

Détection des sursauts radio rapides (FRB) basée sur une architecture neuronale hybride

Benjamin AMBLARD¹ Sara EL-BOUCH² Mauro DALLA MURA^{1,3} André FERRARI² Cédric RICHARD^{2,3}

¹Université Grenoble-Alpes, CNRS, Gipsa-Lab, 11 Rue des Mathématiques, BP 38400, Saint-Martin d'Hères, France

²Université Côte d'Azur, CNRS, Observatoire de la Côte d'Azur, 28 Avenue Valrose, BP 06108, France

³Institut Universitaire de France

Résumé – Les sursauts radio rapides, des phénomènes astrophysiques brefs et intenses, sont difficiles à détecter en raison des interférences radiofréquences. Dans cet article, nous proposons une architecture hybride combinant une tête contrastive et une tête de reconstruction, optimisée pour la détection d'anomalies. Cette approche permet de distinguer efficacement les FRBs des interférences grâce à une modification du coût contrastif et à une reconstruction sélective des données normales uniquement.

Abstract – Fast Radio Bursts are brief, intense astrophysical events whose detection is hampered by radio frequency interferences. In this work, we introduce a hybrid architecture that combines contrastive learning with selective reconstruction, specifically tailored for anomaly detection. By modifying the contrastive loss and restricting reconstruction to normal data, our approach enables robust discrimination between FRBs and interference signals.

1 Introduction

Les sursauts radio rapides (FRB) sont des phénomènes astrophysiques transitoires caractérisés par leur courte durée et leur forte intensité. Ce phénomène se traduit par des sursauts radio brillants d'une durée de quelques millisecondes à des distances extra-galactiques [1]. Découverts pour la première fois en 2007 [2], ces sursauts énergétiques demeurent d'origine inconnue. Néanmoins, ils peuvent être utilisés comme des sondes cosmologiques pour étudier le taux d'expansion de l'Univers et contraindre la composante de matière noire [3]. Une meilleure compréhension des FRBs est donc d'un grand intérêt scientifique.

La détection efficace des FRB est entravée par les interférences radiofréquences (RFI), qui génèrent de nombreux faux positifs. Les algorithmes traditionnels reposent généralement sur une série d'étapes bien définies (voir par ex. [4]). Ces étapes incluent le filtrage des interférences radio grâce à un seuil prédéfini, suivi par l'application de grilles de dispersion temporelle (DM) pour concentrer le support des données. Les données sont ensuite intégrées selon la direction fréquentielle afin de produire une série temporelle pour chaque valeur de DM. Des filtres boxcar sont appliqués à ces séries temporelles, et le rapport signal/bruit (RSB) est calculé. Les signaux dont le RSB dépasse un seuil prédéfini sont alors identifiés comme candidats FRB. Cette méthodologie, mise en œuvre par des outils tels que Presto [5] ou Heimdall [6], constitue la norme actuelle pour les recherches sur les FRBs. Cependant, de telles approches souffrent de trois limitations majeures : a) un taux élevé de faux positifs dû aux RFI, b) une dépendance critique à des paramètres d'atténuation du bruit nécessitant des ajustements manuels, et c) une incapacité à gérer le volume croissant de données générées par les instruments modernes.

Face à ces limitations, les approches issues de l'apprentissage automatique offrent des perspectives prometteuses. La majorité des recherches actuelles s'appuient sur des al-

gorithmes supervisés conçus pour classer ou identifier les candidats parmi les transitoires détectés par les algorithmes traditionnels. Par exemple, [7]-[9] utilisent des architectures de réseaux de neurones convolutifs entraînés sur des jeux de données (RFI/FRB) équilibrés. Cependant, les signaux RFI sont abondants et très diversifiés, tandis que les phénomènes transitoires astrophysiques recherchés demeurent extrêmement rares. Le déséquilibre marqué des classes limite l'efficacité des approches supervisées classiques [10], car la rareté des anomalies connues empêche les modèles d'apprentissage profond de s'adapter correctement à des signaux d'intérêt différents de ceux observés lors de l'entraînement.

D'autre part, plusieurs approches non-supervisées ont été développées, telles que les classificateurs à une seule classe, les auto-encodeurs, les réseaux adverses génératifs (GANs) et les techniques auto-supervisées [11]. Ces approches reposent souvent sur l'hypothèse que seuls des échantillons normaux sont disponibles lors de l'entraînement et que le modèle formé sur ces données devrait produire des performances dégradées face à des anomalies, permettant ainsi leur détection. Toutefois, dans certains cas, il est possible d'accéder à un nombre limité d'échantillons anormaux. Ainsi, des méthodes de détection d'anomalies en ensemble ouvert (*OSAD : Open Set Anomaly Detection*) [12]-[17] ont été proposées pour exploiter cet a priori pendant la phase d'entraînement, tout en permettant aux modèles d'être à la fois généralisables aux anomalies *inconnues* (c-à-d des échantillons provenant de classes d'anomalies non rencontrées auparavant), et d'identifier efficacement les anomalies déjà observées (similaires aux exemples utilisés pour l'entraînement).

En particulier, des travaux [18], [19] ont évalué l'efficacité de l'apprentissage contrastif pour la détection d'anomalies, démontrant que cette approche permet d'améliorer la séparation entre les échantillons normaux et anormaux, tout en généralisant face à de nouvelles anomalies. Ce paradigme a suscité un grand intérêt ces dernières années, surtout dans les approches

auto-supervisées, notamment avec l’algorithme SimCLR [20] qui a démontré que l’apprentissage auto-supervisé pouvait égaler les performances d’un modèle supervisé (ResNet-50) sans utiliser d’étiquettes. Son principe repose sur le rapprochement des représentations de différentes vues d’un même échantillon tout en éloignant celles des autres échantillons du lot. L’apprentissage contrastif supervisé (SupCon) [21] étend cette approche dans un cas supervisé.

Dans cette perspective, cet article propose une architecture hybride combinant une tête contrastive et une tête de reconstruction, conçue pour améliorer à la fois la précision et la robustesse dans un contexte caractérisé par un déséquilibre entre classes. La tête contrastive améliore les performances de classification pour mieux distinguer le bruit des signaux astrophysiques, renforçant ainsi la capacité du modèle à discriminer efficacement les FRBs dans des environnements complexes et bruités. Contrairement au coût SupCon [21] où la similarité entre les échantillons de toutes les classes est maximisée, nous modifions le coût contrastif, dans le même esprit que [16], pour ne pas rapprocher les représentations des anomalies entre elles. La tête de reconstruction du modèle est entraînée exclusivement à reconstruire les données normales. Par conséquent, durant la phase de test, les échantillons anormaux, qu’ils soient connus ou inconnus, sont mal reconstruits.

2 Méthode

Soit $\mathbf{X}_i^1$ une image de taille $n \times m$. Nous définissons un jeu de données d’images $\mathcal{X} = \{\mathbf{X}_i\}_{i=1}^{N+A}$ tel que $\mathcal{X}_n = \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N\}$ représente les données normales (bruit seul) et $\mathcal{X}_a = \{\mathbf{X}_{N+1}, \mathbf{X}_{N+2}, \dots, \mathbf{X}_{N+A}\}$ correspond à l’ensemble limité d’anomalies (signaux FRB d’intérêt) étiquetées avec $A < N$. Le problème que l’on se propose de résoudre ici s’inscrit dans le cadre de la détection d’anomalies *en ensemble ouvert* afin de détecter aussi bien les anomalies connues (similaires aux exemples dans $\mathcal{X}_a$) qu’inconnues, grâce à un score hybride.

L’architecture de notre modèle est schématisé dans la Figure 1. Elle comprend :

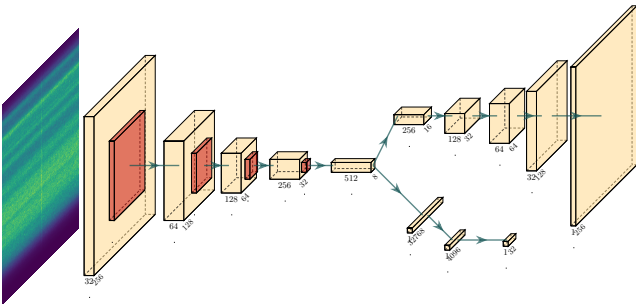


FIGURE 1 : Représentation schématique de l’architecture hybride pour la détection des sursauts radio rapides.

- Un extracteur de caractéristiques f commun aux données normales et anormales où $f(\mathbf{X})$ transforme l’image fournie en entrée $\mathbf{X}$ en une représentation latente $\mathbf{H}$.

¹Les scalaires sont représentés en police normale (par exemple, x), tandis que les vecteurs sont notés en gras minuscule ($\mathbf{x}$) et les matrices en gras majuscule ($\mathbf{X}$).

- Un décodeur représenté par une fonction $g^{(r)}(\mathbf{H})$ qui reconstruit une approximation de l’entrée originale à partir de la représentation latente avec des couches convolutives transposées (architecture miroir de l’encodeur). Le rôle du décodeur est de reconstruire uniquement les données **normales**. Il s’agit d’une architecture générique d’auto-encodeur $g^{(r)}(f(\mathbf{X}))$. La fonction de coût associée est l’erreur quadratique moyenne :

$$\mathcal{L}_{\text{rec}}(\mathbf{X}) = \|g^{(r)}(f(\mathbf{X})) - \mathbf{X}\|_F^2 \quad (1)$$

où $\|\cdot\|_F$ désigne la norme de Frobenius.

- La tête contrastive vise à apprendre un espace de représentation où la dissimilarité entre les échantillons normaux et anormaux est maximisée. Contrairement à la fonction de coût contrastive supervisée proposée par [21] qui cherche à rapprocher tous les échantillons d’une même classe, nous choisissons de ne pas maximiser la similarité entre les échantillons issus de la classe anormale dans le même esprit que [16]. Cette stratégie permet d’obtenir des représentations mieux adaptées à la détection d’anomalie, qui est intrinsèquement un problème en *jeu ouvert*. Dans ce but, chaque lot est d’abord augmenté pour générer des paires positives (autre que les éléments de la même classe) et renforcer la séparabilité des représentations. On applique donc des augmentations choisies de sorte à préserver les propriétés des FRB présents dans les images (voir partie 4). En pratique, les images de chaque lot auront deux augmentations. Un lot $\mathcal{B}$ est donc de taille $2N$. Pour chaque lot $\mathcal{B}$, composé d’échantillons normaux et anormaux, nous extrayons un sous-ensemble noté $\mathcal{B}_n$ de données normales uniquement. Chaque représentation $f(\mathbf{X}_k)$ est transformée en appliquant une tête de projection $g^{(c)}(\cdot)$ [20]. L’application $g^{(c)}(\cdot)$ est un Perceptron Multicouche (MLP) à deux couches cachées avec une fonction d’activation ReLU [20]. Enfin, les vecteurs obtenus en sortie sont normalisés. La fonction de coût est alors calculée de la manière suivante :

$$\mathcal{L}_{\text{con}}(\mathbf{X}) = \sum_{i \in \mathcal{B}_n} \frac{-1}{|\mathcal{Q}_i|} \sum_{q \in \mathcal{Q}_i} \log \frac{\exp(\mathbf{z}_i \cdot \mathbf{z}_q / \tau)}{\sum_{a \in \mathcal{B}_i} \exp(\mathbf{z}_i \cdot \mathbf{z}_a / \tau)}, \quad (2)$$

où $\mathbf{z}_k = g^{(c)}(f(\mathbf{X}_k))$, $\mathcal{Q}_i = \mathcal{B}_n \setminus \{i\}$ et $\mathcal{B}_i = \mathcal{B} \setminus \{i\}$. Enfin, le paramètre τ appelé température agit en tant que facteur d’échelle pour contrôler la sensibilité du modèle aux différences de similarité entre les vecteurs latents projetés.

L’apprentissage est réalisé de manière conjointe pour les deux têtes du réseau, avec pour fonction de coût la somme des fonctions de coût contrastif et de reconstruction :

$$\mathcal{L} = \mathcal{L}_{\text{rec}} + \mathcal{L}_{\text{con}}.$$

Enfin, pendant l’inférence, un score est attribué à chaque échantillon $\mathbf{X}$, en se basant sur les scores de chaque tête. Pour la tête de reconstruction, nous calculons le score avec $s_{\text{rec}} = \mathcal{L}_{\text{rec}}$. D’autre part, le score contrastif est défini par :

$$s_{\text{con}}(\mathbf{X}) = 1 - \frac{1}{|P|} \sum_{i \in P} \mathbf{z} \cdot \mathbf{z}_i, \quad (3)$$

où P est un échantillon aléatoire des données d’entraînement composé uniquement de données normales. On rappelle que les $\mathbf{z}_k$ sont normalisés en sortie de la tête de projection. Le score d’anomalie final attribué au candidat $\mathbf{X}$ est alors donné par $s(\mathbf{X}) = s_{\text{rec}}(\mathbf{X}) + s_{\text{con}}(\mathbf{X})$.

3 Données

Les données considérées sont des spectrogrammes issus d’observations acquises sur le Green Bank Telescope par trois instruments différents. Ces données sont mises à disposition par les auteurs de [9] et sont disponibles pour le téléchargement depuis leur dépôt GitHub². Il s’agit de représentations temps-fréquence de taille 256×256 qui ont été préalablement centrées et réduites dans le but de fournir des données indépendantes du télescope et de l’instrument. De plus, ces spectrogrammes sont dé-dispersés avec l’outil Heimdall [6]. Enfin, seuls les FRBs avec un RSB supérieur à 8 sont conservés.

Les jeux de données disponibles consistent en deux jeux annotés composés respectivement de 40.000 images pour l’entraînement et de 13.983 images pour le test, répartis de manière équilibrée en deux classes : **True** regroupe les images avec un FRB et **False** celles avec seulement des RFI. Les FRBs présents dans le jeu d’entraînement sont composés à 85% d’observations simulées, tandis que le jeu de test contient uniquement des observations réelles. Pour notre application, nous utilisons un jeu d’entraînement déséquilibré avec 80% d’images venant de la classe **False** et 20% de la classe **True**, et un jeu de test modifié pour être d’une part équilibré, et d’autre part pour que la classe **True** contienne à la fois des FRBs simulés et des Pulsars réels en proportions comparables.

4 Implémentation

L’extracteur de caractéristique ou encodeur est un réseau convolutif composé de quatre couches avec un noyau de taille 3×3 . Chaque couche convolutive intermédiaire est suivie de la fonction d’activation ReLU. Le décodeur est composé de couches convolutives transposées, organisées selon l’architecture inverse de celle de l’encodeur. Enfin, la tête contrastive est un Perceptron multi-couche avec deux couches cachées et la fonction d’activation ReLU. Pour générer les paires d’images nécessaires à l’apprentissage contrastif, les données sont augmentées par retournement horizontal. L’apprentissage a été réalisé sur 100 époques avec un taux d’apprentissage de 10^{-3} et une taille de batch de 128. Les hyper-paramètres utilisés sont $\tau = 0.07$ et $|P| = 64$. Le jeu d’entraînement a été divisé en 85% pour l’entraînement et 15% pour la validation.

5 Résultats et discussion

Pour évaluer la performance du modèle proposé, il est comparé tout d’abord à celle de ses deux sous-modèles (branche contrastive et branche de reconstruction prises individuellement), ainsi qu’à plusieurs méthodes de référence. Plus précisément, il est comparé à FETCH [9], un modèle basé sur des réseaux de neurones convolutifs couramment employé pour

la classification de transitoires astrophysiques. Il est nécessaire d’apporter une précision sur le modèle utilisé dans le cas de FETCH. Son architecture met en œuvre deux branches en parallèle, qui prennent respectivement une représentation fréquence-temps (FT), et une représentation *mesure de dispersion*-temps (DMT) des données impliquant une intégration en fréquence. La branche correspondante bénéficie donc d’un filtrage en énergie des données. Comme nous n’utilisons que la première représentation des données dans notre architecture, nous ne nous comparons qu’à la branche de FETCH correspondante, notée FETCH-FT. Ensuite, nous incluons les comparaisons avec DRA [14], une méthode qui améliore les DevNets [22] en proposant un cadre d’apprentissage sur des anomalies connues, des pseudo-anomalies et anomalies résiduelles latentes. Enfin, nous ajoutons une comparaison à un seuillage en énergie. Il est important de rappeler que les données sont filtrées pour ne contenir que des FRBs ayant un RSB à 8, ce qui favorise nettement les approches basées sur l’énergie. Pour comparer les architectures entre elles, les proportions de vrai positifs, faux positifs, vrais négatifs et faux négatifs pour chaque modèle sont résumées dans la figure 2. Nous calculons également la justesse (Accuracy, [23]), ainsi que de l’aire sous la courbe COR (ROC-AUC). Cette dernière permet de comparer les modèles indépendamment du choix de seuil. Les valeurs de ces deux indicateurs pour les trois modèles sont récapitulées dans le tableau 5.

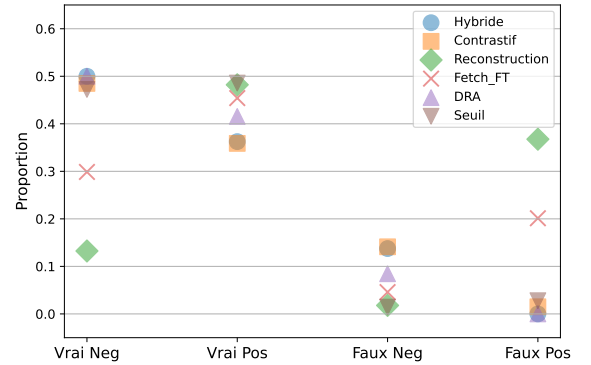


FIGURE 2 : Proportion de vrai positifs, faux positifs, vrais négatifs et faux négatifs pour chaque modèle.

Architecture	Acc (%)	ROC-AUC (%)
Hybride	86.2	85.5
Contrastive seule	84.4	86.8
Reconstruction seule	61.5	52.9
FETCH - FT	75.3	80.9
DRA	91.6	91.6
Seuil en énergie	95.8	97.7

La méthode hybride réduit le plus efficacement les faux positifs par rapport à toutes les autres approches proposées. Elle atteint une justesse (Acc) de 86.2%, surpassant nettement FETCH-FT, ainsi que la tête de reconstruction utilisée séparément. Ce résultat montre que le couplage avec l’apprentissage contrastif est très important pour améliorer la séparabilité entre les FRBs et les RFIs. Nous avons également envisagé un modèle avec deux extracteurs de caractéristiques distincts pour étudier l’intérêt d’une base commune. Ce modèle donne des

²<https://github.com/devanshkv/fetch>

résultats très proches de ceux de l'architecture hybride malgré une augmentation non négligeable du nombre de paramètres. La performance élevée du seuillage en énergie reflète l'effet du filtrage préalable des données, ce qui en fait une borne supérieure pour les méthodes fondées sur l'énergie. DRA surpasse l'ensemble des autres architectures testées. Ces résultats préliminaires ouvrent des perspectives d'amélioration, en intégrant notamment une tête de déviation [16]. -

6 Conclusion

Dans cet article, nous avons introduit une architecture hybride combinant une tête de reconstruction, exploitant les données normales pour apprendre une représentation latente robuste, et une tête contrastive, affinant la frontière de décision à partir des exemples anormaux. Les résultats expérimentaux montrent une réduction significative du taux de faux positifs par rapport à FETCH. Ces travaux ouvrent plusieurs perspectives, telles que l'intégration d'un encodeur variationnel, l'exploration de fonctions de coût contrastives adaptées aux déséquilibres de classes [24], ainsi que l'ajout d'une tête de déviation.

Références

- [1] B. ZHANG, "The physics of fast radio bursts", *Reviews of Modern Physics*, t. 95, n° 3, 2023.
- [2] D. R. LORIMER, M. BAILES, M. A. MCLAUGHLIN, D. J. NARKEVIC et F. CRAWFORD, "A Bright Millisecond Radio Burst of Extragalactic Origin", *Science*, t. 318, n° 5851, nov. 2007.
- [3] V. RAVI, N. BATTAGLIA, S. BURKE-SPOLAOR et al., "Fast Radio Burst Tomography of the Unseen Universe", *Bulletin of the AAS*, t. 51, n° 3, 420, mai 2019.
- [4] J. CORDES et M. A. MCLAUGHLIN, "Searches for fast radio transients", *The Astrophysical Journal*, t. 596, n° 2, 2003.
- [5] S. RANSOM, "PRESTO : pulsar exploration and search toolkit", *Astrophysics source code library*, 2011.
- [6] B. R. BARSDELL, M. BAILES, D. G. BARNES et C. J. FLUKE, "Accelerating incoherent dedispersion", *Monthly Notices of the Royal Astronomical Society*, t. 422, n° 1, 2012.
- [7] J. AKERET, C. CHANG, A. LUCCHI et A. REFREGIER, "Radio frequency interference mitigation using deep convolutional neural networks", *Astronomy and computing*, t. 18, 2017.
- [8] L. CONNOR et J. van LEEUWEN, "Applying deep learning to fast radio burst classification", *The Astronomical Journal*, t. 156, n° 6, 2018.
- [9] D. AGARWAL, K. AGGARWAL, S. BURKE-SPOLAOR, D. R. LORIMER et N. GARVER-DANIELS, "FETCH : A deep-learning based classifier for fast transient classification", *Monthly Notices of the Royal Astronomical Society*, t. 497, n° 2, 2020.
- [10] M. BUDA, A. MAKI et M. A. MAZUROWSKI, "A systematic study of the class imbalance problem in convolutional neural networks", *Neural networks*, t. 106, 2018.
- [11] L. RUFF, J. R. KAUFFMANN, R. A. VANDERMEULEN et al., "A Unifying Review of Deep and Shallow Anomaly Detection", *Proceedings of the IEEE*, t. 109, n° 5, mai 2021.
- [12] W. LIU, W. LUO, Z. LI, P. ZHAO, S. GAO et al., "Margin Learning Embedded Prediction for Video Anomaly Detection with A Few Anomalies.", in *International Joint Conference on Artificial Intelligence*, t. 3, 2019.
- [13] M. SALEHI, H. MIRZAEI, D. HENDRYCKS, Y. LI, M. ROHBAN, M. SABOKROU et al., "A Unified Survey on Anomaly, Novelty, Open-Set, and Out-of-Distribution Detection : Solutions and Future Challenges", *Transactions on Machine Learning Research*, n° 234, 2022.
- [14] C. DING, G. PANG et C. SHEN, "Catching both gray and black swans : Open-set supervised anomaly detection", in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022.
- [15] X. YAO, R. LI, J. ZHANG, J. SUN et C. ZHANG, "Explicit boundary guided semi-push-pull contrastive learning for supervised anomaly detection", in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023.
- [16] T. LAI, T. K. K. HO et N. ARMANFARD, "Open-Set Multivariate Time-Series Anomaly Detection", in *European Conference on Artificial Intelligence*, IOS Press, oct. 2024.
- [17] J. ZHU, C. DING, Y. TIAN et G. PANG, "Anomaly heterogeneity learning for open-set supervised anomaly detection", in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2024.
- [18] V. SEHWAG, M. CHIANG et P. MITTAL, "SSD : A unified framework for self-supervised outlier detection", in *International Conference on Learning Representations*, 2021.
- [19] J. TACK, S. MO, J. JEONG et J. SHIN, "CSI : Novelty detection via contrastive learning on distributionally shifted instances", *Advances in neural information processing systems*, t. 33, 2020.
- [20] T. CHEN, S. KORNBLITH, M. NOROUZI et G. HINTON, "A simple framework for contrastive learning of visual representations", in *International Conference on Machine Learning*, 2020.
- [21] P. KHOSLA, P. TETERWAK, C. WANG et al., "Supervised contrastive learning", *Advances in neural information processing systems*, t. 33, 2020.
- [22] G. PANG, C. SHEN et A. VAN DEN HENGEL, "Deep anomaly detection with deviation networks", in *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, 2019.
- [23] C. M. BISHOP et H. BISHOP, *Deep Learning : Foundations and Concepts*. Springer International Publishing, 2024.
- [24] D. MILDENBERGER, P. HAGER, D. RUECKERT et M. J. MENTEN, "A Tale of Two Classes : Adapting Supervised Contrastive Learning to Binary Imbalanced Datasets", in *Proceedings of the Computer Vision and Pattern Recognition Conference*, 2025.