

Aide à la classification d'objets spatiaux à partir d'un unique TLE

Théo Dorlin¹, Alex Elmaian², Thibault de-la-Villegeorges³, Jérôme Dufour⁴

¹ARIANEGROUP, HELIX

theo.dorlin@ariane.group

Résumé

Nous proposons une méthode de caractérisation probabiliste d'objets spatiaux résidents à partir d'une information volontairement minimale : un unique TLE. La chaîne extrait des paramètres orbitaux, projette l'espace par PaCMAP, puis estime des scores de cohérence par noyaux de densité pour trois familles de classes : type, application principale et constellation. L'évaluation sur un corpus apparié TLE-métadonnées de plus de 25 000 objets montre que l'information orbitale seule suffit à dégager des régularités utiles, en particulier pour les classes fortement structurées. Les résultats montrent la pertinence d'un couplage PaCMAP+KDE comme brique d'aide à l'analyse orbitale et ouvrent la voie à une extension temporelle et multimodale.

Mots-clés

Objets spatiaux (RSO), TLE, PaCMAP, estimation de densité, classification probabiliste.

Abstract

We present a probabilistic method for characterizing resident space objects from intentionally minimal information: a single TLE. The pipeline extracts orbital parameters, projects the feature space with PaCMAP, and estimates coherence scores with kernel density models for three classes families: object type, main application, and constellation. Evaluation on a matched TLE-metadata corpus of more than 25,000 objects shows that orbital information alone already carries useful structure, especially for highly organized classes. These results support the relevance of a PaCMAP+KDE combination as a first operational aid for orbital analysis and motivate future temporal and multimodal extensions.

Keywords

resident space objects, TLE, PaCMAP, kernel density estimation, probabilistic classification.

1 Introduction

Les activités de surveillance de l'espace nécessitent d'orienter rapidement l'identification d'un objet orbital nouvellement catalogué avec une information immédiatement disponible souvent très limitée. Dans de nombreux cas, la seule donnée accessible à court terme est un set de mesures et les éléments orbitaux estimés sur la base de ce set de mesure. Un Two-Line Elements (TLE) est communément dérivé de ces informations orbitales car

facile à diffuser. Cette information minimale porte une signature dynamique : altitude, inclinaison, excentricité et géométrie orbitale qui traduisent partiellement la mission d'un objet, son statut et son éventuelle appartenance à une constellation.

L'objectif est de construire une chaîne de caractérisation probabiliste d'objets spatiaux (Resident Space Objects, RSO) qui exploite uniquement cette signature orbitale. Nous cherchons non pas à produire une décision dure et définitive, mais à fournir des classes plausibles et des scores de cohérence interprétables afin d'aider un analyste à prioriser son enquête. Les trois niveaux de description visés sont le type de l'objet, son application principale et son appartenance éventuelle à une constellation.

Nos contributions sont triples. Premièrement, nous formalisons un protocole de caractérisation fondé uniquement sur cinq variables orbitales dérivées des TLE. Deuxièmement, nous utilisons PaCMAP comme outil de structuration de l'espace orbital, en alternative aux approches de plus classiques [2, 5, 6]. Troisièmement, nous montrons qu'un modèle simple d'estimation de densité à noyau peut déjà produire des résultats utiles dans un cadre opérationnel où l'information d'entrée est limitée.

2 Travaux antérieurs

2.1 Réduction de dimension pour les RSO

Les représentations réduites sont fréquemment utilisées pour explorer des ensembles de RSO. Les méthodes t-SNE et UMAP offrent des visualisations expressives, mais leur stabilité, leur sensibilité aux hyper-paramètres et leur passage à l'échelle restent délicats sur des jeux massifs [2, 5]. PaCMAP a été proposé pour mieux préserver simultanément les voisinages locaux et la structure globale, ce qui est particulièrement intéressant pour des ensembles orbitaux organisés en familles ou en constellations [6].

2.2 Caractérisation probabiliste

En contexte opérationnel, des règles déterministes fondées sur quelques seuils orbitaux ne suffisent pas dès que les classes se recouvrent ou que les métadonnées sont absentes. Les travaux sur la caractérisation automatique de RSO montrent l'intérêt des approches d'apprentissage pour extraire des signatures de mission et de comportement [1]. Dans notre cadre, l'estimation de densité à noyau constitue un compromis intéressant : elle reste non paramétrique, s'adapte à des formes de classes variées et fournit des scores directement interprétables

dans l'espace réduit[3].

3 Données et représentation orbitale

3.1 Corpus de données

Le corpus agrège les TLE issus de Space-Track, et du catalogue Helix couvrant le premier trimestre 2025. Les deux sources totalisent 27 909 identifiants NORAD distincts. Après appariement avec la base de métadonnées consolidée par Helix, 25 563 objets disposent d'une orbite et d'annotations exploitables pour l'évaluation.

La famille de classes « Type » ramène les objets à trois classes : Payload (13 809 RSO), Débris (11 488) et Unknown (266). La famille « MainApplication » fait apparaître des classes comme Unknown (11 748 RSO) et Commercial (10 273 RSO). Enfin, la famille « Constellation » conserve les constellations suffisamment peuplées et regroupe les plus petites dans une classe « <30 » qui regroupe les constellations de moins de 30 éléments.

3.2 Variables orbitales

À partir de chaque TLE, nous calculons des éléments orbitaux et retenons cinq variables numériques : l'inclinaison, l'excentricité, la longitude du nœud ascendant, l'argument du périégée et le demi-grand axe. Ce choix reste volontairement parcimonieux afin de préserver la généralité de la chaîne et sa compatibilité avec une situation d'enquête où aucune autre modalité n'est encore disponible : seulement la dernière désignation TLE connue est considérée [4].

4 Méthode proposée

4.1 Vue d'ensemble

Les TLE et les métadonnées sont d'abord associées et nettoyés, puis transformés en descripteurs orbitaux. L'espace obtenu est projeté dans un plan PaCMAP ; enfin, un estimateur de densité à noyau est appris pour chaque classe de chaque famille. La sortie n'est pas une classe unique imposée, mais un ensemble de scores de cohérence permettant de hiérarchiser les hypothèses.

4.2 Réduction par PaCMAP

PaCMAP est ici utilisé comme une étape de structuration en vue de la classification. Le modèle de réduction est appris sur la partition d'entraînement puis appliqué sans réapprentissage à la partition de test. On retient une projection en dimension 2, ce qui favorise à la fois l'interprétation géométrique et la stabilité de la modélisation KDE en aval.

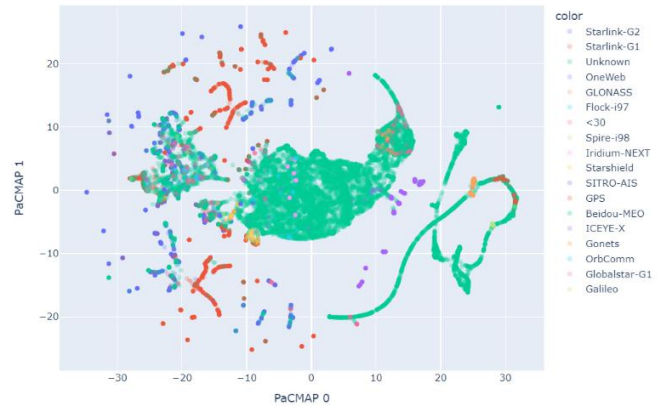


Figure 1. Projection PaCMAP et classification pour la famille de classes « Constellation ».

4.3 Estimation de densité et score de cohérence

Pour chaque famille de classes, un KDE gaussien est ajusté séparément dans l'espace réduit. La bande passante est choisie à partir d'un quantile de distances intra-classe, puis validée par croisée interne. Les densités obtenues sont normalisées par un quantile de référence afin de produire des scores bornés entre 0 et 1. Ces scores ne constituent pas des probabilités calibrées au sens bayésien strict, mais des indicateurs de compatibilité particulièrement lisibles pour un usage d'aide à la décision.

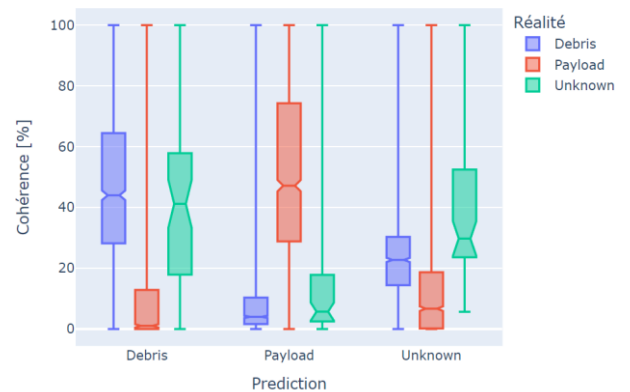


Figure 2. Score de cohérence prédit sur l'ensemble de test pour la famille de classes « Type »

4.4 Configuration

La base est divisée en sets d'apprentissage et de test selon une partition 80/20. Les sorties évaluées sont des matrices de confusion normalisées, des boîtes de cohérence (Figure 2) et plusieurs métriques par classe : précision, accuracy, spécificité, sensibilité et F1 (Figure 3).

5 Résultats

5.1 Vue d'ensemble

La famille de classes « Constellation » présente le meilleur compromis global avec un macro-F1 pondérée par la prévalence de 81%, 75% pour « Type » et 72% pour « MainApplication ». Ces résultats montrent que la seule

structure orbitale extraite d'un TLE contient déjà un signal utile pour l'orientation de l'analyse. La performance décroît lorsque la granularité sémantique augmente, ce qui confirme l'intuition physique : les paramètres orbitaux suffisent souvent à différencier une grande famille ou une constellation très structurée, mais deviennent moins déterminants dès que le RSO encode une mission plus abstraite ou moins corrélée à l'orbite.

5.2 Famille de classes « Type »

La Figure 3 présente les métriques de performances classique pour les différentes classes de la famille « Type ». La classe Payload est la mieux reconnue avec 81 % de F1-score. Debris et Unknown atteignent respectivement 71 % et 9 % de F1-score.

Debris	82	77	88	64	72
Payload	92	82	94	72	81
Unknown	5	76	76	54	9
Mean	60	78	86	63	54
WeightedMean	86	80	91	67	75
	Precision	Accuracy	Specificity	Sensistiviy	F1

Figure 3. Matrice donnant les métriques de performance normalisée [%] pour la famille de classes « Type » sur le jeu de test.

L'erreur dominante concerne la confusion entre Debris et Unknown, ce qui est cohérent avec la pauvreté de l'information disponible pour des objets mal documentés ou atypiques.

Sur le plan applicatif, cette performance est déjà utile : elle permet de séparer une large part des charges utiles des objets plus ambigus, et donc de hiérarchiser les analyses humaines. La classe Unknown ne doit cependant pas être interprétée comme une catégorie physique homogène ; elle mélange au contraire absence d'annotation, objets marginaux et signatures orbitales difficiles à relier à une classe métier claire.

5.3 MainApplication et Constellation

Pour MainApplication, le meilleur F1 est observé sur la classe Commercial (80%) qui possède des signatures orbitales relativement cohérentes. Les catégories Amateur (F1 = 10%) et Civil (F1 = 26%) restent plus difficiles à séparer, avec des effectifs faibles et des régimes orbitaux moins spécifiques. La cible Constellation fournit en revanche des indications très instructives sur la structure

de l'espace réduit permettant à certaines constellations d'être quasiment parfaitement reconnues, par exemple Galileo (F1 = 94%) ou OneWeb (F1 = 97%). Les familles Starlink-G1 et Starlink-G2 restent correctement identifiées mais avec des F1-scores plus modérés, respectivement 87% et 79%. Ces écarts entre familles de classes montrent bien que la méthode valorise avant tout les organisations orbitales stables et fortement contraintes, alors que les classes plus diffuses, moins peuplées ou agrégées dans une classe Unknown demeurent naturellement plus difficilement identifiables.

6 Discussion et limites

Trois enseignements se dégagent de cette étude. Premièrement, l'orbite seule porte une information exploitable pour l'orientation d'une enquête orbitale. Deuxièmement, PaCMAP se comporte ici comme un véritable prétraitement de classification et non comme un simple outil de visualisation : la géométrie obtenue est suffisamment stable pour accueillir une modélisation KDE simple. Troisièmement, la qualité de séparation dépend fortement du niveau sémantique visé et de la structuration physique des classes.

Les limites sont néanmoins importantes. L'entrée ne contient qu'un TLE instantané et ne modélise ni historique orbital, ni manœuvres, ni données photométriques. La qualité des labels externes conditionne également la pertinence des résultats, notamment pour MainApplication. Enfin, les scores fournis sont des scores de cohérence normalisés et non des probabilités calibrées ; un mécanisme explicite d'abstention ou de calibration constituerait une extension naturelle.

À court terme, les pistes les plus prometteuses sont l'ajout de variables liées à la signature des RSO dans les mesures ayant permis l'estimation du TLE (comme la magnitude optique du RSO), ainsi que la comparaison avec des classifieurs probabilistes calibrés plus riches. Ces extensions devraient améliorer la robustesse du système et rendre la caractérisation plus informative dans les cas ambigus.

7 Conclusion

Nous avons présenté une méthode de caractérisation probabiliste de RSO fondée exclusivement sur des paramètres orbitaux dérivés des TLE. Malgré cette entrée minimale, la chaîne proposée distingue correctement le type des objets, fournit des indications utiles sur l'application principale et reconnaît plusieurs constellations fortement structurées. Le couplage PaCMAP+KDE apparaît ainsi comme une brique crédible pour l'aide à l'analyse orbitale.

Cette approche ne vise pas à remplacer l'expertise humaine mais à l'outiller. Elle peut servir de filtre initial, de moteur de priorisation et de support d'explicitation des voisinages orbitaux. Elle fournit enfin un socle cohérent pour une future chaîne multimodale liant données orbitales et de signature de caractérisation des objets spatiaux résidents.

Références

- [1] M. D. Hejduk et al. Resident Space Object Characterization via Machine Learning. AMOS Conference, 2020.
- [2] L. McInnes, J. Healy and J. Melville. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction, 2018.
- [3] D. W. Scott. Multivariate Density Estimation: Theory, Practice, and Visualization. Wiley, 2015.
- [4] D. A. Vallado. Fundamentals of Astrodynamics and Applications. Springer, 2013.
- [5] L. van der Maaten and G. Hinton. Visualizing Data using t-SNE. Journal of Machine Learning Research, 9:2579-2605, 2008.
- [6] Y. Wang, H. Huang, C. Rudin and Y. Shaposhnik. PaCMAP: Pairwise Controlled Manifold Approximation for Dimensionality Reduction. Journal of Machine Learning Research, 22(201):1-29, 2021.