

Évaluation de méthodes basées sur les noyaux de convolution aléatoires pour la détection de défauts mécaniques dans les machines tournantes

Mouhamadou Mansour LO^{*†}, Gildas MORVAN^{*‡}
Mathieu ROSSI^{*†}, Fabrice MORGANTI^{*†}, David MERCIER^{*‡}

* Univ. Artois

† UR 4025, Laboratoire Systèmes Electrotechniques et Environnement (LSEE), F-62400 Béthune, France

‡ UR 3926, Laboratoire de Génie Informatique et d'Automatique de l'Artois (LGI2A), F-62400 Bethune, France

prenom.nom@univ-artois.fr

RESUME - Cet article présente une évaluation des performances de plusieurs algorithmes de classification de séries temporelles basés sur les noyaux de convolution aléatoires dans le cadre d'un problème de diagnostic de machine tournante en temps réel.

ABSTRACT - This article presents a performance evaluation of several time series classification algorithms based on random convolution kernels within the framework of real-time rotating machinery diagnostics.

MOTS-CLES - Diagnostic de machines tournantes, Apprentissage automatique, Classification de séries temporelles, Noyaux de convolution aléatoires

1. Introduction

Le diagnostic de défauts sur des machines électriques tournantes est un problème récurrent sur un process industriel. De nombreuses méthodes ont été développées afin de déterminer des indicateurs permettant de prévenir une éventuelle dégradation d'un constituant mécanique (rupture de barre, roulement, déséquilibre, excentricité) ou d'un défaut électrique (court-circuit inter-spires). Les méthodes classiques reposent souvent sur des analyses fréquentielles et exigent une connaissance experte (physique) de ces différents défauts ainsi que des caractéristiques des machines (nombre de pôles, vitesse de rotation, point de fonctionnement) [1], [8]. Elles peuvent donc être difficiles à mettre en œuvre, peu généralisables et ne sont pas toujours concluantes en pratique. Ces dernières années, de nombreuses méthodes basées sur l'apprentissage automatique, et notamment l'apprentissage profond (*Deep Learning*) ont émergé [7]. Si elles permettent d'obtenir de très bonnes performances, elles nécessitent de disposer d'un grand nombre de données et d'importantes capacités de calcul.

Dans cet article nous proposons de traiter ce problème à l'aide d'algorithmes de classification de séries temporelles légers, c'est-à-dire pouvant être entraînés rapidement sur un jeu de données réduit. En effet, nous nous plaçons dans un contexte d'analyse temps réel et embarqué, excluant des méthodes populaires type apprentissage profond. Plus précisément, nous proposons d'investiguer plusieurs méthodes récentes basées sur les noyaux de convolution aléatoires.

Dans un premier temps, nous présenterons les travaux existants liés à notre approche : le diagnostic de machines tournantes et la classification de séries temporelles. Puis, nous présenterons plus en détail les méthodes que nous allons utiliser. Enfin, nous montrerons et analyserons les résultats obtenus et conclurons.

2. Travaux connexes

2.1 Diagnostic de machines tournantes

Connaître l'état de fonctionnement ou de dysfonctionnement d'une machine peut se faire par l'usage de mesures invasives ou non-invasives. Les méthodes de diagnostic invasives utilisent généralement des mesures comme le courant, la tension, le couple ou la température. L'instrumentation de la machine nécessite alors son arrêt complet. A contrario les méthodes de diagnostic non-invasives peuvent être mises en place sans avoir à arrêter la machine. On s'intéresse alors à des grandeurs électriques, magnétiques ou mécaniques facilement mesurables telles que le flux de dispersion magnétique rayonné par l'entrefer, le bruit acoustique, ou les vibrations.

On peut distinguer trois familles de méthodes de diagnostic [7].

Méthodes statistiques / traitement du signal Ces méthodes supposent une connaissance experte de la machine étudiée. Elles reposent pour la plupart sur l'analyse du spectre harmonique, l'étude de certaines fréquences permettant de détecter un défaut particulier.

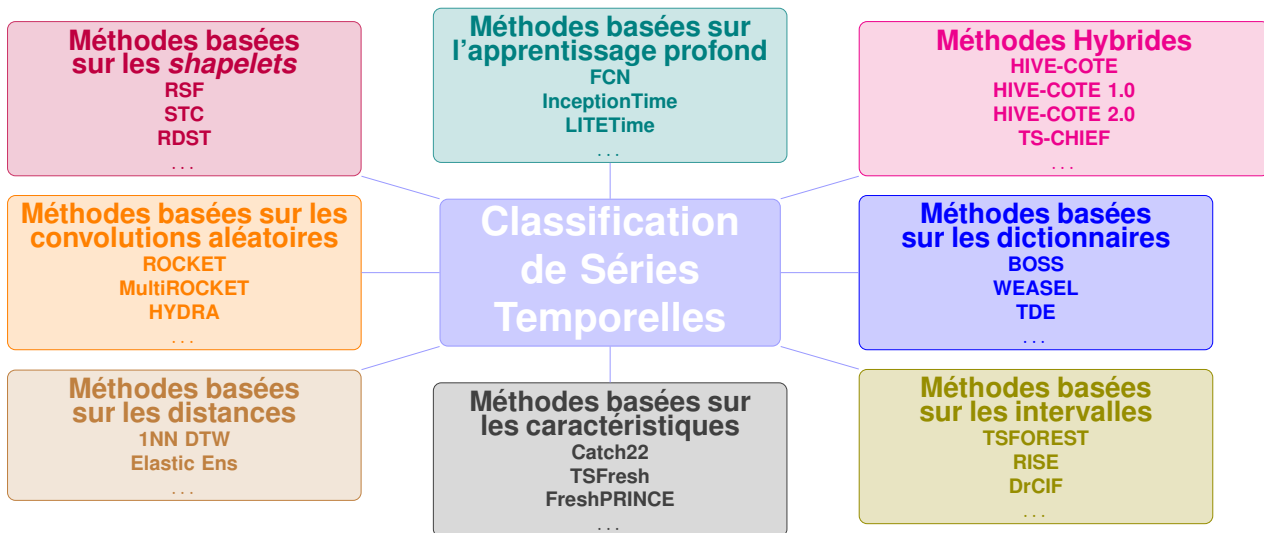


FIGURE 1 – Taxonomie des principales méthodes de classification de séries temporelles

Méthodes basées sur des algorithmes classiques d'apprentissage automatique Ici, comme dans la famille précédente, des caractéristiques sont extraites à l'aide d'une méthode spécifique (e.g. transformée de Fourier, ondelettes) puis injectées dans un classifieur classique (i.e. ne reposant pas sur l'apprentissage profond) comme une machine à vecteurs de support ou une forêt aléatoire.

Méthodes basées sur l'apprentissage profond Contrairement aux deux familles précédentes, ces méthodes prennent en entrée une série temporelle brute, sans aucun pré-traitement spécifique. L'extraction des caractéristiques pertinentes est directement apprise par le réseau de neurones et repose souvent sur la convolution [10]. La phase d'apprentissage consiste ainsi à apprendre les paramètres des noyaux de convolution aux différentes couches du réseau.

Notons qu'à l'exception de l'apprentissage profond, les approches évoquées ne mettent pas en œuvre les techniques les plus récentes de classification de séries temporelles bien que ce champ de recherche ait connu un essor important ces dernières années [6], [11]. Nous présentons brièvement ces méthodes dans la section suivante.

2.2 Classification de séries temporelles

La classification de séries temporelles est un sous-domaine de l'apprentissage automatique avec des applications concrètes dans de nombreux domaines. Il s'agit d'une tâche d'apprentissage supervisé qui vise à trouver une relation entre un ensemble de séries temporelles et une variable cible (classe). Dans la littérature, il existe de nombreuses approches permettant de réaliser cette tâche [11]. Elles peuvent être regroupées en plusieurs familles (cf figure 1), principalement selon la représentation de l'information et les traitements effectués sur les données.

Méthodes basées sur la distance Une fonction de distance comme *Dynamical Time Warping* [12] est utilisée pour mesurer la similarité entre deux séries temporelles. Un classifieur classique comme l'algorithme des k plus proches voisins (k -NN) peut alors être utilisé.

Méthodes basées sur les caractéristiques Elles vont chercher à extraire des statistiques descriptives (caractéristiques) résumant toute la série temporelle pour les utiliser dans un classifieur classique.

Méthodes basées sur les intervalles Elles génèrent des intervalles (aléatoires pour la majorité) puis extraient des caractéristiques sur chacun de ceux-ci. Dans certains cas, la séquence permettant de discriminer la série est assez petite, utiliser des intervalles au lieu de toute la série permet donc d'éviter une trop forte perte d'information locale.

Méthodes basées sur les *shapelets* (petites formes) Les *shapelets* sont de courts motifs capables de discriminer une classe en fonction de leur présence ou absence dans la série temporelle. Ces méthodes sont très efficaces pour détecter des motifs inhabituels qui peuvent apparaître occasionnellement dans une série temporelle.

Méthodes basées sur les dictionnaires Les séries temporelles sont transformées en séquences de mots. La similarité entre séries sera ainsi calculée à partir de la distribution des mots.

Méthodes basées sur l'apprentissage profond L'apprentissage profond suscite beaucoup d'intérêt ces dernières années, y compris dans le domaine de la classification de séries temporelles. Les techniques de cette famille se basent sur différentes architectures de réseaux de neurones profonds existantes. Les méthodes les plus connues font usage de la convolution [10] ou plus récemment, du mécanisme d'attention [6].

Méthodes basées sur les noyaux de convolution aléatoires À la différence des méthodes précédentes où les poids des noyaux de convolution sont appris, l'idée est ici de combiner des caractéristiques extraites de très nombreuses convolutions aléatoires afin de capturer des motifs complexes à différentes échelles [2].

Méthodes hybrides Elles agrègent les résultats de plusieurs méthodes issues des différentes familles évoquées ci-dessus. De par leur nature composite, elles obtiennent généralement de meilleures performances que les autres mais nécessitent un temps d'exécution relativement long.

Parmi toutes ces méthodes, nous nous intéresserons plus particulièrement à celles basées sur les noyaux de convolution aléatoires. En effet, elles permettent d'obtenir de très bonnes performances tout en étant peu gourmandes en temps de calcul [2]. Par ailleurs, comme pour les méthodes basées sur l'apprentissage profond, il n'y a pas de phase d'extraction de caractéristiques : on classe des séries temporelles brutes. En revanche, contrairement à ces dernières, elles permettent de travailler avec un jeu de données réduit. Elles semblent donc très prometteuses et, à notre connaissance, n'ont pas été utilisées dans ce contexte. Nous présentons les principales méthodes dans la section suivante.

3. Méthodes basées sur les noyaux de convolution aléatoires : une approche performante et rapide

Comme nous l'avons dit ci-dessus, ces méthodes font usage de convolution aléatoires pour discriminer les séries temporelles. Les noyaux de convolution peuvent être vus comme des mini-séquences aléatoires dont la convolution avec des séries temporelles génère des cartes d'activation indiquant la position des sous-séquences les plus similaires à ces noyaux. Ces cartes sont ensuite exploitées pour obtenir plusieurs caractéristiques synthétiques capables de discriminer les séries. Ces caractéristiques sont ensuite injectées dans un simple classifieur linéaire. La phase d'apprentissage consiste donc seulement à apprendre les bons poids à assigner à chaque caractéristique, cela permet aux membres de cette famille d'être aujourd'hui les méthodes de classification de séries temporelles les plus rapides tout en garantissant de très bonnes performances [11].

ROCKET (*Random Convolutional Kernel Transform*) [2] est la méthode pionnière de cette famille. Elle génère aléatoirement un grand nombre de noyaux de convolution (typiquement 10000) qu'elle utilise pour créer des cartes d'activation. Ces dernières sont ensuite résumées par deux opérateurs (*pooling operators*) : le PPV (*Proportion of Positive Values*), qui consiste à calculer le pourcentage de valeurs positives et le GMP (*Global Maximum Pooling*), la valeur maximale contenue dans la carte. Pour chaque série temporelle, $2k$ caractéristiques sont ainsi extraites où k représente le nombre de noyaux. Bien qu'aléatoires, les noyaux sont paramétrisés :

- La longueur est choisie aléatoirement entre $\{7, 9, 11\}$ avec équi-probabilité.
- Les poids $w \sim \mathcal{N}(0, 1)$ sont sélectionnés aléatoirement puis normalisés $w = W - \bar{W}$.
- Un biais $b \sim \mathcal{U}(-1, 1)$ est ajouté à la carte d'activation.
- Les valeurs de dilatation s'obtiennent avec $d = 2^x$, $x \sim \mathcal{U}(0, A)$ avec $A = \log_2(\frac{l_{série}-1}{l_{noyau}-1})$

ROCKET peut être vu comme un cas particulier de réseau de neurones convolutif avec une seule couche cachée, des poids de noyaux aléatoires, sans fonction d'activation (e.g. RELU) et sans normalisation (e.g. *Batch Normalization*).

MINIROCKET [3] est une variante de ROCKET avec quelques modifications clés permettant de grandement accélérer le temps d'apprentissage sans pour autant subir des pertes de performance. La méthode est rendue beaucoup plus déterministe en opérant les changements suivants :

- La longueur des noyaux est fixée à 9 au lieu de $\{7, 9, 11\}$.
- L'utilisation d'un unique ensemble de 84 noyaux ne contenant que $\{-1, 2\}$ comme valeurs. Ce changement majeur permet d'optimiser les opérations de convolution et ainsi réduire significativement le temps de calcul.
- Le biais est maintenant issu du résultat de la convolution avec un couple noyau/dilatation.
- Le GMP n'est plus calculé, uniquement le PPV.

MultiROCKET [13] étend MINIROCKET en apportant quelques améliorations :

- La différence du premier ordre entre deux unités de temps, i.e. la vitesse de variation, est utilisée comme représentation supplémentaire de la série temporelle.
- En plus du PPV, on extrait les caractéristiques suivantes : la valeur moyenne des valeurs positives (MPV - *Mean of Positive Value*), la valeur moyenne des indices des valeurs positives (MIPV - *Mean of Indice of Positive Values*), la valeur de la plus longue série de valeurs positives (LSPV - *Longest Stretch of Positive Values*).

HYDRA (*Hybrid Dictionary-ROCKET Architecture*) [4] est la dernière méthode en date de cette famille. Elle combine des aspects propres aux méthodes de type dictionnaire et convolutions aléatoires. L'idée générale est de regrouper l'ensemble des noyaux en g groupes de k noyaux. Pour chaque série temporelle du jeu de données, on effectue une convolution avec tous les noyaux d'un groupe donné. À chaque point de la série, le noyau conduisant à la plus grande valeur de convolution est sélectionné. Ainsi, pour chaque groupe, un histogramme des réponses maximales des noyaux est créé, faisant office de caractéristiques pour le classifieur. La différence du premier ordre est là encore utilisée. Les noyaux sont paramétrisés ainsi :

- La longueur est fixée à 9.
- Les poids $w \sim \mathcal{N}(0, 1)$ sont sélectionnés aléatoirement.
- On n'utilise pas de biais, ni d'opérateur à l'issue de la convolution.

4. Expérimentations

Dans cette section, des résultats expérimentaux sont présentés afin de tester et comparer ces méthodes sur un problème de détection de défauts mécaniques (roulements et dégradations de la barre de rotor) d'un moteur asynchrone. Les implémentations des méthodes de base sont celles fournies par leurs auteurs respectifs.

4.1 Données

Le jeu de données utilisé est disponible en ligne et diffusé sous licence CC BY 4.0 [9]. Il contient des enregistrements du courant triphasé mesuré sur un moteur asynchrone avec un taux d'échantillonnage de 10 kHz.

On dispose de 39 enregistrements correspondants :

- à la combinaison de 3 niveaux de charge (100W, 200W, 300W), 2 type de problèmes (bague intérieure et extérieure de roulement) et 6 niveaux de criticité du défaut (0.7mm, 0.9mm, 1.1mm, 1.3mm, 1.5m, 1.7mm), soit 36 enregistrements,
- 2 niveaux de charge (100W, 300W) pour la détection d'un problème de dégradation d'une barre de rotor,
- et un enregistrement de fonctionnement du moteur sain.

Chaque enregistrement contient les mesures des 3 phases du courant délivré par le moteur en fonction du temps. Afin de travailler sur des séries temporelles univariées, nous allons seulement nous intéresser à la première phase pour évaluer les performances des différentes méthodes considérées.

En fonction de l'état de fonctionnement du moteur, un nombre différent de sous-séquences est extrait : 1800 sous-séquences pour l'état sain, 1800 sous-séquences (2×900) pour l'état de défaut de la barre du rotor, 4500 sous-séquences (250×18) pour l'état défaut de bague interne et 4500 sous-séquences (250×18) pour l'état défaut de bague externe, soit 12600 sous-séquences au total. Les sous-séquences sont de grande taille (2500 points) afin que chacune d'elles contienne suffisamment d'information pour identifier l'absence ou la présence d'un défaut.

L'ensemble des 12600 sous-séquences est divisé en 2 : le premier sous-ensemble (75% des données) est utilisé pour apprendre les paramètres du modèle, tandis que les 25% restants permettent de valider sa capacité à classifier correctement des données qui n'ont pas été utilisées pour l'apprentissage. Le découpage entraînement / validation est randomisé.

Les méthodes basées sur les noyaux de convolution aléatoires étant non-déterministes, 15 jeux de données de 12600 sous-séquences ont été échantillonnées des 39 enregistrements d'origine afin d'obtenir des résultats statistiquement significatifs.

4.2 Méthodes mises en œuvre

Plusieurs variantes des méthodes présentées dans la section 3. ont été mises en œuvre. Elles sont présentées dans le tableau I.

MINIROCKET (9996 caractéristiques) et MultiRocket (49728 caractéristiques) n'ont été testés qu'avec leurs versions de base.

Pour ROCKET, nous avons testé la version de base (PPV + GMP = 20000 caractéristiques) et deux variantes avec pour seul opérateur le PPV ou le GMP (10000 caractéristiques).

HYDRA a également été testé avec différentes variantes en concaténant les caractéristiques produites par HYDRA et celles des autres méthodes : HYDRA de base (9216 caractéristiques), HYDRA + ROCKET (29216 caractéristiques), HYDRA + MINIROCKET (19212 caractéristiques), HYDRA + MultiRocket (58944 caractéristiques).

TABLE I – Caractéristiques des différentes méthodes

Algorithmes	Complexité des Transformations	Complexité du Classifieur	Nb de Caractéristiques
ROCKET	$\mathcal{O}(k \cdot n \cdot \ell)$	$\mathcal{O}(n^2 \cdot f)$	$2 \times k$
ROCKET PPV	$\mathcal{O}(k \cdot n \cdot \ell)$	$\mathcal{O}(n \cdot f^2)$	k
ROCKET GMP	$\mathcal{O}(k \cdot n \cdot \ell)$	$\mathcal{O}(n \cdot f^2)$	k
MINIROCKET	$\mathcal{O}(k \cdot n \cdot \ell)$	$\mathcal{O}(n \cdot f^2)$	$84 \times \lfloor \frac{k}{84} \rfloor \simeq k$
MultiRocket	$\mathcal{O}(k \cdot n \cdot \ell)$	$\mathcal{O}(n^2 \cdot f)$	$2 \times 4 \times k$
HYDRA	$\mathcal{O}(k \cdot n \cdot \ell)$	$\mathcal{O}(n \cdot f^2)$	$2 \times g \times k_g \times d$
HYDRA + ROCKET	$\mathcal{O}(k \cdot n \cdot \ell)$	$\mathcal{O}(n^2 \cdot f)$	$2 \times g \times k_g \times d + 2 \times k$
HYDRA + MINIROCKET	$\mathcal{O}(k \cdot n \cdot \ell)$	$\mathcal{O}(n^2 \cdot f)$	$2 \times g \times k_g \times d + 84 \times \lfloor \frac{k}{84} \rfloor$
HYDRA + MultiRocket	$\mathcal{O}(k \cdot n \cdot \ell)$	$\mathcal{O}(n^2 \cdot f)$	$2 \times g \times k_g \times d + 2 \times 4 \times k$

Avec n , le nombre de séries temporelles, ℓ , la longueur des séries temporelles, k , le nombre de noyaux voulus, k_g , le nombre de noyaux par groupe, d , la valeur maximale de dilatation, g , le nombre de groupes, f , le nombre de caractéristiques générées

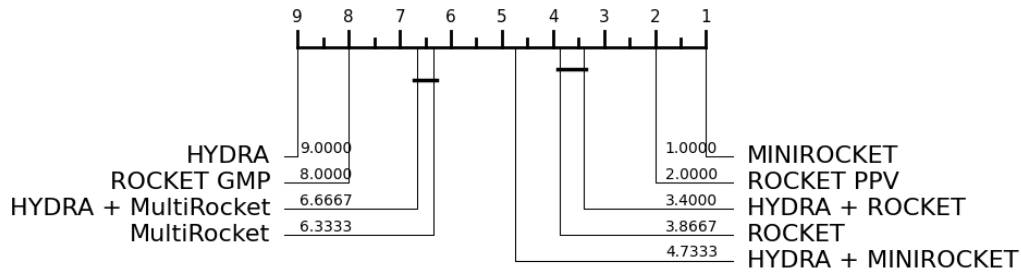


FIGURE 2 – Rang moyen des méthodes utilisant des noyaux de convolution aléatoires sur les 15 jeux de données créés

4.3 Résultats

Les résultats d'une classification multi-classes peuvent être présentés sous la forme d'une matrice de confusion. Elle comprend pour chaque classe 4 valeurs : les Vrais Positifs (VP) : les exemples prédits comme appartenant à une classe X et appartenant vraiment à cette classe, les Faux Positifs (FP) : les exemples prédits comme appartenant à la classe X mais appartenant en réalité à une autre classe, les Vrais Négatifs (VN) : les exemples prédits comme appartenant à une autre classe et appartenant vraiment à une autre classe, et enfin les Faux Négatifs (FN) : les exemples prédits comme appartenant à une autre classe mais en réalité appartenant à la classe X.

L'exactitude globale (*overall accuracy*) représente le pourcentage de bonnes classifications sur l'ensemble des données de validation.

$$\text{Exactitude} = \frac{\text{Nombre de bonnes prédictions}}{\text{Nombre total de prédictions}}$$

Le nombre de bonnes prédictions peut s'obtenir en sommant les Vrais Positifs de toutes les classes (valeurs présentes sur la diagonale de la matrice de confusion complète).

Le tableau II présente les résultats obtenus. Le matériel d'expérimentation est un Dell Precision 5480 équipé d'un CPU Intel Core i9-13900H vPro 2.6 GHz et 32GB de RAM opérant sous Debian 12.

On remarque que, de façon générale, l'écart-type de l'exactitude est très faible. Cela démontre la robustesse des méthodes testées.

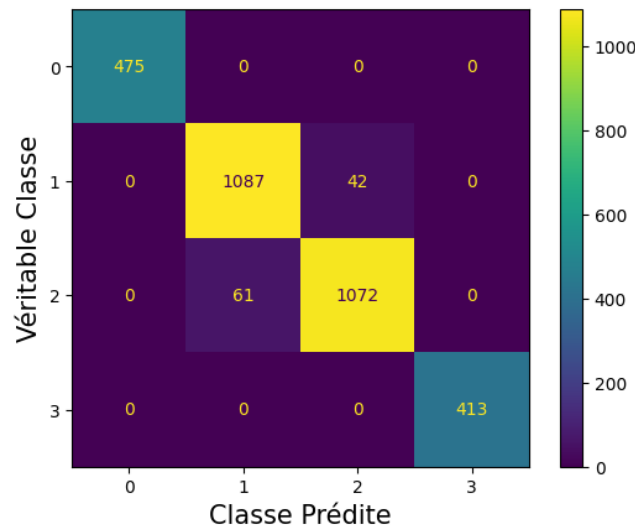
MINIROCKET domine largement les autres méthodes. Elle permet d'obtenir la meilleure exactitude moyenne avec le plus faible écart-type. De plus, les temps d'entraînement et de prédiction moyens sont très courts, tout particulièrement pour la prédiction, ce qui en fait un candidat idéal pour notre application (classification en temps réel).

La figure 2 présente le diagramme de différence critique (un outil classiquement utilisé en apprentissage automatique pour visualiser les différences significatives de performance entre des modèles) [5]. On constate que les deux meilleurs résultats sont obtenus par des modèles n'utilisant que la PPV comme opérateur (MINIROCKET et ROCKET PPV) et que les deux moins bons par des modèles qui ne l'utilisent pas (HYDRA et ROCKET GMP).

Un exemple de matrice de confusion obtenue avec MINIROCKET est donné dans la figure 3. Les sous-séquences saines (classe 0) et avec des défauts de dégradation de la barre du rotor (classe 3) n'ont généralement très peu, voir pas de FP et FN. La principale difficulté (minime) que rencontre MINIROCKET est de correctement distinguer les défauts de la bague interne et externe du roulement (resp. classes 1 et 2). Notons que cette matrice de confusion est globalement similaire (en termes de difficultés de classification) à celles obtenues avec les autres méthodes.

TABLE II – Performances des différents modèles testés sur les 15 jeux de données

Modèles testés	Exactitude moyenne (en %) et Écart-Type	Temps d'entraînement moyen (secondes)	Temps de prédiction moyen (secondes)
ROCKET	90,078 (±0,717)	241,174	52,919
ROCKET PPV	95,151 (±0,556)	226,255	52,348
ROCKET GMP	84,158 (±0,615)	216,269	50,872
MINIROCKET	96,362 (±0,491)	78,268	1,700
MultiRocket	86,603 (±0,659)	181,491	21,670
HYDRA	81,259 (±0,583)	306,856	48,222
HYDRA + ROCKET	90,250 (±0,558)	391,899	101,518
HYDRA + MINIROCKET	89,236 (±0,527)	224,708	48,755
HYDRA + MultiRocket	86,463 (±0,676)	367,636	86,593



Etats de la machine : machine saine (classe 0), défauts de bague roulement interne (classe 1) et externe (classe 2), dégradation de la barre du rotor dégradée (classe 3)

FIGURE 3 – Une des matrices de confusion obtenue pour MINIROCKET

5. Conclusion

Dans cet article, nous avons évalué la capacité de différentes méthodes de classification de séries temporelles basées sur les noyaux de convolution aléatoires à diagnostiquer l'état d'une machine tournante en temps réel à partir de l'analyse d'une phase du courant. L'une d'entre elle se distingue tout particulièrement : MINIROCKET.

Dans la continuité de ce travail, nous envisageons d'exploiter cette méthode pour le diagnostic de l'ensemble de la chaîne de traction d'un véhicule électrique.

Remerciements

Ce travail de recherche, soutenu et financé par l'ANR (Agence Nationale pour la Recherche), entre dans le cadre du Labcom (Laboratoire Commun) MYEL (MobilitY and Reliability of Electrical chain Lab) associant le LSEE, le LGI2A et la société CRITTM2A.

Références

- [1] Andrian Ceban, Remus Pusca, and Raphaël Romary. Study of rotor faults in induction motors using external magnetic field analysis. *IEEE Transactions on industrial electronics*, 59(5) :2082–2093, 2011.
- [2] Angus Dempster, François Petitjean, and Geoffrey I Webb. ROCKET : exceptionally fast and accurate time series classification using random convolutional kernels. *Data Mining and Knowledge Discovery*, 34(5) :1454–1495, 2020.
- [3] Angus Dempster, Daniel F Schmidt, and Geoffrey I Webb. MINIROCKET : A very fast (almost) deterministic transform for time series classification. In *Proceedings of the 27th ACM SIGKDD conference on knowledge discovery & data mining*, pages 248–257, 2021.
- [4] Angus Dempster, Daniel F Schmidt, and Geoffrey I Webb. HYDRA : Competing convolutional kernels for fast and accurate time series classification. *Data Mining and Knowledge Discovery*, 37(5) :1779–1805, 2023.
- [5] Janez Demšar. Statistical comparisons of classifiers over multiple data sets. *The Journal of Machine learning research*, 7 :1–30, 2006.
- [6] Johann Faouzi. Time series classification : A review of algorithms and implementations. *Machine Learning (Emerging Trends and Applications)*, 2022.
- [7] Muhammed Ali Gultekin and Ali Bazzi. Review of fault detection and diagnosis techniques for ac motor drives. *Energies*, 16(15) :5602, 2023.
- [8] Miftah Irhoumah, Remus Pusca, Eric Lefevre, David Mercier, and Raphael Romary. Detection of the stator winding inter-turn faults in asynchronous and synchronous machines through the correlation between harmonics of the voltage of two magnetic flux sensors. *IEEE Transactions on Industry Applications*, 55(3) :2682–2689, 2019.
- [9] Dileep Kumar, Muhammad Zakir Shaikh, Majid Hussain, Sanaullah Mehran, Tanweer Hussain, and Bhawani Shankar Chowdhry. Current signature dataset of three-phase induction motor under varying load conditions. <https://data.mendeley.com/datasets/gxdd74czwh/1>, 2022.
- [10] Yann LeCun and Yoshua Bengio. Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks*, 3361(10) :1995, 1995.
- [11] Matthew Middlehurst, Patrick Schäfer, and Anthony Bagnall. Bake off redux : a review and experimental evaluation of recent time series classification algorithms. *arXiv preprint arXiv :2304.13029*, 2023.
- [12] H. Sakoe and S. Chiba. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26(1) :43–49, 1978.
- [13] Chang Wei Tan, Angus Dempster, Christoph Bergmeir, and Geoffrey I Webb. MultiRocket : multiple pooling operators and transformations for fast and effective time series classification. *Data Mining and Knowledge Discovery*, 36(5) :1623–1646, 2022.