un ensemble non fiable contenant des capteurs susceptibles de

défaillance et un ensemble fiable composé de capteurs considérés comme fiables. La méthodologie repose sur trois étapes principales : la création de capteurs virtuels pour estimer les mesures des capteurs non fiables en excluant les données du capteur en question, le calcul des résidus qui mesure la dissimilarité entre les estimations des capteurs virtuels et les mesures réelles, et enfin, la classification et la prise de décision basée sur l'analyse des résidus pour détecter et isoler les défaillances des capteurs. Cette architecture utilise des réseaux de neurones perceptrons multicouches (MLP NN) pour les capteurs virtuels et le classificateur, optimisés par des algorithmes avancés et formés sur des ensembles de données réels. Il se concentre sur les défaillances tels que les biais et les dérives, en utilisant trois ensembles de données réelles pour valider l'architecture.

Pour la détection de défaillances de biais, étincelle et gelé des capteurs, [Jan et al., 2020] décrivent une méthodologie pour détecter ces défaillances dans les systèmes de moteurs d'avion. Cela implique la création d'un modèle, l'introduction de défaillances de capteurs simulés, la création d'un modèle de diagnostic de défaillance et la mise en œuvre d'un système de détection. L'approche proposée utilise un auto-encodeur (AE) pour extraire les caractéristiques du signal brut du capteur. Le vecteur de fonctionnalités est transmis par la suite à la machine à un algorithme de classification qui utilise les vecteurs de support (SVM) pour l'identification des défaillances. Un AE est un type de réseau neuronal, il se compose d'un encodeur suivi d'une réplique en miroir de l'encodeur appelée décodeur. Les nœuds d'entrée et de sortie (première couche d'encodeur et couche finale de décodeur) sont égaux. L'AE apprend les représentations du signal brut en essayant de générer une copie exacte du signal d'entrée au niveau de la couche de sortie.

[de Silva et al., 2021] présentent une approche hybride pour la détection des défaillances des capteurs, particulièrement appliquée dans le cadre des essais en vol. Ce modèle combine des techniques d'apprentissage automatique avec des algorithmes basés sur les données pour améliorer la détection des défaillances des capteurs dans les systèmes régis par des principes physiques sous-jacents. À l'aide de données de vols commerciaux, la méthode se concentre sur la détection de divers types de défaillances de capteurs, notamment les défaillances catastrophiques, les oscillations lentes, l'augmentation du bruit et les dérives lentes. La méthodologie proposée fonctionne en deux phases : une phase hors ligne pour entraîner le modèle, en utilisant la décomposition modale dynamique avec contrôle (DMDc) pour approximer l'évolution des mesures d'intérêt et entraîner un arbre de décision, et une phase en ligne pour la détection des défauts en temps réel, appliquant le modèle entraîné à de nouvelles données pour prédire les défaillances de capteurs sensorielles. Cette approche est efficace pour détecter avec précision les défaillances des capteurs dans les données d'essais en vols commerciaux, mais a également montré son applicabilité dans les modèles simulés d'aérodynamique instable et

de dynamique de vol longitudinal. La méthodologie est principalement appliquée dans le domaine des essais en vol et de la surveillance de la fiabilité des systèmes de capteurs dans l'industrie aérospatiale.

Un autre exemple est fourni par [Chen et al., 2022]. Ils utilisent une machine à vecteurs de support (SVM) comme classificateur pour catégoriser les caractéristiques provenant de plusieurs capteurs de moteurs d'avion. Le but est de déterminer le type de défaillance à partir des données enregistrées par les capteurs du moteur. SVM se distingue comme un type de machine d'apprentissage qui présente une forte capacité de généralisation, un apprentissage puissant et convient aux petits ensembles de données. Les auteurs disposent d'une base de données étiquetée avec 500 signaux identifiés par chaque type de défaillance. Parmi les défaillances existantes figurent les défaillances de type biais, dérive et étincelle. De plus, les signaux sans défaillances sont inclus dans la base de données. Ils effectuent un processus de formation modèle, prenant en compte les étiquettes et utilisent ensuite une base de données de tests pour la détection.

[Fong et al., 2023] présentent une approche pour la détection et le diagnostic des défaillances de capteurs dans les installations de refroidissement, en utilisant un algorithme hybride qui combine plusieurs techniques de détection de défaillances. Cette étude se distingue par sa méthodologie complète qui non seulement identifie les défaillances des capteurs, mais diagnostique et reconstruit également les informations affectées par ces défaillances. La méthodologie est basée sur la fusion des données de plusieurs capteurs et l'analyse de ces informations pour détecter des incohérences ou des anomalies qui indiqueraient une défaillance. Le modèle est appliqué à un ensemble de données comprenant des mesures de température, de pression et de débit au sein d'une installation de refroidissement, permettant une détection précise des défaillances des composants critiques du système. Les défaillances détectées sont le biais, la dérive des capteurs et les défaillances brusques. Le domaine d'application de cette étude se concentre sur les installations de refroidissement, un domaine critique des systèmes CVC (chauffage, ventilation et climatisation) dans les bâtiments commerciaux et industriels. Cette application est vitale pour la maintenance préventive et le fonctionnement efficace de ces systèmes, garantissant non seulement le confort mais également l'efficacité énergétique et la réduction des coûts d'exploitation. Concernant le type de méthode utilisée, le document propose une approche hybride qui intègre des techniques d'apprentissage automatique ainsi que des méthodes statistiques pour l'analyse des données des capteurs.

Il existe différentes approches pour extraire des informations à partir des signaux dans le but est de détecter des défaillances de capteurs. Parmi ces méthodes, on peut mentionner le calcul de paramètres statistiques, l'utilisation de filtres, les modèles d'autorégression, la

transformation de Fourier et la transformation en ondelettes. Parallèlement, les réseaux de neurones artificiels, qui sont des systèmes de traitement basés sur des neurones interagissant par de connexions numériquement pondérées, sont des outils efficaces pour anticiper les défaillances, à condition de disposer d'une quantité adéquate de données pour garantir un apprentissage "optimal".

2.4 Synthèse des méthodes selon les défaillances détectées

Le tableau 2.1 synthétise les travaux de recherche sur la détection des défaillances des capteurs. Il montre la méthode de détection et le domaine d'application en ce qui concerne les types de défaillances des capteurs. Il répertorie plusieurs références de recherche, chacune associée à un domaine spécifique, un type de défaillance de capteur et la méthode de détection utilisée. Par exemple, la référence [Zhao et al., 2016] se rapporte au domaine aéronautique, où la défaillance de type biais est détectée en utilisant une méthode statistique. De même, d'autres références sont répertoriées avec des combinaisons de domaines, de types de défaillance et de méthodes de détection.

TABLEAU 2.1 Relation entre la méthode et le domaine avec le type de défaillance des capteurs

Référence	AS	Sat	Biais	Gel	Bruit	Eti	Domaine	Méthode
[Balaban et al., 2009]	x		x		x		Électrique	AA
[Zhang et al., 2014]			x	x		x	Électrique	AA
[Abid et al., 2015]			x				IdO	AA
[Zhao et al., 2016]			x				Aéronautique	Statistique
[Ehlenbröker et al., 2016]	x						Manufacture	AA-Statistique
[Darvishi et al., 2020]			x				Industrie 4.0	AA
[Jan et al., 2020]			x	x		x	SCP	AA
[Feng et al., 2019]				x		x	Surveillance de la santé	AA-Statistique
[Kim et al., 2019b]					x		Énergie	AA
[Reddy et al., 2020]		x					Surveillance de la santé	Statistique
[de Silva et al., 2021]	x				x		Aéronautique	AA
[Chen et al., 2022]			x	x	x	x	Aéronautique	AA
[Fong et al., 2023]			x				Énergie	AA
Proposition	x	x	x	x	x	x	Aéronautique	AA

Où AS : Absence de signal, Sa : Saturation, Eti : Étincelle, et IdO : Internet des objets, SCP : Systèmes cyber-physique et AA : Apprentissage automatique.

Les domaines incluent des secteurs variés tels que l'aéronautique, l'Internet des objets (IdO), la surveillance de la santé, l'électrique, les systèmes cyber-physiques (SCP) et l'énergie. Chacun de ces domaines présente des enjeux particuliers en ce qui concerne la détection de défaillances de capteurs.

Le tableau identifie également différents types de défaillances de capteurs, parmi lesquels le étincelle, le biais, la saturation, et le bruit. Les méthodes de détection de défaillances varient considérablement en fonction des références. Certaines s'appuient sur des approches statistiques, tandis que d'autres adoptent des techniques d'apprentissage automatique (AA). L'utilisation croissante de l'apprentissage automatique reflète l'importance grandissante de cette discipline dans le domaine de la détection de défaillances de capteurs.

On note la diversité des approches dans le domaine de la détection des défaillances de capteurs, en tenant compte des spécificités de chaque domaine. Cependant, il est important de souligner qu'aucune des références examinées ne présente une méthode capable de couvrir l'ensemble des types de défaillances de capteurs (Absence de signal, biais, saturation, gelé, bruit et étincelle). Cette observation met en évidence le besoin continu de recherches approfondies pour développer des méthodes pour la détection de défaillances de capteurs.

CHAPITRE 3 PROBLÉMATIQUE

Dans ce projet, l'étude s'appuie sur l'analyse des données de certains capteurs embarqués dans des systèmes aéronautiques, à bord d'un avion de ligne civil.

Le modèle d'avion concerné est équipé de divers systèmes électromécaniques. Le projet actuel se concentre sur ces systèmes comme étude de cas pour le développement et la validation. Ultérieurement, les algorithmes et approches développés seront étendus à d'autres systèmes. L'objectif est de développer une méthode pour la détection d'anomalies dans de signaux de capteurs.

À ce jour, nous disposons d'une base de données contenant des séries chronologiques de capteurs d'un des systèmes embarqués. Les données sont stockées sous forme de series chronologiques. Les données collectées comprennent des informations sur des capteurs qui mesurent différents phénomènes physiques.

Dans de très rares cas, un capteur peut avoir produit un signal jugé anormal par rapport aux conditions réelles. Ce signal peut être un indicateur d'un dysfonctionnement à venir et que l'on souhaite prévenir dans le cadre d'une maintenance préventive. Il s'agit d'identifier le signal anormal en question dans un énorme volume de données qui possède une très grande variabilité interne et très peu de signaux identifiés comme anormaux.

Comme mentionné précédemment, nous disposons d'une base de données qui couvre certains numéros de événements, pour un système électromécanique embarqué qui contient different types des capteurs (certains capteurs sont redondants d'autres non). L'événement est le moment où le system et ses composants sont soumis aux plus fortes contraintes. Nous allons nous concentrer sur cette période. Pour cela, nous nous basons sur les données extraites sur une période de temps fixe autour d'un événement. Chaque événement étant unique, il présente donc une durée et des conditions particulières. Cela se traduit par une fenêtre d'analyse temporelle variable pour chaque événement, à l'intérieur d'une fenêtre d'extraction de données plus large et identique pour tous les événements.

Parmi ces événements, les experts du domaine ont identifié un petit nombre de situations anormales à étudier. Le nombre de situations identifiées comme anormales étant très faible et le temps d'analyse pour les identifier très long, il a été envisagé de développer une méthode automatique pour identifier ces événements anormaux et de les reoutner aux experts qui se chargeront in fine de l'analyse en profondeur avec leurs outils actuels.

Comme déjà mentionné, dans les systèmes aéronautiques, la fiabilité des capteurs joue un rôle

critique dans la sécurité des événements. Toutefois, les capteurs sont susceptibles de présenter une gamme variée de défaillances. Comme on observe dans l'état de l'art, il n'existe pas d'outils d'apprentissage automatique capables de détecter de manière automatique et précise l'ensemble de ces défaillances. Cette lacune constitue un enjeu pour notre partenaire industriel.

La problématique repose donc sur la conception d'une méthodologie pour aborder ce défi. Une telle méthodologie devra être en mesure de détecter les différents types de défaillances déjà mentionnés.

Comme il a été indiqué dans l'état de l'art, il n'existe actuellement pas d'outils d'apprentissage automatique permettant de détecter l'ensemble des défaillances (absence de signal, biais, saturation, capteur gelé, bruit, et étincelle) dans les capteurs.

CHAPITRE 4 MÉTHODOLOGIE PROPOSÉE

Dans ce chapitre, on présente la méthode proposée pour la détection des défaillances de capteurs. En premier lieu, nous traitons les données pour extraire les principaux descripteurs statistiques.

4.1 Analyse statistique des données

La base de données composée des valeurs de 77 capteurs ($i=1...77$) correspondant à plus de 8000 événements ($j=1...8000$). Les événements dans la base de données correspondent à une année. Chaque capteur i enregistre des informations à l'instant t_k . La fréquence d'échantillonnage f_c est liée à l'intervalle de temps Δt_k dans notre cas égal à 1 sec, par conséquent $f_c = \frac{1}{\Delta t_k} = \frac{1}{1 \text{ seg}} = 1 \text{ Hz}$. Les données sont extraites dans les fenêtres de la région d'intérêt de longueur ΔT_j , à noter qu'elles peuvent être différentes pour chaque événement j .

4.1.1 Nomenclature

Voici la nomenclature des différentes variables qui interviennent dans le calcul statistique.

i	numéro de capteur $i = 1, \dots, N_c$
j	numéro de événement $j = 1, \dots, N_v$
\mathbf{X}_{ij}	matrice contenant les informations du capteur i et du événement j
t_k	temps à l'incrément k , $k = 0, \dots, N_d - 1$
N_d	nombre de données d'un événement j dans la région d'intérêt
N_v	nombre de événements retenus pour l'analyse, $N_v = 8038$
N_c	nombre de capteurs, $N_c = 77$
ΔT_j	Longueur de la région d'intérêt du événement j
f_c	fréquence d'échantillonnage (=1 Hz)
\bar{x}_{ij}	moyennes des mesures d'un capteur i , correspondant à un événement j
$\text{VAR}(x_{ij})$	variance des mesures d'un capteur i , correspondant à un événement j
$S(x_{ij})$	coefficient d'aplatissement (obliques) du signal d'un capteur i , correspondant à un événement j
$K(x_{ij})$	kurtosis du signal du capteur i , correspondant à un événement j
$H(x_{ij})$	entropie du signal du capteur i , correspondant à un événement j
$p(x_{ij}(t_k))$	probabilité d'apparition de x_{ij} au temps t_k

$CORR(x_{nj}, x_{mj})$	corrélation de Pearson entre les capteurs m et n du événement j
$x_{mj}(t_k)$	valeur de la variable du capteur m dans le événement j à l'instant t_k
$x_{nj}(t_k)$	valeur de la variable du capteur n dans le événement j à l'instant t_k
\bar{x}_{mj}	moyenne des valeurs du capteur m du événement j
\bar{x}_{nj}	moyenne des valeurs du capteur n du événement j
N_d	nombre total de données

4.1.2 Descripteurs statistiques

On discute maintenant les méthodes de calcul et l'utilité des descripteurs statistiques (moyenne, variance, obliquité, kurtosis, entropie et corrélation.) dans le contexte spécifique de la détection de défaillances dans les capteurs. Ces derniers fournissent des signaux sous la forme de séries temporelles. Ces descripteurs nous permettent de résumer et d'identifier les tendances dominantes, et aussi d'identifier les comportements suspects. Par exemple, la moyenne et la variance offrent des pistes sur les conditions normales de fonctionnement, tandis que l'obliquité et le kurtosis révèlent des asymétries et des pointes (étincelles) dans les données. Quant à l'entropie, elle fournit une mesure du degré de désordre (ou d'incertitude) dans les données, elle peut s'avérer intéressante pour détecter des variations subtiles qui pourraient signaler une défaillance [Iglesias-Martínez et al., 2022]. Ensemble, ces descripteurs fournissent une piste potentielle pour réaliser la détection de défaillances dans les capteurs.

Chaque capteur i de chaque événement j mesure un phénomène physique, et il génère des données à l'instant t_k , représentées par le vecteur d'information $X_{ij} = \{x_{ij}(t_k)\}_{N_d \times 1}$.

La figure 4.1 nous montre ces caractéristiques statistiques.

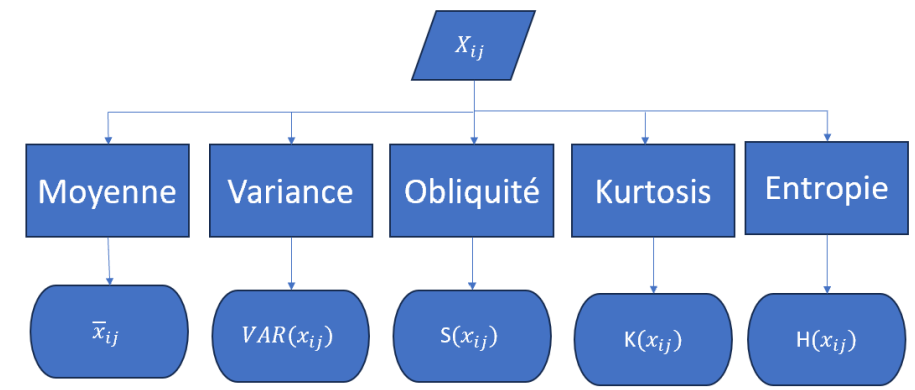


FIGURE 4.1 Descripteurs statistiques à partir de la série chronologique i de l'événement j

L'équation pour la moyenne des mesures d'un seul capteur i , correspondant à un seul événe-

ment j est présentée dans l'Éq. 4.1.

$$\bar{x}_{ij} = \frac{1}{N_d - 1} \sum_{k=0}^{N_d-1} x_{ij}(t_k) \quad (4.1)$$

L'équation pour la variance des mesures d'un seul capteur i , correspondant à un seul événement j est présentée dans l'Éq. 4.7.

$$\text{VAR}(x_{ij}) = \frac{1}{N_d - 2} \sum_{k=0}^{N_d-1} (x_{ij}(t_k) - \bar{x}_{ij})^2 \quad (4.2)$$

L'équation pour l'obliquité d'un capteur i , correspondant à un seul événement j est présentée dans l'Éq. 4.3.

$$S(x_{ij}) = \frac{\frac{1}{N_d-1} \sum_{k=0}^{N_d-1} (x_{ij}(t_k) - \bar{x}_{ij})^3}{\left(\frac{1}{N_d-1} \sum_{k=0}^{N_d-1} (x_{ij}(t_k) - \bar{x}_{ij})^2 \right)^{\frac{3}{2}}} \quad (4.3)$$

L'équation pour le kurtosis des mesures du capteur i et de l'événement j est présentée dans l'Éq. 4.4.

$$K(x_{ij}) = \frac{1}{N_d - 1} \sum_{k=0}^{N_d-1} \left(\frac{x_{ij}(t_k) - \bar{x}_{ij}}{\sqrt{\text{VAR}(x_{ij})}} \right)^4 \quad (4.4)$$

Finalement, l'équation pour estimer l'entropie du signal d'un capteur i pour le événement j est présentée dans l'Éq. 4.5.

$$H(x_{ij}) = - \sum_{k=0}^{N_d-1} p(x_{ij}(t_k)) \cdot \log(p(x_{ij}(t_k))) \quad (4.5)$$

4.1.3 Matrices des descripteurs statistiques

Après le calcul des descripteurs statistiques pour chaque capteur et chaque événement, il est possible de stocker ces informations dans des matrices. Ces matrices, de taille $N_c \times N_v$ (nombre de capteurs par nombre de événements), offrent une perspective pour l'examen des données. En analysant les colonnes de ces matrices, nous pouvons identifier des défaillances potentielles à certains événements, ce qui permet une évaluation ciblée de la performance ou de problèmes potentiels au cours des événements. De manière similaire, l'analyse des lignes permet de détecter des défaillances au niveau des capteurs, identifiant ainsi les capteurs qui pourraient présenter des défaillances ou des comportements anormaux sur plusieurs événements. Pour mieux comprendre on illustre ce processus dans la figure 4.2.

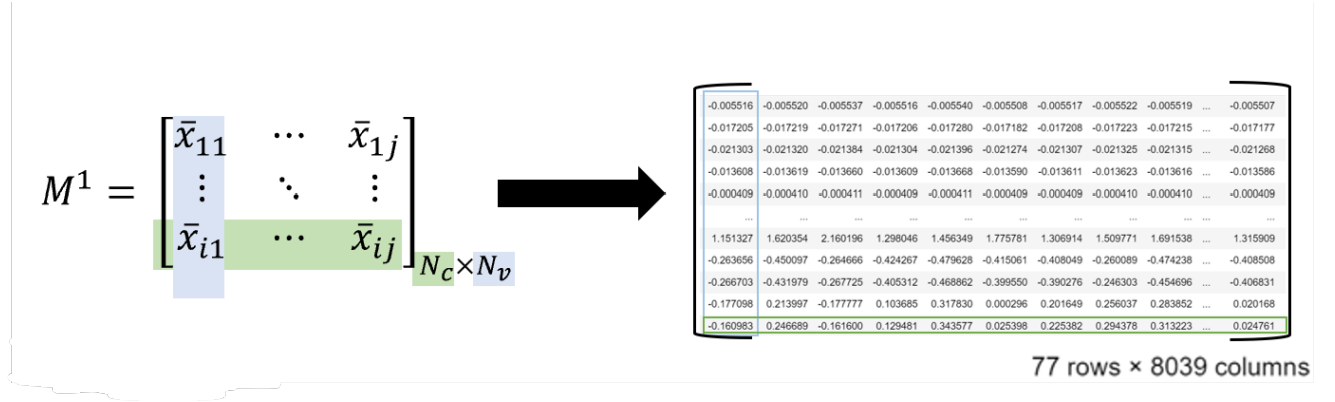


FIGURE 4.2 Construction d'une matrice de moyennes et sa visualisation à partir d'une base de données d'un système industriel réel

Les matrices résultantes contiennent les moyennes \bar{x}_{ij} , les variances $VAR(x_{ij})$, les obliques $S(x_{ij})$, les kurtosis $K(x_{ij})$ et les entropies $H(x_{ij})$ des signaux de chaque capteur i et pour chaque événement j . Ainsi, à partir du calcul de la moyenne des capteurs de la base de données, nous pouvons définir la matrice des moyennes qui est présentée dans l'Éq. 4.6.

$$\mathbf{M}^1 = \begin{bmatrix} \bar{x}_{11} & \cdots & \bar{x}_{1j} \\ \vdots & \ddots & \vdots \\ \bar{x}_{i1} & \cdots & \bar{x}_{ij} \end{bmatrix}_{N_c \times N_v} \quad (4.6)$$

L'Éq. 4.7 présente la matrice des variances.

$$\mathbf{M}^2 = \begin{bmatrix} VAR(x_{11}) & \cdots & VAR(x_{1j}) \\ \vdots & \ddots & \vdots \\ VAR(x_{i1}) & \cdots & VAR(x_{ij}) \end{bmatrix}_{N_c \times N_v} \quad (4.7)$$

L'Éq. 4.8 présente la matrice des obliquités.

$$\mathbf{M}^3 = \begin{bmatrix} S(x_{11}) & \cdots & S(x_{1j}) \\ \vdots & \ddots & \vdots \\ S(x_{i1}) & \cdots & S(x_{ij}) \end{bmatrix}_{N_c \times N_v} \quad (4.8)$$

L'Éq. 4.9 présente la matrice des kurtosis.

$$\mathbf{M}^4 = \begin{bmatrix} K(x_{11}) & \cdots & K(x_{1j}) \\ \vdots & \ddots & \vdots \\ K(x_{i1}) & \cdots & K(x_{ij}) \end{bmatrix}_{N_c \times N_v} \quad (4.9)$$

Et finalement, l'Éq. 4.10 présente la matrice des entropies.

$$\mathbf{M}^5 = \begin{bmatrix} H(x_{11}) & \cdots & H(x_{1j}) \\ \vdots & \ddots & \vdots \\ H(x_{i1}) & \cdots & H(x_{ij}) \end{bmatrix}_{N_c \times N_v} \quad (4.10)$$

4.1.4 Corrélation de Pearson

La redondance de plusieurs capteurs (ex. 4 capteurs de force, etc.) nous suggère d'examiner les corrélations entre eux. En effet, on s'attend à avoir de très fortes corrélations linéaires entre les capteurs redondants. Et si jamais, un de ces capteurs est défaillant, on s'attend à une plus faible corrélation. La formule pour calculer la corrélation de Pearson (corrélation linéaire) entre deux variables de séries temporelles (dans notre cas les variables X_{mj} et X_{nj} où j représente qu'elles appartiennent au même événement, et m et n sont les indices des capteurs), est la suivante :

$$CORR(x_{nj}, x_{mj}) = \frac{\sum_{k=0}^{N_d-1} (x_{mj}(t_k) - \bar{x}_{mj})(x_{nj}(t_k) - \bar{x}_{nj})}{\sqrt{\sum_{k=0}^{N_d-1} (x_{mj}(t_k) - \bar{x}_{mj})^2 \sum_{k=0}^{N_d-1} (x_{nj}(t_k) - \bar{x}_{nj})^2}} \quad (4.11)$$

4.2 Méthodologie proposée

On aborde le développement d'une méthode visant à détecter les changements de comportement d'un capteur. Cette stratégie repose sur la génération et l'analyse des résultats d'une base de données constituée d'images construites à partir de signaux de séries temporelles bruts. Dans le cadre du développement de la méthode visant à détecter les changements de comportement d'un capteur, l'utilisation d'images (qui contiennent les descripteurs statistiques et/ou les matrices de corrélation) prend une importance particulière. Ces images sont construites pour représenter la durée d'un événement (ex. un système aéronautique), capturant ainsi le comportement et l'évolution temporelle des signaux de séries temporelles provenant des capteurs. Chaque image représente ainsi un résumé de la séquence temporelle. Ces représentations sont ensuite exploitées pour détecter des anomalies, ou des altérations

dans le comportement du capteur.

Suivant une idée proche de la détection d'anomalies basée sur l'AE, l'approche VAE a également été utilisée pour résoudre des problèmes de détection d'anomalies [Sun et al., 2018]. La principale différence est que, alors que la détection des anomalies basée sur l'AE repose sur une erreur de reconstruction qui se comporte de manière prévisible, la détection des anomalies basée sur la VAE repose sur la probabilité de reconstruction, ce qui implique un certain degré d'incertitude. La notion de probabilité de reconstruction permet une approche probabiliste dans la détection des anomalies, qui s'avère plus efficace pour établir des critères de classification statistique visant à discerner si l'entrée est anormale ou non [An and Cho, 2015].

Cela est dû à deux caractéristiques dont il dispose le VAE pour détecter les anomalies :

- L'erreur de reconstruction du modèle.
- La position dans l'espace latent (qui pourrait indiquer le type d'anomalie).

Pour aller plus loin, le Variational Autoencoder utilise des architectures de réseaux de neurones convolutifs (CVAE) pour encoder et décoder les données de manière plus efficace.. Cette approche, associée à des réseaux de neurones convolutifs, permet d'incorporer des informations contextuelles, telles que l'état d'un capteur. En utilisant des architectures CNN, le CVAE améliore la représentation latente et la détection des anomalies en tenant compte des facteurs externes qui pourraient influencer le comportement du capteur, conduisant ainsi à une meilleure distinction entre les comportements normaux et anormaux.

Erreur de reconstruction du CVAE

La figure 4.3 représente le processus de reconstruction d'une image qui passe par l'encodeur et par la suite par le décodeur.

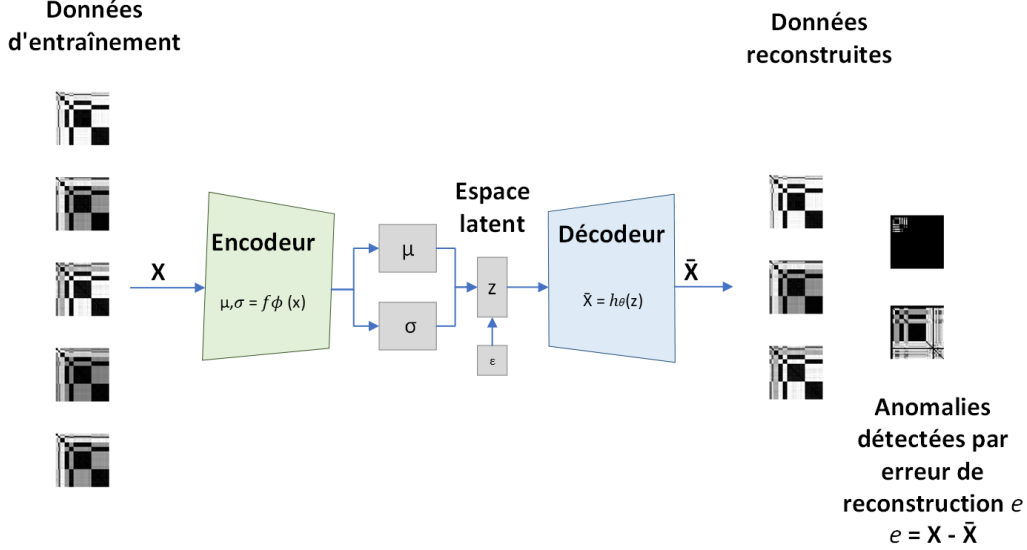


FIGURE 4.3 Schéma de reconstruction et erreur de reconstruction e du modèle CVAE proposé

Sur la figure, on peut observer comment le CVAE capture et compresse les caractéristiques importantes des images d'entrée. Les différences entre les données d'entraînement et les données reconstruites mettent en évidence la capacité à bien reconstruire les images que le modèle utilise. L'erreur de reconstruction souligne la capacité du CVAE à apprendre des motifs significatifs dans les données, ce qui est crucial pour son application dans des tâches telles que la détection d'anomalies ou la génération de nouvelles représentations visuelles. L'erreur de reconstruction est utilisée pour évaluer dans quelle mesure le modèle peut reproduire les données originales à partir de leur représentation dans l'espace latent [Kim et al., 2019a]. L'erreur de reconstruction (e) dépend alors des données reconstruites ($\bar{\mathbf{X}}$) par le CVAE et les originales (\mathbf{X}) tel qu'illustré dans l'équation 4.12.

$$\mathbf{e} = \mathbf{X} - \bar{\mathbf{X}} \quad (4.12)$$

L'encodeur est un réseau neuronal avec des paramètres ϕ qui place les éléments de \mathbf{X} dans l'espace latent avec une moyenne $\mu \in \mathbb{R}^{n_L}$ et un écart-type $\sigma \in \mathbb{R}^{n_L}$, comme indiqué dans l'équation 4.13, où $n_L \ll n_F$, n_L est la dimension de l'espace latent, n_F est la dimension des données d'entrée. Où ϕ sont les paramètres de l'encodeur du réseau neuronal, μ est la moyenne de l'espace latent, σ est l'écart-type de l'espace latent.

$$f_\phi : \mathbf{X} \mapsto \{\mu, \sigma\}, \mathbb{R}^{n_F} \rightarrow \mathbb{R}^{n_L} \times \mathbb{R}^{n_L} \quad (4.13)$$

La méthode de reparamétrage introduit une estimation bayésienne variationnelle de l'espace latent. Elle utilise les valeurs μ et σ dans la variable latente $\mathbf{z} \in \mathbb{R}^{n_L}$ selon l'Équation 4.14, où $\epsilon \sim \mathcal{N}(0, 1)$ est un vecteur gaussien de dimension n_L avec un espace d'échantillonnage \mathcal{E} , et \odot est le produit élément par élément. \mathbf{z} est la variable qui représente l'espace latent, ϵ est un vecteur gaussien utilisé pour la reparamétrisation, et $\mathcal{N}(0, 1)$ est la distribution normale standard.

$$g : \{\mu, \sigma, \epsilon\} \mapsto \mathbf{z} = \mu + \sigma \odot \epsilon, \mathbb{R}^{n_L} \times \mathbb{R}^{n_L} \times \mathcal{E} \rightarrow \mathbb{R}^{n_L}. \quad (4.14)$$

Enfin, le décodeur h_θ du CVAE est un réseau neuronal profond (DNN en anglais) qui reproduit l'espace latent \mathbf{z} dans la sortie CVAE $\bar{\mathbf{X}}$, comme indiqué dans l'équation 4.15, où θ est l'ensemble des paramètres du décodeur, et g fait référence au processus de reparamétrisation dans le CVAE [Oliveira-Filho et al., 2023].

$$h_\theta : \mathbf{z} \mapsto \bar{\mathbf{X}}, \mathbb{R}^{n_L} \rightarrow \mathbb{R}^{n_F}. \quad (4.15)$$

Lors de la détection d'anomalies, un seuil est défini pour l'erreur de reconstruction. Les données qui produisent une erreur de reconstruction supérieure à ce seuil sont considérées comme anormales. En effet, si le modèle a du mal à reconstruire avec précision ces données, ceci suggère qu'elles sont inhabituelles par rapport aux instances normales de l'ensemble de données.

 Algorithme de détection d'anomalies basé sur un autoencodeur variationnel

Entrée :Jeu de données X_n Données considérées comme anormales $\{x_{a,i} | i = 1, \dots, I\} \subseteq X_a = 12$ Seuil de probabilité de reconstruction α **Sortie :**

Résultats de détection d'anomalies (classification) pour toutes les données anormales

Entraînement de l'autoencodeur variationnel : $f_{enc}, f_{dec} \leftarrow$ Entraîner un autoencodeur variationnel en utilisant X_n **for** $i = 1$ à I **do** $\mu_{z,i}, \sigma_{z,i} \leftarrow f_{enc}(z|x_{a,i})$ **for** $k = 1$ à K **do**Tirer K échantillons de z à partir de $N(\mu_i, \sigma_i)$ $\mu_{x,i,k}, \sigma_{x,i,k} \leftarrow f_{dec}(x|z_{i,k})$ **end for**

Calculer la probabilité de reconstruction :

 $P(i) = \frac{1}{K} \sum_{k=1}^K p(x_{a,i} | \mu_{x,i,k}, \sigma_{x,i,k})$ **if** $P(i) \leq \alpha$ **then** $x_{a,i}$ est une anomalie**else** $x_{a,i}$ n'est pas une anomalie**end if****end for**

Représentation dans l'espace latent

L'espace latent dans un CVAE est donc une représentation de dimension réduite des données d'entrée. Il s'agit d'un espace abstrait où le modèle cherche à coder les informations les plus importantes et les plus pertinentes contenues dans les données originales. La principale raison de l'utilisation d'un espace latent est cette compression des données et l'extraction de fonctionnalités à partir de cet espace.

Plus spécifiquement, dans notre cas, lorsqu'une image est différente (anormal), elle sera située dans une région inhabituelle dans l'espace latent. En effet, le CVAE a appris à représenter des données normales dans des régions densément peuplées de l'espace latent. Si un point de données est situé dans une région peu peuplée ou éloignée de l'espace latent, il est considéré comme anormal [Marimont and Tarroni, 2021].

La capacité de CVAE à apprendre des représentations des données le rend utile dans la détection d'anomalies d'images. Pendant la phase de formation, le CVAE reçoit un ensemble d'images et il apprend à les représenter dans son espace latent et à les reconstruire [Marimont and Tarroni, 2021]. La capacité du CVAE à apprendre des représentations efficaces et compactes des données le rend utile dans la détection d'anomalies d'images. Pendant la phase de

formation, il reçoit un ensemble d'images considérées comme normales ou typiques du système. Par conséquent, le CVAE apprend à les représenter dans son espace latent (comprimé) et à les reconstruire avec précision.

Ainsi, la position géographique dans l'espace latent et l'erreur de reconstruction du VAE représentent une excellente option pour atteindre l'objectif spécifique OS1 de notre projet. L'espace latent permet une représentation commode des images statistiques des signaux d'un événement, l'erreur de reconstruction met en évidence les différences entre les données d'entrée et leur reconstruction. Ces aspects font du CVAE un choix pertinent pour la détection de changements de comportement des capteurs.

4.3 Étapes proposées

Quatre (4) étapes séquentielles sont proposées pour la méthodologie. La dernière étape (#4) repose sur deux branches pour effectuer la détection des défaillances. Dans la première branche, la détection est réalisée en utilisant l'erreur de reconstruction, à la fois dans les événements normaux et dans les événements avec des capteurs défectueux. Dans la deuxième branche, la détection des événements avec des capteurs anormaux est réalisée grâce à l'utilisation de l'espace latent et une métrique basée sur la distance euclidienne des événements anormaux vers le centroïde du groupe de événements normaux. La méthodologie proposée est présentée dans la Figure 4.4.

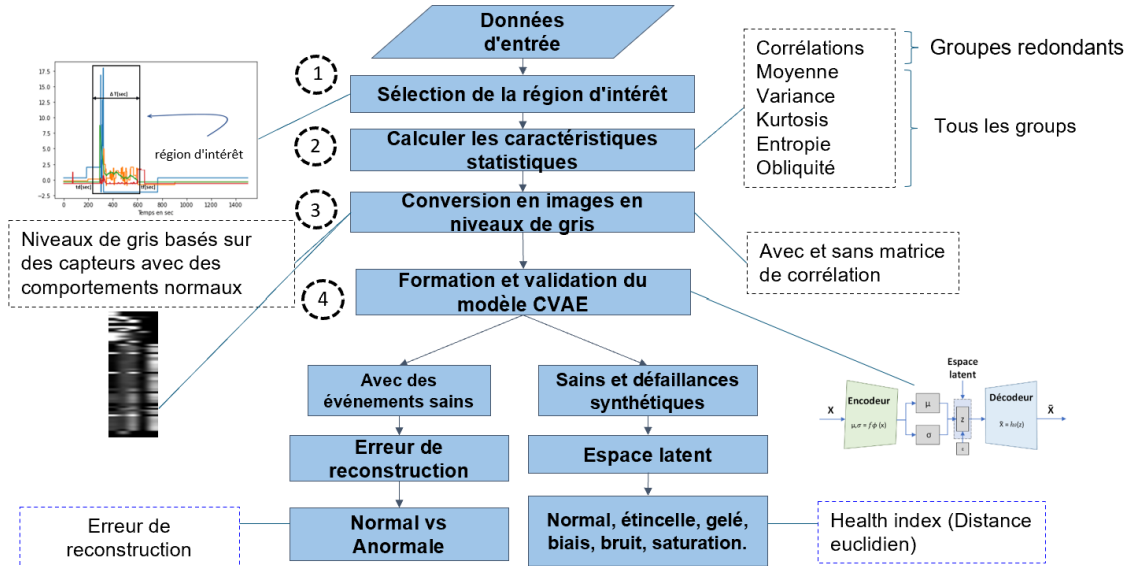


FIGURE 4.4 Organigramme illustrant les étapes de la méthodologie proposée pour la détection des défaillances de capteurs dans les systèmes industriels

4.3.1 Étape 1 : Identifier et extraire la région d'intérêt par événements

Détecter la zone d'intérêt où varient les signaux des séries chronologiques est important pour fournir uniquement les informations utiles au modèle à développer. Nous disposons de données détaillées sur l'évolution des mesures pendant un événement commercial. La zone d'intérêt de ce projet correspond à la portion du temps où les signaux varient pendant un processus bien défini, excluant certains moments de la séquence. On commence par détecter et isoler cette région d'intérêt. Pour y arriver, on identifie les instants clés du début du processus et celui de la stabilisation. Cette fenêtre de temps représente la région d'intérêt et est différente pour chaque événement, détectée par les signaux mesurés au cours de l'analyse. La figure 4.5 montre différents signaux. Un signal en **bleu**, **vert**, **orange** et **rouge**. Ce sont des signaux des capteurs qui sont anonymisés pour des raisons de confidentialité. L'un de ces signaux est utilisé pour réaliser le processus de détection de la région d'intérêt pour chaque événement. Le choix de ce signal particulier pour la détection de la région d'intérêt s'explique par un pré-analyse qui montre que les signaux des capteurs présentent une activité variée élevée dans cette fenêtre temporelle.

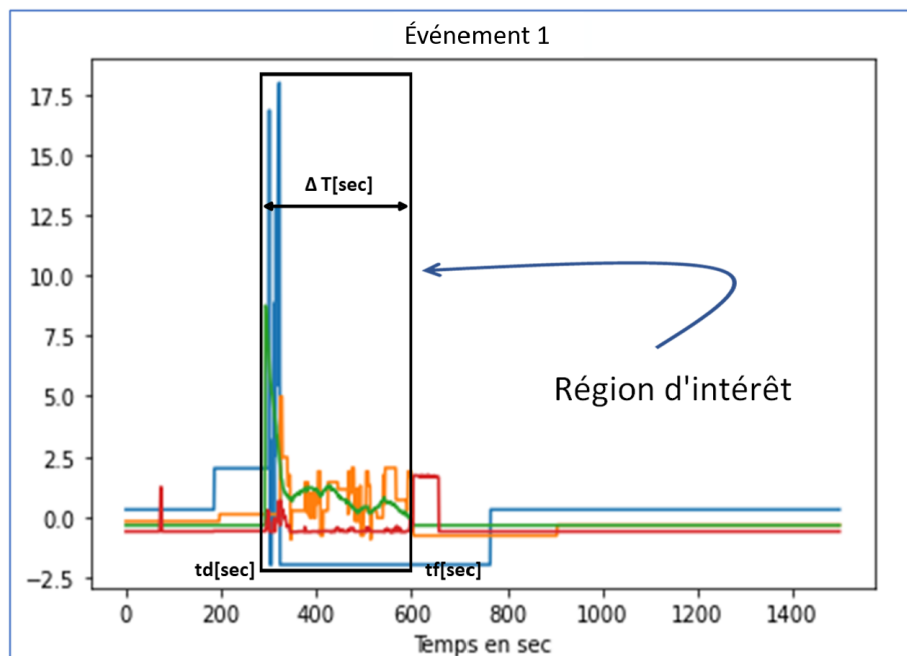


FIGURE 4.5 Exemple de détection de la région d'intérêt d'un événement $j = 1$ avec des séries temporelles de i capteurs

L'équation 4.16 fournit le temps de durée de la région d'intérêt par événement (j).

$$\Delta T_j = t_f - t_d \quad (4.16)$$

Où t_d est le temps début de la région d'intérêt et t_f est le temps de la fin. On a donc comme résultat la durée de la région d'intérêt $\Delta T[sec]$ qui représente la période propre de l'événement. Il est important de souligner que les fenêtres d'intérêt varient d'un événement à un autre. La figure 4.6 illustre un exemple de mesure pendant cette période de temps. Nous pouvons apprécier la figure correspondant au signal détecté dans la région d'intérêt, avec une valeur ΔT de 300 secondes qui représente la longueur de la région d'intérêt.

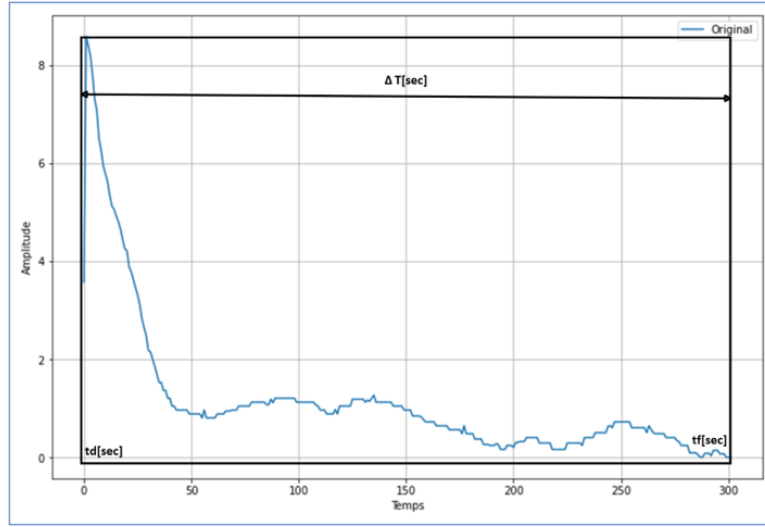


FIGURE 4.6 Signal d'un capteur i utilisé pour effectuer le processus de détection de la région d'intérêt

4.3.2 Étape 2 : Création de base de données des images

Pour atteindre l'objectif spécifique OS1, nous avons utilisé des descripteurs statistiques tels que la moyenne, la variance, l'obliquité (skewness), le kurtosis, l'entropie, et la corrélation de Pearson. L'utilisation de ces descripteurs est pertinente dans un contexte où des capteurs redondants sont présents, car ils permettent d'extraire des caractéristiques distinctives des signaux enregistrés par ces capteurs.

Nous avons choisi de représenter les descripteurs sous forme d'images en échelle de gris pour faciliter la visualisation des variations entre les capteurs. L'utilisation de cette échelle de gris se justifie par sa simplicité et son efficacité à révéler les différences d'intensité des descripteurs calculés pour chaque capteur. Les nuances de gris permettent ainsi d'illustrer clairement les différences de comportement entre les capteurs et leurs signaux respectifs,

rendant les variations plus visibles et plus aisément interprétables.

La taille des images est un paramètre flexible, et elle peut être configurée en fonction du nombre de capteurs et des descripteurs considérés. Chaque image correspond donc à un évènement donné, et chaque pixel de l'image représente la valeur normalisée d'un descripteur pour un capteur spécifique. Par exemple, dans la figure 4.7, on observe différents capteurs qui font partie d'un système lié à un évènement. Cette figure montre également des lignes représentant les connexions entre les capteurs, ainsi qu'une matrice de corrélation qui a été établie après l'évènement. La corrélation entre les signaux des capteurs est visualisée sous forme de pixels dans une échelle de gris. Chaque pixel représente la corrélation entre une paire de capteurs, allant du noir (corrélation parfaite, $|r| = 1$) au blanc (absence de corrélation, $|r| = 0$). Les zones sombres indiquent une corrélation forte, tandis que les zones claires signalent des écarts plus significatifs entre les capteurs.

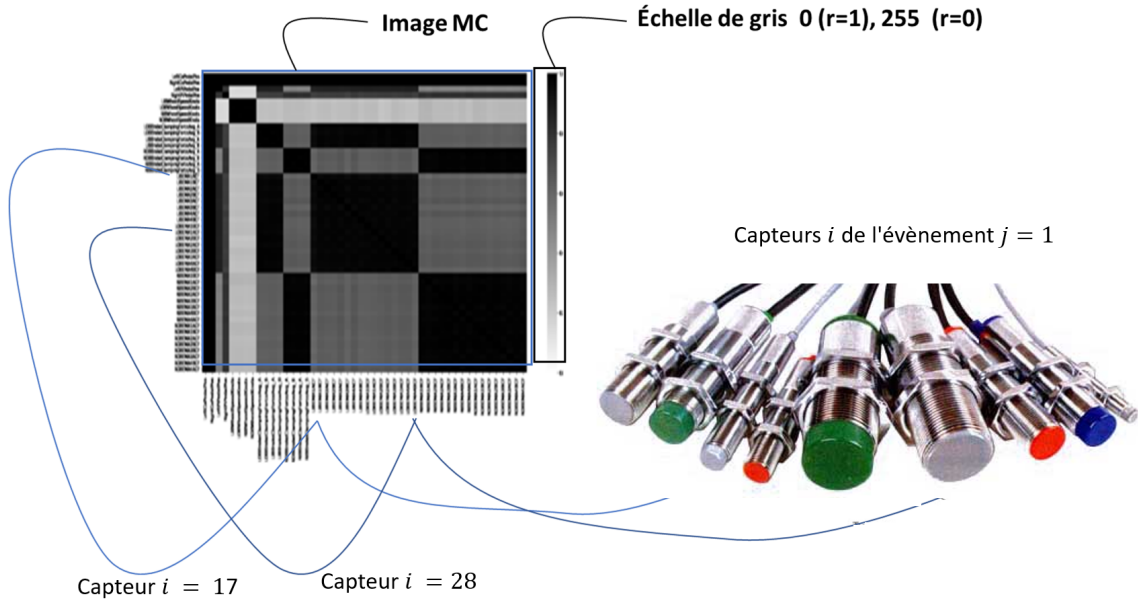


FIGURE 4.7 Relation entre les capteurs d'un évènement et l'image avec pixels de corrélation

À ce stade, nous avons principalement utilisé les pixels de corrélation pour représenter les relations entre les capteurs. Toutefois, nous prévoyons d'intégrer d'autres descripteurs, tels que la kurtosis, la médiane et la déviation standard, afin d'enrichir chaque image avec des informations plus complètes. Cela permettra de capturer non seulement les corrélations entre les capteurs, mais aussi d'autres aspects statistiques de leurs signaux.

La taille des images sera ajustée en fonction du nombre de descripteurs k et du nombre de capteurs n . Ainsi, une image résultante aura une taille de $kn \times kn$, permettant une représenta-

tion plus détaillée des données et des relations complexes entre les capteurs. L'augmentation du nombre de descripteurs et de capteurs dans chaque image permettra de visualiser des variations plus subtiles et des interactions complexes au sein des systèmes de capteurs.

Le processus de création de la base de données d'images repose sur deux étapes clés : la normalisation et la conversion en valeurs de pixels. Pour chaque événement, nous convertissons les valeurs de chaque descripteur en niveaux de gris, en normalisant ces valeurs afin de garantir une échelle cohérente entre les différents événements.

La normalisation est effectuée sur tous les événements afin d'assurer une cohérence dans l'échelle des descripteurs, permettant ainsi des comparaisons significatives entre différents événements. La formule de normalisation est donnée dans 4.17 comme suit :

$$y_{ij} = \frac{x_{ij} - \min(\mathbf{X}_{ij})}{\max(\mathbf{X}_{ij}) - \min(\mathbf{X}_{ij})} \quad (4.17)$$

Où y_{ij} représente la valeur normalisée des mesures x_{ij} du capteur i pour l'événement j , et \mathbf{X}_{ij} est la matrice contenant les informations du capteur i et de l'événement j .

La valeur normalisée est ensuite multipliée par 255 pour la convertir en une intensité de pixel, où 0 correspond à noir (la valeur minimale) et 255 correspond à blanc (la valeur maximale). Si les valeurs minimale et maximale sont égales, le pixel est défini sur 0 (noir). La valeur de pixel pour les mesures est donnée dans 4.18 comme suit :

$$P_{ij} = \frac{x_{ij} - \min(\mathbf{X}_{ij})}{\max(\mathbf{X}_{ij}) - \min(\mathbf{X}_{ij})} \times 255 \quad (4.18)$$

Où P_{ij} représente la valeur de pixel pour les mesures x_{ij} du capteur i pour l'événement j . La normalisation garantit que les valeurs des pixels représentent des différences relatives entre tous les événements, et pas seulement au sein d'un seul événement, offrant ainsi une vue plus complète des données.

Une fois que toutes les images sont créées, elles sont regroupées dans une base de données d'images, permettant ainsi une analyse ultérieure en vue de la détection des défaillances. La base de données contient ainsi une collection d'images représentant différents événements, avec des niveaux de gris qui reflètent la magnitude des descripteurs statistiques pour chaque capteur. Cela nous fournit une représentation visuelle intuitive et riche des données, facilitant la détection des anomalies ou des tendances spécifiques dans le comportement des capteurs au cours du temps.

4.3.3 Étape 3 : Création du modèle d'apprentissage automatique

Nous avons décidé d'implémenter un Autoencodeur variationnel avec filtres de réseau neuronal convolutif (CVAE). Ce modèle a été retenu en raison de son adéquation à l'apprentissage automatique avec des données semi-supervisées sous forme des images et de sa large application dans la détection de défauts dans différents types de capteurs dans différents domaines.

Pour mieux comprendre le sujet, commençons par examiner les différences entre un Variational Autoencoder (VAE), qui constitue la base du CVAE, et un Autoencoder (AE). L'auto-encodeur (AE) représente un type d'architecture de réseau neuronal formé par des processus d'apprentissage non supervisés ou semi-supervisés. Essentiellement, la formation d'un EA est orienté vers l'obtention d'une entrée et d'une sortie présentant la plus grande similitude possible, tout en incorporant des couches internes avec un nombre de nœuds inférieur aux dimensions des données d'origine [Qin et al., 2021].

L'AE se compose de deux éléments fondamentaux : l'encodeur et le décodeur. La configuration de l'encodeur se caractérise par sa nature étroite, ce qui signifie que le nombre de nœuds dans la couche suivante est inférieur à celui de la couche précédente. En revanche, la conception du décodeur est de nature large, ce qui implique que le nombre de nœuds dans la dernière couche est supérieur à celui de la couche précédente. Grâce à ces attributs structurels du codeur et du décodeur, le codeur réduit la dimension des données d'origine à une taille plus compacte, tandis que le décodeur est chargé de restaurer les données d'entrée d'origine sur la base de la représentation condensée obtenue à partir de l'encodeur [Ballard, 1987].

Comme son nom l'indique, l'auto-encodeur variationnel (VAE) a une structure similaire à celle de l'AE, bien que ses propriétés soient un peu différentes. Comme les AE, les VAE se composent d'un encodeur et d'un décodeur. Cependant, l'encodeur d'un VAE déduit les paramètres de la distribution de probabilité à partir des entrées du décodeur plutôt que de déduire directement les entrées du décodeur. En d'autres termes, l'entrée du décodeur d'un VAE devient une variable aléatoire avec une distribution de probabilité continue. En conséquence, alors que le décodeur d'un AE reçoit toujours la même entrée (déterministe) de l'encodeur, qui correspond à l'entrée d'origine du modèle dans son ensemble, le décodeur d'un VAE reçoit plusieurs entrées (dans un sens probabiliste) bien que l'entrée d'origine du modèle soit identique [Kim et al., 2019b].

Pour générer des entrées, c'est-à-dire des vecteurs latents aléatoires, à partir des paramètres de la distribution déduite, une étape d'échantillonnage est incorporée entre le encodeur et le décodeur. Grâce à la nature stochastique du VAE, il est non seulement utilisée pour reproduire des entrées, mais également pour générer des ensembles de données complètement

nouveaux présentant des similitudes avec les entrées, tout comme d'autres modèles génératifs. En tant que type de modèle génératif, la VAE a été appliquée dans divers domaines, tels que l'augmentation des données et le traitement du langage naturel [Gunduz, 2021].

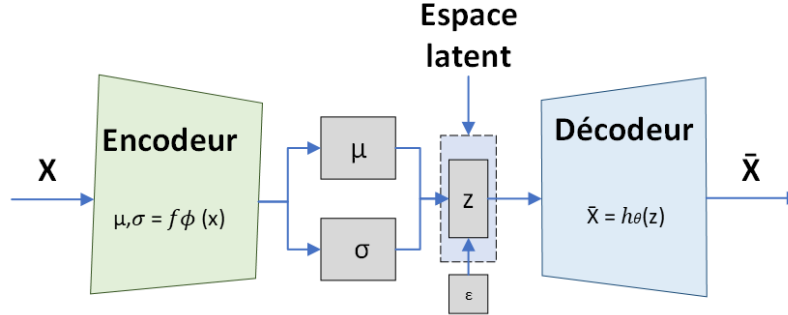


FIGURE 4.8 Structure de l'auto-encodeur variationnel

La figure 4.8 montre la structure générale d'un auto-encodeur variationnel. Où z représente le vecteur latent résultant, généralement utilisé pour encoder l'information ou la représentation latente d'une donnée, μ représente la moyenne de la distribution normale, σ est l'écart-type, et ϵ est un terme d'erreur aléatoire. Ainsi, μ contrôle la position centrale de la distribution (c'est-à-dire, la moyenne), et σ contrôle la dispersion ou la variabilité (c'est-à-dire, l'écart-type) de la distribution normale multivariée. L'équation 4.19 complète est :

$$z = \mu + \sigma \cdot \epsilon \quad (4.19)$$

4.3.4 Étape 4 : Apprentissage du modèle

L'objectif principal de la formation est d'optimiser les paramètres du modèle de telle manière qu'il soit capable de reconstruire efficacement fidèlement les images d'entrée. Pendant l'entraînement, le modèle cherche à minimiser la différence entre les images d'entrée et les images générées par le décodeur.

L'entraînement peut être plus au moins chronophage en fonction de la taille de la base de données d'images et de la complexité du modèle. Une fois l'entraînement terminé, le modèle est prêt à être utilisé pour la phase de test.

Le processus d'apprentissage suit une approche itérative basée sur la rétropropagation de l'erreur, où les poids des couches du modèle sont ajustés à chaque itération afin de minimiser la fonction de perte. Dans le cas du CVAE, la fonction de perte totale est composée de l'erreur de reconstruction, souvent mesurée par une métrique telle que l'erreur quadratique moyenne

(MSE), et du terme KL qui contrôle la proximité de la distribution latente par rapport à une distribution normale standard. L'équation de la perte peut s'exprimer comme suit :

$$\mathcal{L}_{total} = \mathcal{L}_{reconstruction} + \beta \cdot \mathcal{L}_{KL} \quad (4.20)$$

Où, $\mathcal{L}_{reconstruction}$ correspond à l'erreur entre les images d'entrée et les images reconstruites, et \mathcal{L}_{KL} est le terme de régularisation KL qui permet de maintenir une distribution latente gaussienne standard. Le facteur β est un hyperparamètre qui permet de contrôler l'importance relative de ces deux termes.

L'entraînement du modèle peut être plus ou moins chronophage (demande beaucoup de temps) en fonction de plusieurs facteurs, notamment la taille de la base de données d'images, la complexité du modèle, ainsi que la taille des images. La base de données d'images étant constituée de signaux capteurs transformés en images de différentes tailles, chaque image représente des informations essentielles pour la détection de défaillances. Le modèle CVAE apprend à identifier les schémas dans ces images et à produire des reconstructions précises dans la phase de décodeur.

Une fois l'entraînement terminé, le modèle est prêt à être évalué sur des données de test. Cette phase permet de valider les performances du modèle, notamment en termes de capacité à reconstruire correctement des images provenant de événements normaux et à générer des représentations latentes cohérentes. En outre, dans le cadre de ce projet, l'utilisation de l'espace latent pour la détection des anomalies constitue un aspect crucial. Les événements présentant des comportements anormaux dans leurs signaux de capteurs devraient se positionner à l'extérieur de la distribution latente des événements normaux, facilitant ainsi la détection de défaillances via une analyse de la distance euclidienne dans cet espace latent.

Processus de création de défaillances synthétiques

Pour créer des défaillances synthétiques, les informations et les équations de l'état de l'art, qui sont présentées dans le chapitre 2 de ce travail, sont utilisées. Parmi les défaillances nous avons :

- Étincelle
- Gelé
- Biais
- Bruit
- Saturation

.

Pour la création des défaillances synthétiques, nous prenons la théorie des défaillances étudiée dans l'état de l'art et utilisons la méthodologie suivante, présentée ci-dessous :

1. Tout d'abord, la base de données est divisée en groupes de manière équilibrée entre le nombre d'états d'un capteur, c'est-à-dire qu'il faut avoir la même quantité de données pour chacune des défaillances synthétiques. Chacun de ces états représente une défaillance synthétique en plus de l'état sain. Ainsi, la base de données est divisée de manière aléatoire en fonction du nombre d'états.
2. Les défaillances synthétiques qui seront créées sont celles mentionnées précédemment dans le chapitre 2. Ensuite, on prend chacun des groupes de données sains de la base de données originale et on commence à injecter à chacun de ces groupes les différents types de défaillances. On laisse un de ces ensembles sans altération, car nous avons également besoin de l'état sain pour notre modèle. Nous aurons alors un groupe avec un état de défaillance de type étincelle, un autre groupe avec une défaillance de type gelé, un autre groupe avec une défaillance de type saturation, etc.
3. À partir de cela, nous prenons les différentes équations des défaillances et les injectons dans un signal sain. Par exemple, pour le signal défectueux avec une erreur de type biais, on ajoute une valeur de biais au signal sain. Un autre exemple est celui du signal avec une étincelle, où un delta de Dirac est injecté dans le signal sain, produisant ainsi une étincelle dans le signal en résultat.

La série temporelle saine après l'injection de la défaillance synthétique sera un état du type de la défaillance synthétique injectée. Pour une meilleure compréhension, une série chronologique saine avec une injection de défaillance de type étincelle est présentée ci-dessous.

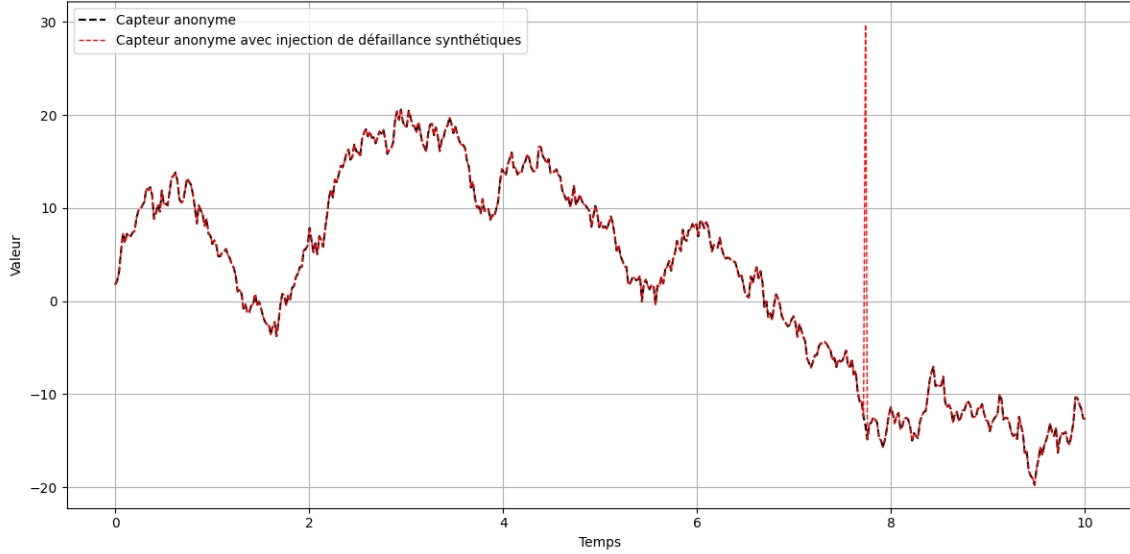


FIGURE 4.9 Capteur anonyme en noir dans le système avec injection de défaillance synthétique de type étincelle en rouge

Dans la figure 4.9, nous observons deux signaux d'un capteur anonyme. Le premier signal, représenté en noir avec des lignes pointillées, correspond à la série temporelle d'origine étiquetée comme série chronologique saine.

Le second signal, en rouge et également en lignes pointillées, représente le capteur anonyme avec une injection de défaillance synthétique de type étincelle. Un seul pic significatif a été ajouté à ce signal, illustrant une défaillance sur les mesures du capteur.

4.3.5 Étape 5 : Test du modèle

La phase de test consiste à évaluer la performance du modèle CVAE sur des données qu'il n'a pas encore vues. Elle permet de vérifier si le modèle a bien appris à généraliser ses capacités de reconstruction et de détection des anomalies à partir des données qu'il n'a jamais vues. Nous utilisons donc un ensemble de données de test distinct de celui de la formation, un ensemble de données de test distinct de l'ensemble de formation, composé d'images de événements normaux ainsi que d'images représentant des défaillances introduites de manière synthétique ou réelle dans les capteurs. Les images de test sont fournies en entrée au modèle, et le modèle tente de les reconstruire. Nous évaluons la qualité de la reconstruction en utilisant des métriques telles que la perte de reconstruction, qui mesure la différence entre les images d'entrée et les images reconstruites.

Le test du modèle nous permet de vérifier si le modèle est capable de généraliser à de nou-

velles données et s'il peut capturer efficacement les caractéristiques importantes de la région d'intérêt.

Les images de test sont fournies au modèle, et celui-ci tente de les reconstruire via son décodeur. L'évaluation des performances du modèle se fait principalement par l'analyse de la qualité des reconstructions, en utilisant des métriques appropriées. La perte de reconstruction, constitue une mesure fondamentale dans cette évaluation. Elle quantifie la différence entre les images d'entrée originales et les images reconstruites par le modèle. Plus cette perte est faible, meilleure est la capacité du modèle à reproduire fidèlement les données d'entrée. Une reconstruction de mauvaise qualité pourrait indiquer que le modèle n'a pas suffisamment capturé les caractéristiques importantes des données.

En parallèle à la mesure de la perte de reconstruction, nous utilisons également l'espace latent pour identifier des comportements anormaux dans les signaux capteurs. Dans cette approche, nous analysons la position des images dans l'espace latent, comparant la distance euclidienne des points associés aux images de test par rapport au centroïde de la distribution des événements normaux. Les images correspondant à des événements normaux devraient être proches du centroïde dans cet espace, tandis que celles associées à des événements défectueux devraient s'en écarter de manière significative. La distance euclidienne sert ainsi de métrique pour évaluer l'ampleur de l'anomalie : plus la distance est grande, plus il est probable que l'événement présente des défaillances.

Pour quantifier les performances du modèle dans la détection des anomalies, plusieurs indicateurs sont calculés :

- Taux de détection des anomalies (True Positive Rate) : Mesure la capacité du modèle à identifier correctement les événements avec défaillances parmi les images de test.
- Taux de fausses alarmes (False Positive Rate) : Mesure le nombre de événements normaux qui sont à tort détectés comme anormaux.

En outre, dans le cadre de la méthodologie proposée, l'utilisation des deux branches (erreur de reconstruction et distance dans l'espace latent) est particulièrement importante dans cette phase de test. L'approche basée sur l'erreur de reconstruction permet de repérer les anomalies en analysant directement la qualité des reconstructions des images de test, tandis que l'approche basée sur l'espace latent aide à détecter les événements anormaux grâce à une mesure de distance plus globale. Ces deux méthodes, lorsqu'elles sont combinées, renforcent la robustesse du modèle en fournissant deux perspectives complémentaires pour la détection des défaillances.

La phase de test du modèle ne se limite pas seulement à évaluer la capacité du CVAE à reconstruire les images d'entrée, mais elle permet aussi de valider la capacité du modèle à

détecter des anomalies dans les données des capteurs, que ce soit via une analyse directe des reconstructions ou en explorant l'espace latent. Les résultats obtenus lors de cette phase fournissent des informations essentielles pour juger de l'efficacité du modèle et de sa capacité à être utilisé dans des scénarios réels de détection des défaillances.

4.3.6 Étape 6 : Validation

Une fois que le modèle a été testé, nous procédons à son évaluation. L'évaluation vise à déterminer si le modèle est adapté à la tâche pour laquelle il a été construit. Pour évaluer le modèle, nous pouvons utiliser diverses métriques en fonction de l'objectif spécifique.

Parmi les métriques d'évaluation, la courbe ROC (Receiving Operating Characteristic) se distingue comme un outil d'analyse pour les tâches de classification binaire, telles que la détection d'anomalies avec un CVAE. La courbe ROC est tracée en portant le taux de vrais positifs (TPR) contre le taux de faux positifs (FPR) à divers seuils de décision. Ces seuils sont déterminés par l'erreur de reconstruction générée par le CVAE pour chaque image testée. Dans notre contexte, un seuil plus bas dans l'erreur de reconstruction indiquerait une image normale, tandis qu'un seuil plus élevé signalerait une anomalie.

L'AUC (Area Under the Curve) de la courbe ROC est un indicateur quantitatif de la performance du modèle. Une valeur d'AUC de 1 indiquerait une performance parfaite où le modèle distingue avec précision toutes les images normales et anormales. À l'inverse, une valeur de 0.5 suggérerait une performance identique à celle obtenue par une classification aléatoire.

La courbe ROC est tracée en calculant deux indicateurs statistiques à différents seuils de discrimination : Le taux de vrais positifs (TPR), qui est défini comme le rapport entre le nombre de vrais positifs (VP) et le nombre total de cas positifs réels (la somme des vrais positifs et des faux négatifs, $VP + FN$) :

$$TPR = \frac{VP}{VP + FN} \quad (4.21)$$

Le taux de faux positifs (FPR), qui est calculé comme le rapport entre le nombre de faux positifs (FP) et le nombre total de cas négatifs réels (la somme des faux positifs et des vrais négatifs, $FP + VN$) :

$$FPR = \frac{FP}{FP + VN} \quad (4.22)$$

Pour chaque seuil choisi, ces deux taux sont calculés en fonction de l'erreur de reconstruction produite par le CVAE pour chaque image. Les images avec une erreur supérieure au seuil sont classées comme anomalies (positifs), tandis que celles avec une erreur inférieure sont classées

comme normales (négatifs). En faisant varier ce seuil, nous obtenons différents couples de valeurs (FPR, TPR), qui nous permettent de tracer la courbe ROC.

4.4 Indices de santé

Un indice de santé (Health Index HI en anglais) est défini comme une mesure utilisée pour quantifier l'état d'un système par rapport à son état de santé connu. L'HI est généralement dérivé de la norme euclidienne des points de données dans un espace latent normalisé, tel que l'espace Nataf standard. Il permet de détecter les écarts par rapport aux conditions de fonctionnement normales, indiquant des dégradations ou des défauts potentiels. Différents types d'HI peuvent être utilisés pour identifier des changements rapides ou progressifs dans le comportement du système, en fonction des exigences de surveillance [Oliveira-Filho et al., 2024].

Pour calculer le seuil HI, nous utilisons le test statistique du chi carré et le degré de liberté déterminé. Lorsque nous traçons les résultats en 3D, nous avons 3 degrés de liberté, lorsque nous le faisons en 2D, nous avons 2 degrés de liberté. Pour évaluer la qualité de l'espace latent généré par un autoencodeur variationnel avec des filtres de réseau neuronal convolutif (CVAE), nous utilisons le processus d'indice de santé. Cela se fait en utilisant le test statistique du chi carré pour déterminer si les échantillons générés correspondent à une distribution attendue. Ce processus valide les performances du modèle génératif. Le processus est décrit ci-dessous :

1. Définition de la distribution attendue : Sur la base du modèle CVAE et de la configuration de l'espace latent, une distribution attendue est définie pour les échantillons générés. Dans le cas d'un espace latent tridimensionnel (3D), la distribution attendue est basée sur l'hypothèse que les échantillons doivent suivre une distribution normale multivariée standard avec trois degrés de liberté, correspondant aux trois dimensions de l'espace latent.
2. Génération d'échantillons : Un ensemble significatif d'échantillons est obtenu à partir de l'espace latent CVAE. Ces échantillons représentent les variables latentes apprises par le modèle pendant l'apprentissage. Avant cette étape, les échantillons sont centrés et réduits. Cela signifie qu'ils sont ajustés à l'aide de la moyenne et de la variance pour normaliser les données et permettre une comparaison précise avec la distribution attendue.
3. Catégorisation des données : Les échantillons générés sont divisés en catégories. Dans un espace 3D, cela peut impliquer de discrétiser chaque dimension en intervalles ou

d'utiliser une grille tridimensionnelle pour organiser les échantillons.

4. Calcul des fréquences observées et attendues : Les fréquences observées des échantillons de chaque catégorie sont calculées. Ces fréquences sont ensuite comparées aux fréquences attendues selon la distribution attendue précédemment définie. Les fréquences attendues sont basées sur le modèle théorique de distribution normale pour l'espace latent.
5. Application du test du chi carré : Pour déterminer s'il existe une différence notable entre les fréquences observées et attendues, nous effectuons le test du chi carré. Ce test calcule la statistique du chi carré à l'aide de la formule :

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

où O_i représente les fréquences que nous avons observées dans chaque catégorie, tandis que i et E_i désignent les fréquences que nous attendions en fonction de notre distribution théorique. L'ensemble de ces étapes forme un processus cohérent pour extraire des informations utiles à partir des capteurs et pour entraîner un modèle d'apprentissage automatique capable de détecter des défaillances.

CHAPITRE 5 RÉSULTATS ET DISCUSSIONS

Dans ce chapitre, nous présentons nos résultats obtenus pour la détection automatique de défaillances de capteurs. Les discussions sont aussi présentées.

5.1 Résultats

Dans ce chapitre, nous explorons deux approches pour l'extraction de descripteurs : l'une qui inclut la corrélation des capteurs et l'autre qui l'exclut. La première méthode calcule des descripteurs statistiques standard tels que la moyenne, la variance, le kurtosis, l'entropie et l'obliquité de chaque capteur. En revanche, la deuxième approche prend en compte en plus les corrélations entre les capteurs. En générant des images distinctes à partir des deux méthodes et en les utilisant comme entrées pour le modèle CVAE, nous comparons leurs performances en termes de détection des défaillances des capteurs. L'efficacité de chaque méthode sera évaluée en examinant leur impact sur l'erreur de reconstruction du modèle et la distribution des événements dans l'espace latent.

5.1.1 Images basées sur les descripteurs à partir des données des capteurs

La figure 5.1 représente trois images générées à partir des descripteurs statistiques, de leur normalisation (0,1) et de l'utilisation de niveaux de gris. Cinq descripteurs sont employés : moyenne, kurtosis, variance, asymétrie et entropie.

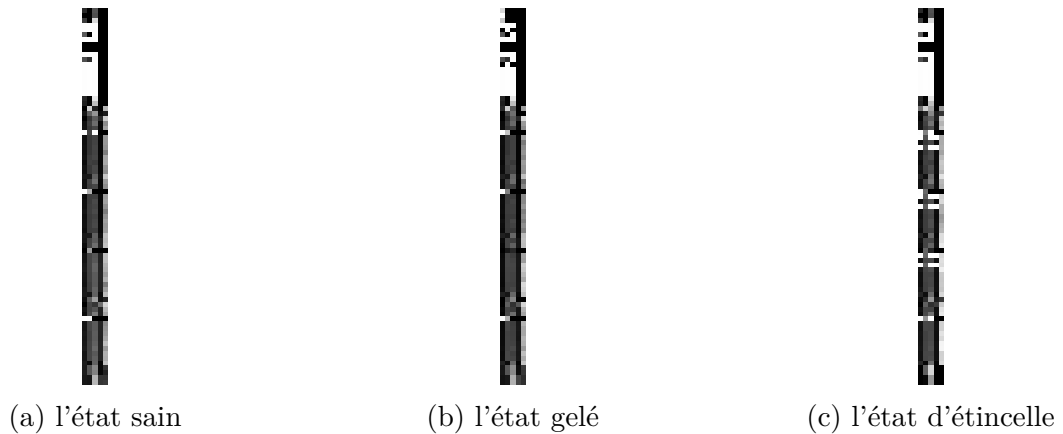


FIGURE 5.1 Images en niveaux de gris : représentation visuelle des caractéristiques des données pour différentes conditions d'état

La première correspond à un fonctionnement en condition saine, tandis que les deux autres sont destinées aux capteurs gelés et en présence d'un défaut de type étincelle. Chaque image a une taille de pixel de 5×77 , ce qui correspond à l'utilisation de 5 descripteurs sur 77 capteurs de la base de données. Par exemple, pour la première image, le pixel de la première ligne et de la première colonne de la matrice d'image représente le calcul moyen pour le capteur i pendant l'événement j . La première colonne de la matrice reflète les valeurs moyennes des données de séries temporelles des 77 capteurs. De même, la deuxième colonne représente la variance calculée à partir du même ensemble de capteurs. Les colonnes suivantes intègrent des mesures statistiques supplémentaires : la troisième colonne correspond à la kurtosis, la quatrième à l'entropie et la cinquième à l'obliquité.

5.1.2 Architecture et formation CVAE

Ce travail implémente l'autoencodeur variationnel convolutionnel [Zhang et al., 2022, Zemouri et al., 2019]. L'architecture CVAE et les hyperparamètres de formation ont été sélectionnés pour optimiser la convergence et la généralisation du modèle sur les ensembles de données de formation et de validation.

La dimension d'entrée est 5×77 . L'espace latent est défini avec la dimension $latent_dim = 2$ pour permettre le tracé de l'espace latent.

L'encodeur convolutionnel se compose de 5 couches Conv2D définies avec des tailles de filtre et des pas progressivement croissants pour le sous-échantillonnage. Après les couches convolutionnelles, une couche Flatten est appliquée pour transformer la sortie en un vecteur 1D, qui est ensuite introduit dans une couche Dense comprenant 256 unités. Chaque couche Conv2D utilise la fonction d'activation ReLU (Rectified Linear Unit) pour introduire la non-linéarité. Pour éviter le sur-ajustement, une couche Dropout avec un taux de Dropout de 0,5 est incorporée. Le décodeur reconstruit l'entrée d'origine à partir de la représentation de l'espace latent. Sa couche d'entrée est définie avec la dimension $latent_dim$, suivie d'une couche Dense avec $128 \times 5 \times 77$ unités et l'activation ReLU. Cette couche est remodelée en un tenseur 3D $5 \times 77 \times 128$ pour faciliter les couches Conv2DTranspose suivantes. Le décodeur reflète la structure de l'encodeur, en utilisant des couches Conv2DTranspose avec des tailles de filtre et des pas décroissants pour suréchantillonner la représentation aux dimensions de l'image d'origine. La couche Conv2DTranspose finale génère l'image reconstruite à l'aide d'une fonction d'activation sigmoïde.

Le processus de formation s'est étendu sur 20 époques avec une taille de lot de 32.

5.1.3 Résultats : Erreur de reconstruction - image composite à 5 descripteurs

La figure 5.2 montre la distribution de probabilité cumulée de l'erreur de reconstruction lors de l'utilisation d'une image composite à 5 caractéristiques pour évaluer les performances du modèle. L'axe x représente l'erreur de reconstruction, calculée comme la différence entre l'image d'entrée et l'image reconstruite à partir du CVAE, et l'axe y représente la probabilité cumulée de événements normaux et anormaux.

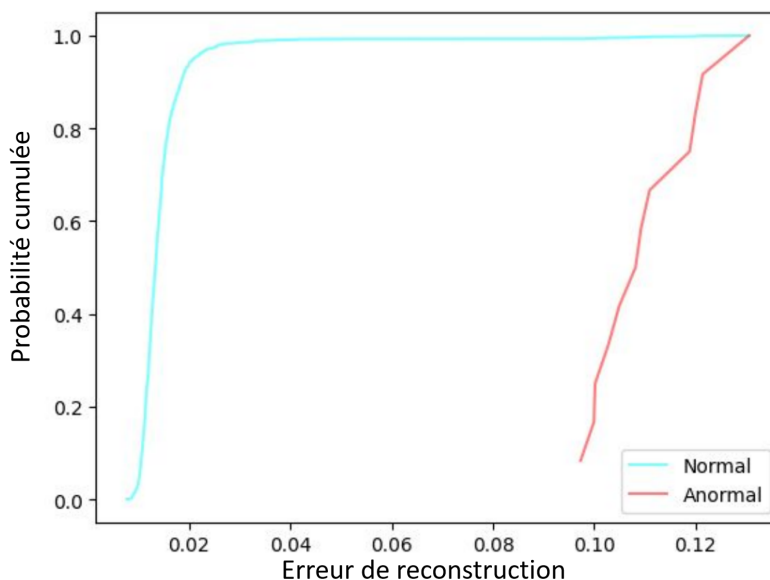


FIGURE 5.2 Erreur de reconstruction en utilisant des images à 5 descripteurs

Comme le montre le graphique, l'erreur de reconstruction est plus faible pour les événements normaux (ligne bleue), et la probabilité cumulée atteint presque 1,0, ce qui donne une valeur d'erreur relativement faible (environ 0,025). En revanche, pour les événements anormaux (ligne rouge), l'erreur de reconstruction est nettement plus élevée. La distance entre les deux courbes indique l'efficacité du CVAE dans l'identification des défauts sur la base d'erreurs de reconstruction dues à des événements anormaux qui s'écartent considérablement de la distribution apprise des données normales du modèle.

5.1.4 Résultats : Erreur de reconstruction - image composite à 6 descripteurs

La figure 5.3 montre la distribution de probabilité cumulée lorsque six caractéristiques sont utilisées pour créer l'image composite. Les tendances observées sur cette figure sont similaires à celles des 5 descripteurs, les événements normaux présentant des erreurs de reconstruction inférieures aux événements anormaux. Dans ce résultat, comme dans le résultat précédent,

l'axe des x représente l'erreur de reconstruction et l'axe des y correspond à la probabilité cumulée.

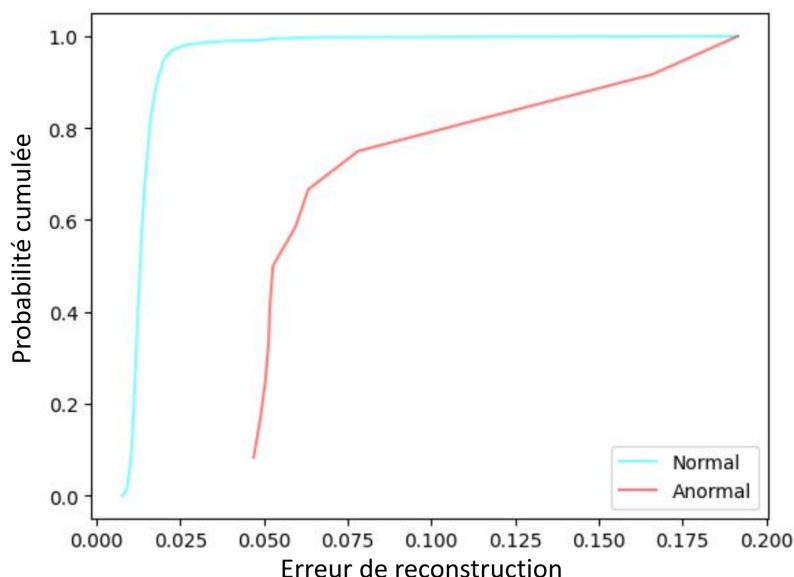


FIGURE 5.3 Erreur de reconstruction en utilisant des images à 6 descripteurs

Pour les événements normaux (ligne bleue), la probabilité cumulée augmente rapidement et atteint presque 1,0 avec une erreur de reconstruction d'environ 0,05. Cela indique que la plupart des événements normaux comportent des erreurs relativement faibles. En revanche, les événements anormaux (ligne rouge) montrent une erreur de reconstruction plus importante et la courbe penche davantage vers la droite. Cela indique que les erreurs de reconstruction ont tendance à être plus importantes lors des événements anormaux, car l'écart par rapport au comportement normal est plus prononcé. Les deux courbes montrent une distinction entre les événements normaux et anormaux, et l'erreur de reconstruction est significativement plus grande pour les événements anormaux, nous permettant de distinguer les deux classes.

5.1.5 Résultats : Espace latent - image composite à 5 descripteurs

La base de données a été divisée en parties égales pour la création des défaillances synthétiques, c'est-à-dire que les données sont équilibrées. Les défaillances synthétiques créées incluent le biais, l'étincelle, le gel, le bruit et la saturation. De cette manière, le modèle a été formé en utilisant 75% des données et les 25% restants ont été utilisés pour les tests. Les résultats sont présentés dans la figure 5.4.

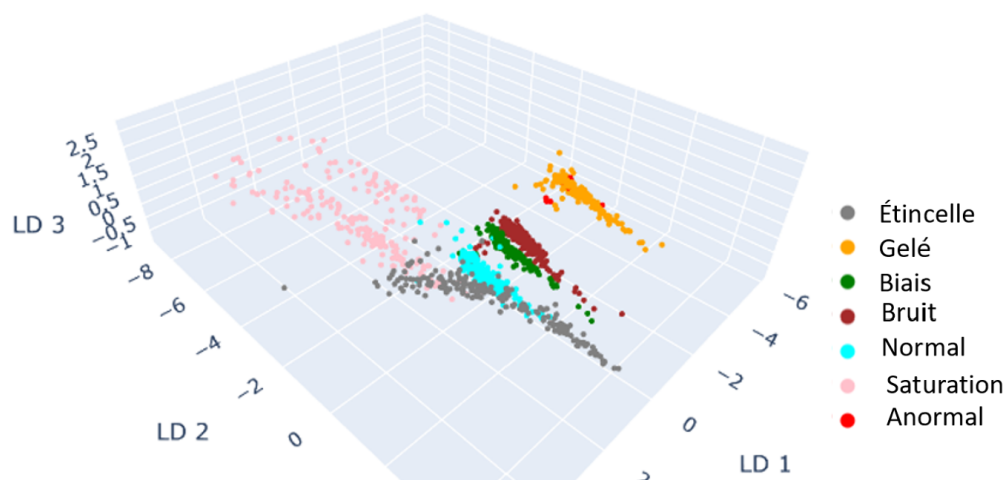


FIGURE 5.4 Espace latent utilisant des images à 5 descripteurs

Nous pouvons voir sur la figure que les différents types d'images de la base de données sont regroupés dans l'espace latent. Nous avons la région des images normales, ainsi que chaque groupe d'images de défaillances synthétiques. Le nom des axes est noté LD 1, 2 et 3, où LD signifie dimension latente.

Après avoir visualisé les données dans l'espace latent, chacune des positions de coordonnées des images a été réduite et centralisée. Cela a été fait afin de calculer l'indice de santé. C'est-à-dire la distance entre chacune des images étiquetées comme anormales et le centre du groupe étiqueté comme normal. Les résultats sont présentés dans la figure 5.5.

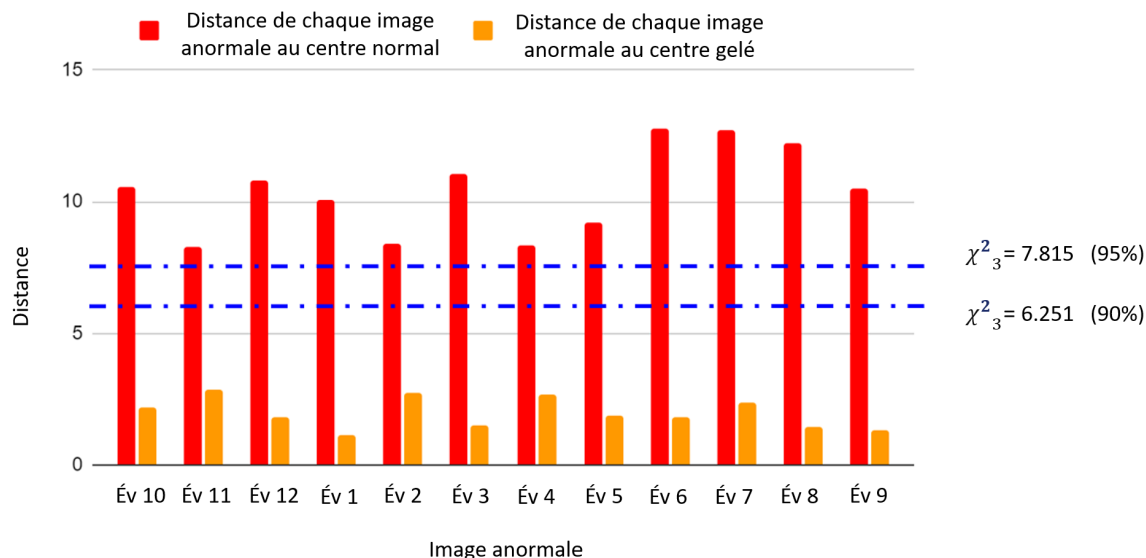


FIGURE 5.5 Distance entre une image anormale et le centre du groupe normal (graphique rouge) et une image anormale et le centre du groupe gelé (graphique orange). Utilisation d'images à 5 descripteurs

Le graphique de la figure 5.5 montre que la distance de chacune des images étiquetées comme anormales au centre du groupe étiqueté comme normal est supérieure au seuil défini par l'indice de santé. De plus, la distance de chaque image anormale au centre du groupe d'images étiquetées figées est inférieure au seuil défini par l'indice de santé.

La figure 5.6 montre l'espace latent qui résulte de l'utilisation de la corrélation comme sixième descripteur. Nous pouvons voir que la qualité de l'espace latent n'est pas meilleure par rapport à l'espace latent qui résulte de l'utilisation uniquement des cinq premiers descripteurs.

5.1.6 Résultats : Espace latent - image composite à 6 descripteurs

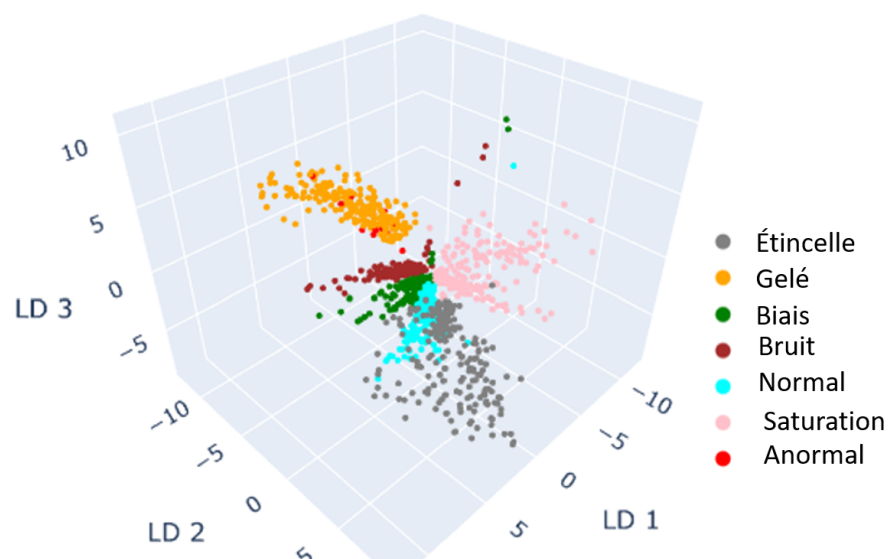


FIGURE 5.6 Espace latent utilisant des images à 6 descripteurs

Les résultats de l'indice de santé sont présentés dans la Figure 5.7 pour l'espace latent des images avec six descripteurs. On peut voir sur l'image que la qualité de l'espace latent en utilisant 6 descripteurs n'est pas meilleure qu'en utilisant 5 descripteurs, car il existe une région où il cohabitent différents types d'événements avec différents types de défaillances dans leurs capteurs. Malgré cela, les événements avec des capteurs défectueux sont localisés avec succès dans la région de faille gelée.

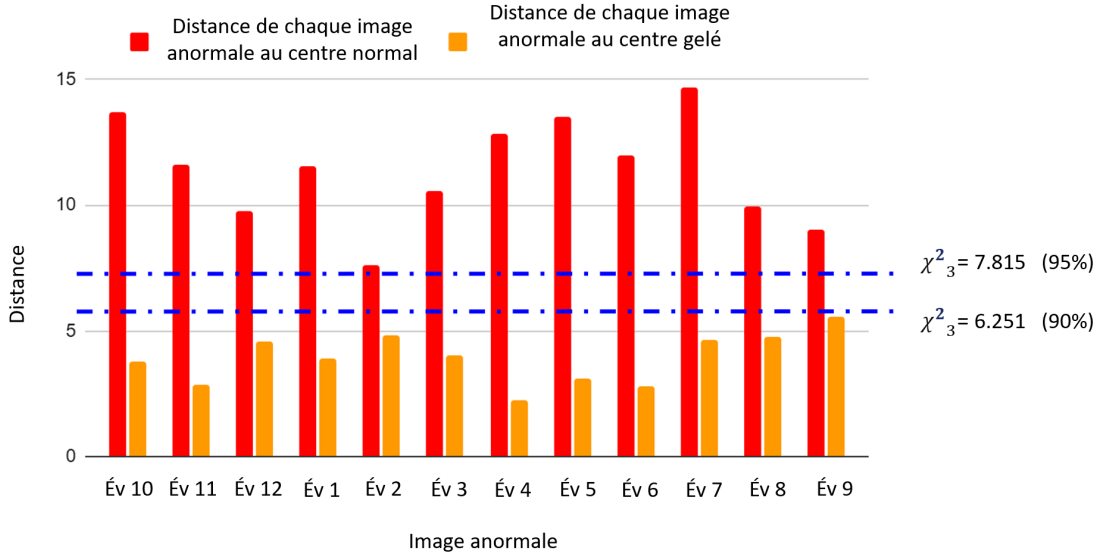


FIGURE 5.7 Distance entre une image anormale et le centre du groupe normal (graphique rouge) et une image anormale et le centre du groupe gelé (graphique orange). Utilisation d'images à 6 descripteurs

Bien que la qualité de l'espace latent utilisant des images à 6 descripteurs ne soit pas la même que celle du modèle utilisant des images à 5 descripteurs, nous pouvons visualiser dans la figure 5.7 que l'indice de santé fonctionne.

5.1.7 Évaluation de l'approche de détection proposée

La figure 5.8 présente la courbe ROC du modèle. L'aire sous la courbe est égale à 1, ce qui suggère la bonne performance de l'approche de détection proposée. Plus la courbe ROC épouse le coin supérieur gauche du graphique, plus le modèle réussit à classer les données en catégories. Pour quantifier cela, nous pouvons calculer l'AUC (aire sous la courbe) qui nous indique quelle partie du graphique est située sous la courbe. Plus l'AUC est proche de 1, meilleur est le modèle. Remarque

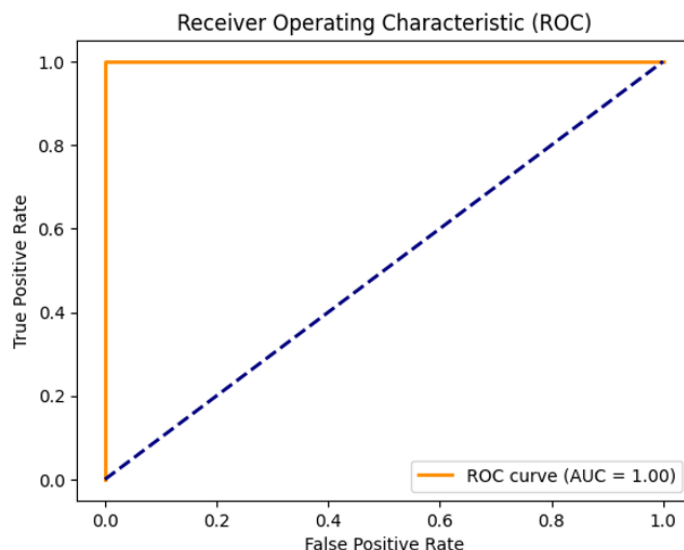


FIGURE 5.8 Courbe ROC de la méthodologie proposée

Ce score presque parfait signifie que le modèle démontre des performances optimales de classification. La courbe ROC illustre le taux de vrais positifs (TPR) par rapport au taux de faux positifs (FPR) sur différentes valeurs de seuil. Dans la figure 5.8, la courbe s'aligne sur le coin supérieur gauche du graphique, atteignant le point (0,1). Cela indique un scénario de classification sans faille, où toutes les instances positives sont correctement identifiées (vrais positifs) sans faux positifs.

$AUC = 1,0$ indique que le modèle distingue parfaitement les deux classes (événements normaux et événements avec des défaillances dans leurs capteurs). Cela signifie que pour toute instance positive choisie au hasard et toute instance négative choisie au hasard, le modèle classera toujours l'instance positive plus haut, ce qui implique qu'il n'y a aucun chevauchement entre les prédictions des deux classes.

Pour obtenir ces résultats, un ensemble de capteurs a été prélevé et des défaillances synthétiques ont été injectées. Bien que notre modèle a démontré une grande efficacité, il est essentiel de noter que l'ajout de données provenant d'autres capteurs avec défaillances différents de ceux du groupe sélectionné pour ces tests, pourrait entraîner une diminution de l'AUC. Cela est dû au fait que le modèle n'a pas encore été entraîné pour reconnaître ces nouvelles situations.

Pour maximiser la robustesse de notre approche, il serait idéal de former le modèle sur une diversité de types de défaillances, y compris des capteurs avec des défaillances synthétiques, pour chaque capteur de l'ensemble du système.

Bien qu'il s'agisse d'un excellent résultat, il est important de s'assurer que cette performance n'est pas due à un surapprentissage. En effet, grâce au graphique de perte, il est vérifié qu'il n'y a pas de surapprentissage.

5.2 Déformation temporelle dynamique : méthodologie indépendante

Pour comparer les résultats du modèle, d'autres techniques et outils ont été choisis pour détecter les défaillances des capteurs qui disposent de données de séries chronologiques. L'une de ces techniques est la déformation temporelle dynamique (Dynamic Time Warping DWT en anglais), cette technique est définie ci-dessous et les résultats obtenus en les utilisant dans la base de données sont présentés.

Déformation temporelle dynamique : definition

La déformation temporelle dynamique (DTW) est une mesure de distance classique et bien établie, particulièrement adaptée à la comparaison de séries chronologiques [Górecki and Łuczak, 2014]. Cette méthode mesure la similarité d'un motif avec différentes zones du signal. Plus la distance produite est petite, plus les deux modèles de signaux seront similaires [Permanasari et al., 2019].

Méthodologie en utilisant DTW

Dans cette étude, nous avons adopté une méthodologie pour comparer les résultats obtenus via la technique DTW avec ceux du modèle que nous proposons. La figure 5.9 montre ce méthodologie. L'approche consiste à sélectionner un événement de référence au hasard parmi tous les événements disponibles. Pour cet exemple, nous avons choisi le événement 1092. Les données analysées proviennent d'un capteur anonyme, qui mesure un phénomène physique. Par la suite, nous avons sélectionné aléatoirement 1000 autres événements, tout en excluant le événement 1092. Des défaillances synthétiques ont été injectées dans ce événement pour comparer les signaux d'un événement sans défaillance (cas 1) avec ceux de 1000 événements sains, ainsi que les signaux contenant ces défaillances synthétiques (cas 2) avec ces mêmes événements sains. Nous avons ensuite calculé les distances DTW entre ces signaux pour chaque cas, et représenté les résultats sous forme d'histogrammes.

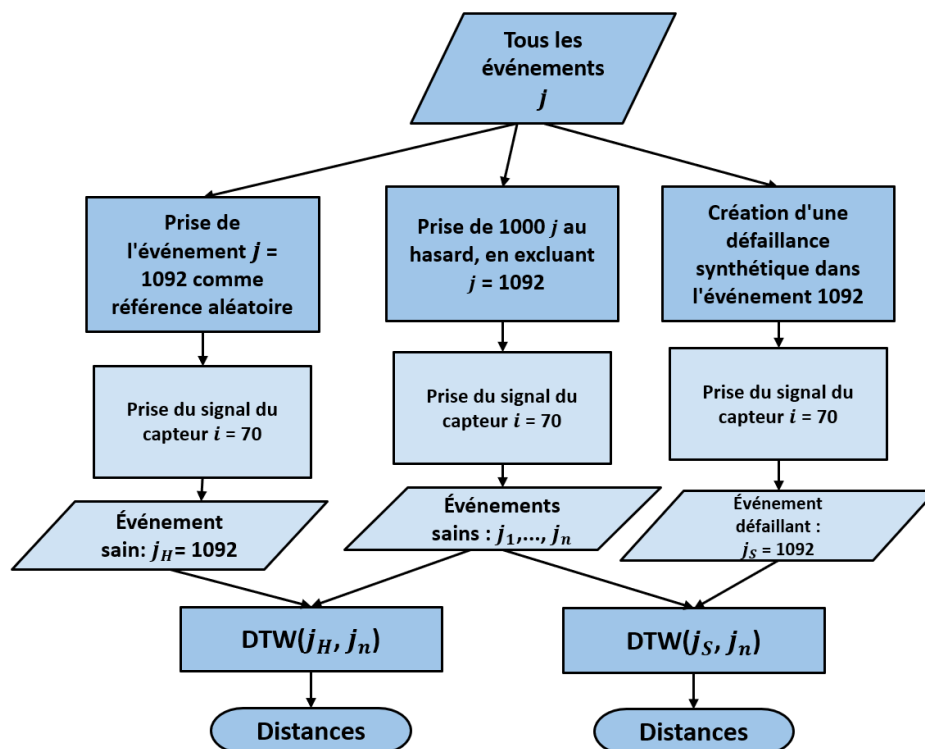


FIGURE 5.9 Organigramme présentant les étapes de la méthodologie proposée en utilisant déformation temporelle dynamique (DTW) pour la détection des défaillances de capteurs dans les systèmes industriels

A la fin du processus méthodologique, une superposition des deux histogrammes sera réalisée, observant et analysant les différences de distances. Ce processus nous amène à conclure si la technique DTW est adaptée ou non à la détection de défauts dans les séries temporelles de capteurs de notre base de données.

Pour valider ces résultats, on a introduit deux types de défaillances synthétiques : les étincelles et de la saturation.

Résultats et discussions en utilisant DTW

Des tests ont été réalisés avec les deux types de défaillances synthétiques mentionnés : étincelle et saturation. Les résultats sont représentés dans les histogrammes de la figure 5.10. Dans cette dernière, les résultats montrent l'analyse pour une défaillance de type étincelle. On y remarque que les distances DTW entre les signaux sains et les signaux altérés par des étincelles sont très similaires, sans différence significative.

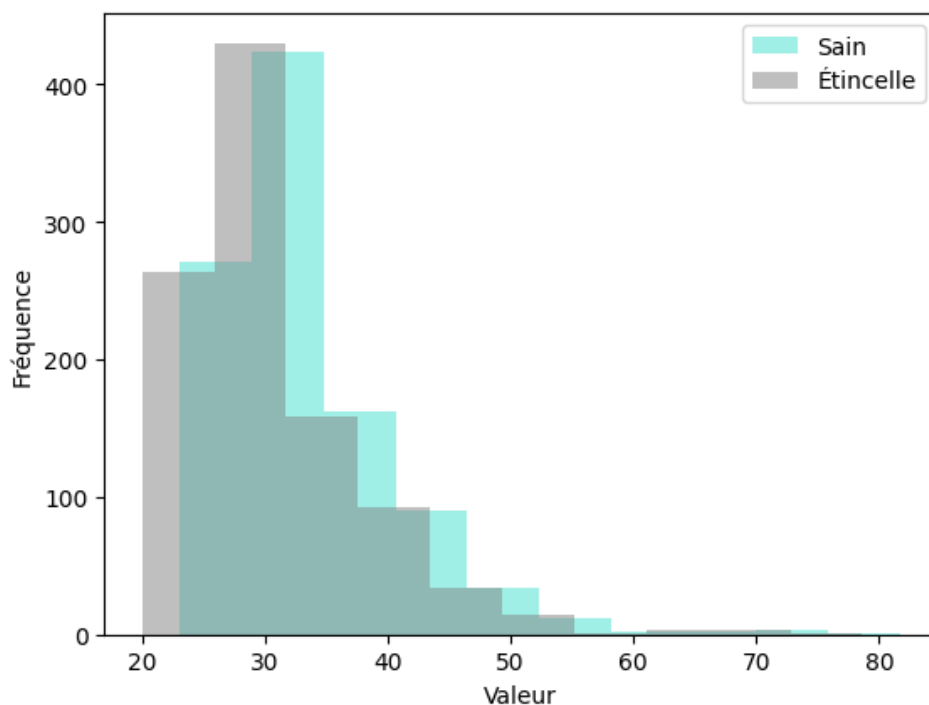


FIGURE 5.10 Histogramme présentant les résultats de la méthodologie, avec les échantillons de type 'étincelle' en gris et les échantillons de type 'sain' en bleu

De façon similaire, la figure 5.11 présente les résultats obtenus pour une défaillance de type saturation. Les distances calculées ne permettent toujours pas de différencier les signaux sains de ceux affectés par cette défaillance.

La technique DTW ne donne pas de bons résultats pour détecter les défaillances dans les signaux de base de données actuels. Les tests qu'ont été effectués montrent qu'il n'y a pas de différence significative dans les valeurs calculées DTW pour la saturation synthétique et les défaillances d'étincelles.

Les résultats obtenus montrent que la technique DTW ne permet pas de détecter les défaillances dans les signaux de notre base de données. Les tests réalisés ne montrent pas aucune différence notable entre les valeurs DTW calculées pour les deux types de défaillances, qu'il s'agisse de saturation ou d'étincelle. Cela suggère que cette méthode n'est pas adéquate dans notre contexte particulier.

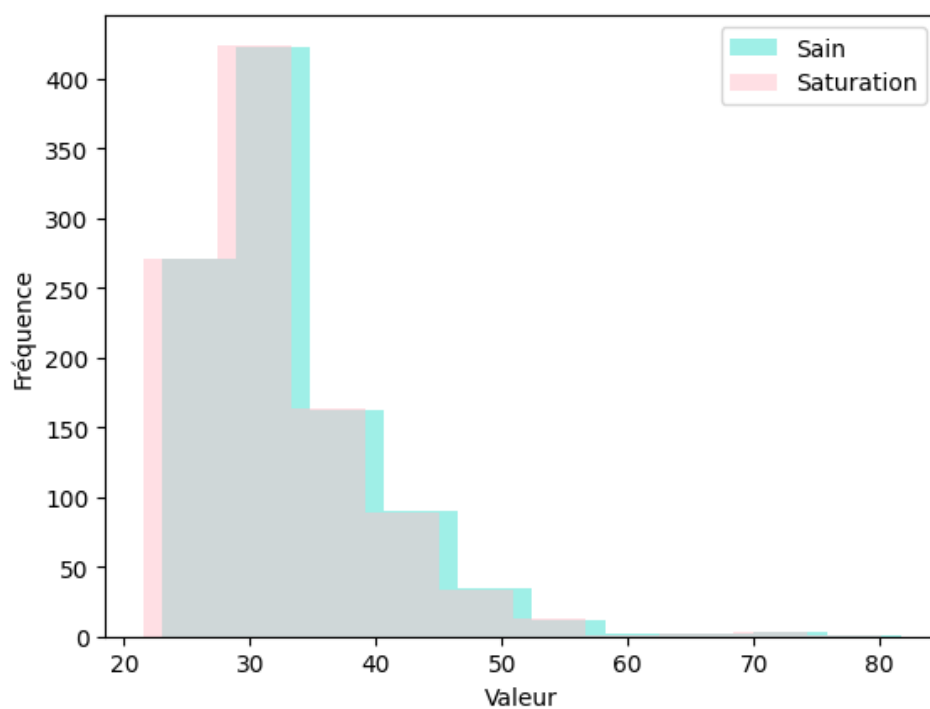


FIGURE 5.11 Histogramme présentant les résultats de la méthodologie, avec les échantillons de type 'saturation' en rose et les échantillons de type 'sain' en bleu

CHAPITRE 6 CONCLUSION

Dans ce travail de recherche, nous présentons une démarche pour le développement de méthodes pour la détection des défaillances (anomalies) des capteurs dans les systèmes électromécaniques, en mettant l'accent sur l'utilisation de l'apprentissage automatique. Les enjeux liés à la sécurité et à la fiabilité des capteurs sont primordiaux dans de nombreux contextes, et notre approche vise à combler un vide existant en matière de détection de défaillances.

À cette fin, nous avons défini une méthodologie basée sur l'apprentissage automatique. Les résultats préliminaires obtenus ont démontré le potentiel de notre approche dans la détection des changements de comportement des capteurs.

Les méthodes basées sur l'apprentissage automatique sont en constante évolution, ouvrant ainsi de nouvelles perspectives dans le domaine de la détection des défaillances des capteurs. Pour l'industrie aéronautique, une solution complète pour la détection des défaillances des capteurs est souhaitée. Actuellement, une telle solution n'existe pas, et notre travail apporte une contribution significative pour renforcer la sécurité et la fiabilité de ces aéronefs. L'application de l'apprentissage automatique dans ce contexte démontre son potentiel pour résoudre des problèmes complexes et cruciaux, permettant de nouvelles avancées dans l'industrie aéronautique.

Les retombées du projet présentent une étape essentielle pour garantir la performance et la sécurité des capteurs dans le domaine de l'aviation, et son impact pourrait s'étendre à d'autres secteurs qui exigent une surveillance précise des capteurs.

De plus, la comparaison des images composites à 5 et 6 descripteurs souligne l'importance de bien sélectionner les descripteurs utilisés dans le modèle. Bien que les deux configurations montrent une distinction claire entre les événements normaux et anormaux, l'utilisation de la corrélation comme fonctionnalité supplémentaire n'a pas amélioré la séparation spatiale potentielle. Le modèle à cinq fonctionnalités maintient des limites claires entre les différents types d'erreurs, comme le montrent l'indice de santé et la visualisation de l'espace latent. Cela suggère que les cinq premières caractéristiques contiennent suffisamment d'informations pour faire la distinction entre le fonctionnement normal du capteur et son fonctionnement défaillant. Le modèle à six descripteurs continue de fonctionner, mais le processus de détection des défauts devient plus complexe sans améliorer significativement sa précision. Ce résultat met en évidence la nécessité d'équilibrer la sélection des fonctionnalités et les performances du modèle pour éviter des rendements décroissants.

Bien que notre modèle a démontré une grande efficacité dans la détection des défaillances, il est crucial de reconnaître les limitations actuelles liées à l'intégration de données provenant d'autres capteurs, notamment ceux sujets au gel, saturation, bruit ou à d'autres types de défaillances. Nous prévoyons une diminution de l'aire sous la courbe (AUC) dans ces cas, car le modèle n'a pas encore été entraîné pour identifier ces nouvelles situations.

Pour maximiser la robustesse de notre approche, il serait idéal de former le modèle sur une diversité de types de défaillances pour chaque capteur de l'ensemble du système. Cela garantirait une performance optimale, permettant ainsi de maintenir une aire sous la courbe d'une valeur très élevée, même dans des conditions variées. Néanmoins, l'injection de multiples types de défaillances peut rendre l'apprentissage plus complexe pour le modèle, car il doit être capable de distinguer non seulement les défaillances des comportements normaux, mais aussi de différencier toutes les types de défaillances eux-mêmes

Aussi, dans le cadre de cette étude, l'utilisation de la technique DTW (Dynamic Time Warping) pour la détection des défaillances dans les séries temporelles des capteurs s'est révélée peu pertinente. Les résultats obtenus, que ce soit pour des défaillances de type étincelle ou saturation, n'ont pas permis de différencier efficacement les signaux altérés des signaux sains. Bien que la DTW soit couramment utilisée pour mesurer la similarité entre des séries temporelles, les faibles variations observées dans les distances calculées indiquent que cette approche n'est pas adaptée aux types de données et de défaillances analysées dans ce travail. En conséquence, cette méthode n'est pas appropriée pour la détection des défaillances dans les systèmes de capteurs aérospatiaux.

Les caractéristiques des défaillances injectées, telles que l'étincelle et la saturation, posent des défis supplémentaires. L'étincelle, bien que générant une perturbation brève, peut se produire de manière très localisée, ce qui complique l'identification de l'anomalie à travers une méthode d'alignement global comme la DTW. La saturation, quant à elle, peut produire un effet de signal plat, mais si elle ne dure pas assez longtemps ou si elle n'affecte qu'une portion spécifique du signal, elle peut ne pas être perçue de façon évidente par la méthode.

RÉFÉRENCES

- [Abid et al., 2015] Abid, A., Kachouri, A., Ben Fradj Guiloufi, A., Mahfoudhi, A., Nasri, N., and Abid, M. (2015). Centralized knn anomaly detector for wsn. *International Multi-Conference on Systems, Signals Devices*, 12(1) :51–54.
- [Ali et al., 2023] Ali, A. M. A., Tawfik, E. Z., Shaeb, M., and Mousbah, A. M. A. (2023). Sensor-to-sensor networking concepts based on local mobile agent platform. *International Journal of Advances in Engineering and Management (IJAEM)*, 5(7) :199–204.
- [An and Cho, 2015] An, J. and Cho, S. (2015). Variational autoencoder based anomaly detection using reconstruction probability. *Special lecture on IE*, 2(1) :1–18.
- [Balaban et al., 2009] Balaban, E., Saxena, A., Bansal, P., Goebel, K. F., and Curran, S. (2009). Modeling, detection, and disambiguation of sensor faults for aerospace applications. *IEEE Sensors Journal*, 9(12) :1907–1917.
- [Ballard, 1987] Ballard, D. H. (1987). Modular learning in neural networks. *Proceedings of the sixth National Conference on artificial intelligence*, 1 :279–284.
- [Bordoni and D’Amico, 1990] Bordoni, F. and D’Amico, A. (1990). Noise in sensors. *Sensors and Actuators A : Physical*, 21(1-3) :17–24.
- [Caruso et al., 2003] Caruso, G., Galeani, S., and Menini, L. (2003). Active vibration control of an elastic plate using multiple piezoelectric sensors and actuators. *Simulation Modelling Practice and Theory*, 11(5-6) :403–419.
- [Chen et al., 2022] Chen, Y., Gou, L., Li, H., and Wang, J. (2022). Sensor fault diagnosis of aero engine control system based on honey badger optimizer. *IFAC-PapersOnLine*, 55(3) :228–233.
- [Cook et al., 2020] Cook, A. A., Mısırlı, G., and Fan, Z. (2020). Anomaly detection for iot time-series data : A survey. *IEEE Internet of Things Journal*, 7(7) :6481–6494.
- [Dang and Suh, 2014] Dang, Q. K. and Suh, Y. S. (2014). Sensor saturation compensated smoothing algorithm for inertial sensor based motion tracking. *Sensors*, 14(5) :8167–8188.
- [Darvishi et al., 2020] Darvishi, H., Ciuonzo, D., Eide, E. R., and Rossi, P. S. (2020). Sensor-fault detection, isolation and accommodation for digital twins via modular data-driven architecture. *IEEE Sensors Journal*, 21(4) :4827–4838.
- [de Silva et al., 2021] de Silva, B. M., Callaham, J., Jonker, J., Goebel, N., Klemisch, J., McDonald, D., Hicks, N., Nathan Kutz, J., Brunton, S. L., and Aravkin, A. Y. (2021). Hybrid learning approach to sensor fault detection with flight test data. *AIAA Journal*, 59(9) :3490–3503.

- [Ehlenbröcker et al., 2016] Ehlenbröcker, J.-F., Mönks, U., and Lohweg, V. (2016). Sensor defect detection in multisensor information fusion. *Journal of Sensors and Sensor Systems*, 5(2) :337–353.
- [Feng et al., 2019] Feng, J., Hajizadeh, I., Yu, X., Rashid, M., Samadi, S., Sevil, M., Hobbs, N., Brandt, R., Lazaro, C., Maloney, Z., et al. (2019). Multi-model sensor fault detection and data reconciliation : A case study with glucose concentration sensors for diabetes. *AIChE Journal*, 65(2) :629–639.
- [Fong et al., 2023] Fong, K., Lee, C., Leung, M., Sun, Y., Zhu, G., Baek, S. H., Luo, X., Lo, T. K. K., and Leung, H. S. Y. (2023). A hybrid multiple sensor fault detection, diagnosis and reconstruction algorithm for chiller plants. *Journal of Building Performance Simulation*, 16 :588–608.
- [Górecki and Łuczak, 2014] Górecki, T. and Łuczak, M. (2014). Non-isometric transforms in time series classification using dtw. *Knowledge-based systems*, 61 :98–108.
- [Gunduz, 2021] Gunduz, H. (2021). An efficient dimensionality reduction method using filter-based feature selection and variational autoencoders on parkinson’s disease classification. *Biomedical Signal Processing and Control*, 66 :102–452.
- [Iglesias-Martínez et al., 2022] Iglesias-Martínez, M. E., Velasco-Pla, P. M., Antonino-Daviu, J., Carmenate, J. G., Dunai, L., Conejero, J. A., and Fernández de Córdoba, P. (2022). Multifractal 1-d wavelet leader based on spectral kurtosis of armature currents for sparking detection in dc motors. In *2022 International Conference on Electrical Machines (ICEM)*, volume 3, pages 1589–1594.
- [Jan et al., 2020] Jan, S. U., Saeed, U., and Koo, I. (2020). Machine learning for detecting drift fault of sensors in cyber-physical systems. 20 :389–394.
- [Kim et al., 2019a] Kim, M. S., Yun, J. P., Lee, S., and Park, P. (2019a). Unsupervised anomaly detection of lm guide using variational autoencoder. *International Symposium on Advanced Topics in Electrical Engineering (ATEE)*, 19 :10–15.
- [Kim et al., 2019b] Kim, S. G., Chae, Y. H., and Seong, P. H. (2019b). Signal fault identification in nuclear power plants based on deep neural networks. *Annals of DAAAM and Proceedings*, 30.
- [Li and Li, 2013] Li, C. and Li, Y. (2013). A spike-based model of neuronal intrinsic plasticity. *IEEE Transactions on Autonomous Mental Development*, 5(1) :62–73.
- [Li et al., 2020] Li, D., Wang, Y., Wang, J., Wang, C., and Duan, Y. (2020). Recent advances in sensor fault diagnosis : A review. *Sensors and Actuators A : Physical*, 309 :111990.
- [Marimont and Tarroni, 2021] Marimont, S. N. and Tarroni, G. (2021). Anomaly detection through latent space restoration using vector quantized variational autoencoders. *Inter-*

- national Symposium on Biomedical Imaging (ISBI)*, 18 :1764–1767.
- [Moreno Haro et al., 2023] Moreno Haro, L., Tahan, A., and Agard, B. (2023). Aperçu des méthodes de prédiction de pannes. *CIGI Qualita MOSIM, Trois-Rivières, Québec, Canada, Art. 3369*.
- [Noshad et al., 2019] Noshad, Z., Javaid, N., Saba, T., Wadud, Z., Saleem, M. Q., Alzahrani, M. E., and Sheta, O. E. (2019). Fault detection in wireless sensor networks through the random forest classifier. *Sensors*, 19(7) :1568.
- [Oliveira-Filho et al., 2023] Oliveira-Filho, A., Zemouri, R., Cambron, P., and Tahan, A. (2023). Early detection and diagnosis of wind turbine abnormal conditions using an interpretable supervised variational autoencoder model. *Energies*, 16(12) :4544.
- [Oliveira-Filho et al., 2024] Oliveira-Filho, A., Zemouri, R., Pelletier, F., and Tahan, A. (2024). System condition monitoring based on a standardized latent space and the nataf transform. *IEEE Access*, 12 :32637–32659.
- [Permanasari et al., 2019] Permanasari, Y., Harahap, E. H., and Ali, E. P. (2019). Speech recognition using dynamic time warping (dtw). In *Journal of physics : Conference series*, volume 1366, page 012091. IOP Publishing.
- [Qin et al., 2021] Qin, K., Zhou, Y., Tian, B., and Wang, R. (2021). Attentionae : Autoencoder for anomaly detection in attributed networks. *International Conference on Networking and Network Applications (NaNA)*, pages 480–484.
- [Reddy et al., 2020] Reddy, G. N. K., Manikandan, M. S., and Murty, N. N. (2020). On-device integrated ppg quality assessment and sensor disconnection/saturation detection system for iot health monitoring. *IEEE Transactions on Instrumentation and Measurement*, 69(9) :6351–6361.
- [Sun et al., 2018] Sun, J., Wang, X., Xiong, N., and Shao, J. (2018). Learning sparse representation with variational auto-encoder for anomaly detection. *IEEE Acces*, 6 :33353–33361.
- [Zemouri et al., 2019] Zemouri, R., Levesque, M., Amyot, N., Hudon, C., Kokoko, O., and Tahan, S. A. (2019). Deep convolutional variational autoencoder as a 2d-visualization tool for partial discharge source classification in hydrogenerators. *IEEE Access*, 8 :5438–5454.
- [Zhang et al., 2014] Zhang, K., Du, K., and Ju, Y. (2014). Algorithm of railway turnout fault detection based on pnn neural network. 1(1) :544–547.
- [Zhang et al., 2022] Zhang, Y., Xie, X., Li, H., and Zhou, B. (2022). An unsupervised tunnel damage identification method based on convolutional variational auto-encoder and wavelet packet analysis. *Sensors*, 22(6) :2412.

- [Zhao et al., 2016] Zhao, Z., Sun, Y.-g., and Zhang, J. (2016). Fault detection and diagnosis for sensor in an aero-engine system. 1(1) :2977–2982.
- [Zidi et al., 2018] Zidi, S., Moulahi, T., and Alaya, B. (2018). Fault detection in wireless sensor networks through svm classifier. *IEEE Sensors Journal*, 18(1) :340–347.