



Encodage 2D de séries temporelles d'imagerie satellitaires pour l'amélioration

Azza Abidi, Dino Ienco, Ali Ben Abbas, Imed Riadh Farah

► To cite this version:

Azza Abidi, Dino Ienco, Ali Ben Abbas, Imed Riadh Farah. Encodage 2D de séries temporelles d'imagerie satellitaires pour l'amélioration. TAIMA 2023 - 12e Traitement et Analyse de l'Information : Méthodes et Applications, May 2023, Hammamet, Tunisia. hal-04689663

HAL Id: hal-04689663

<https://hal.science/hal-04689663v1>

Submitted on 5 Sep 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/371276319>

Encodage 2D de séries temporelles d'imagerie satellitaires pour l'amélioration de la classification de l'occupation des sols

Conference Paper · June 2023

CITATIONS

0

READS

66

4 authors:



[Azza Abidi](#)

Université de la Manouba

5 PUBLICATIONS 18 CITATIONS

SEE PROFILE



[Dino Ienco](#)

French National Institute for Agriculture, Food, and Environment (INRAE)

212 PUBLICATIONS 3,182 CITATIONS

SEE PROFILE



[Ali Ben Abbes](#)

83 PUBLICATIONS 1,400 CITATIONS

SEE PROFILE



[Imed Riadh Farah](#)

Université de la Manouba

316 PUBLICATIONS 3,193 CITATIONS

SEE PROFILE

Encodage 2D de séries temporelles d'imagerie satellitaires pour l'amélioration de la classification de l'occupation des sols

Azza Abidi^{1;2}, Dino Ienco², Ali Ben Abbes¹, et Imed Riadh Farah¹

¹ Laboratoire RIADI
École Nationale des Sciences de l'Informatique
Université de La Manouba, Tunisie
`azza.abidi@ensi-rnu.tn, ali.benabbes@yahoo.fr`
`, imedriadh.farah@mse.uma.tn`
² UMR, TETIS
Institut national de la recherche agronomique
Montpellier, France
`dino.ienco@inrae.fr`

Abstract Dans cet article, nous proposons Encodeep, une nouvelle méthode pour classer l'occupation des sols à partir de séries temporelles d'images satellites (STIS). Notre objectif est d'explorer la transformation 2D des STIS multivariée à l'aide de plusieurs techniques d'encodage, notamment le Gramian Angular Summation Field (GASF), le Gramian Angular Difference Field (GADF), le Markov Transition Field (MTF) et le Recurrence Plot (RP). Cette méthode permet de dériver des images à partir d'une combinaison de techniques d'encodage, qui seront ensuite utilisées en entrée d'un modèle de classification de réseau de neurones convolutifs (CNN). Nous avons mené des expériences sur une sous-région du département de la Dordogne, dans le sud-ouest de la France en utilisant des données STIS de Sentinel-2. Nous avons comparé les résultats obtenus par notre méthode avec ceux obtenus par plusieurs approches standard utilisées pour la cartographie de la couverture terrestre et de l'occupation du sol basée sur les STIS. Nos résultats ont démontré la qualité de cette méthode par rapport aux approches standards pour la cartographie de l'occupation des sols à partir des STIS.

Mots clés Deep Learning, Réseaux de Neurones Convolutionnels (CNN), Séries Temporelles Multivariées, Encodage 2D, classification.

Abstract In this paper, we present Encodeep, a new method for Land Use Land Cover (LULC) mapping from Satellite Image Time Series (SITS). Our focus is on the 2D encoding of SITS using various encoding techniques, including Gramian Angular Summation field (GASF), Gramian angular difference field (GADF), Markov Transition Field (MTF), and Recurrence Plot (RP) to improve the classification process. This method allows us to derive multivariate images that are used as input for a convolutional neural network (CNN) classification model. We conducted experiments on a real datasets named Dordogne, using Sentinel-2 SITS data. We compared the results obtained by our method with those obtained by several standard approaches used for land cover mapping based on SITS. Our results demonstrated the superiority of this method over standard approaches for LULC mapping from SITS.

Key words Deep Learning, Convolutional Neural Networks (CNN), Multivariate Time-Series, Encoding representation, classification,.

1 Introduction

Les programmes d'observation de la Terre ont accumulé une grande quantité de données de télédétection au fil des ans. La mission Sentinel, dans le cadre du programme Copernicus, fournit des images satellites qui permettent de créer des STIS pour surveiller la surface terrestre. Cependant, l'utilisation de ces données est complexe en raison de leur nature multivariée. Des approches de classification ont été proposées pour cartographier l'occupation des sols, mais leur mise en œuvre reste difficile [5,4]. Pour surmonter ces défis, une solution courante consiste à encoder les STIS en image 2D afin d'extraire des informations temporelles et spatiales à partir des données de télédétection. Cette approche permet d'obtenir des représentations de données plus efficaces pour améliorer les résultats de la cartographie de la couverture terrestre et de l'occupation du sol [7]. Pour ce faire, plusieurs techniques de transformation des STIS en images 2D ont été proposées, notamment le Gramian Angular Field (GAF) [12], le Markov Transition Field (MTF) [12] et le Recurrence Plot (RP) [2]. Regardant la technique GAF qui comprend à la fois la méthode GASF et GADF [12], Chen et al. ont présenté GAF-CNN, une méthode pour identifier automatiquement les motifs de chandelier. Ce travail consiste à encoder la ST en image 2D en utilisant la méthode GAF. Les résultats de classification de ces images en utilisant un CNN ont montré que l'approche proposée est plus performante que le modèle de mémoire à court terme (LSTM). En ce qui concerne la combinaison de plusieurs techniques d'encodage, Buz et al. [3] ont proposé une approche d'encodage des signaux de ST en utilisant trois transformations mathématiques pour classer les objets sous-marins. Dans cette approche, les ST multivariées ont été encodées en format 2D en utilisant les techniques GASF/GADF et MTF avant d'appliquer une technique de fusion de canaux pour créer une nouvelle image. Comparée à d'autres études, la transformation d'encodage proposée pour les ST évite la nécessité d'une extraction de caractéristiques et permet d'obtenir de meilleurs résultats de classification en utilisant des algorithmes d'apprentissage en profondeur sur les données sonar.

La transformation des STIS en représentation d'images a amélioré la précision de la cartographie de l'occupation des sols. Cependant, la plupart des études ont examiné l'encodage de signaux de ST univariés en tant qu'images 2D, interprétant différemment les dynamiques spectrales et temporelles dans chaque plage de fréquence.

Nous proposons dans ce papier une nouvelle méthode de cartographie de l'utilisation et de l'occupation des sols qui utilise une transformation d'image 2D pour intégrer les informations temporelles et spectrales des données STIS multivariées. Contrairement aux études précédentes qui ont principalement examiné l'encodage de signaux ST univariés en tant qu'images 2D, nous avons exploré quatre méth-

odes d'encodage d'images: MTF[12], GASF[12], GADF[12] et RP[2]. Ces méthodes ont été largement étudiées dans la littérature et ont montré des performances remarquables dans la classification de ST multivariées. En utilisant ces techniques d'encodage, les données STIS multivariées peuvent créer une représentation globale de l'évolution temporelle de la région étudiée. Ce qui permet d'exploiter et à réutiliser des cadres de vision par ordinateur largement établis dans le contexte de la classification de ST multivariées. A la connaissance des auteurs, c'est la première fois que des techniques encodées en 2D pour des données ST multivariées sont utilisées dans le contexte de la classification de données de télédétection.

2 Méthode

Notre méthode se compose de trois étapes. La première consiste à encoder les données STIS multivariées d'origine en images 2D à l'aide de quatre méthodes d'encodage de ST. La deuxième étape vise à fusionner les différentes informations encodées pour obtenir une représentation unique en combinant les images des quatre techniques d'encodage en une seule. La troisième étape est dédiée à la prise de décision finale, réalisée à l'aide d'un modèle CNN. La Figure 1 présente la méthodologie proposée dans cet article. Le principal avantage de la méthode proposée est l'utilisation directe de plusieurs schémas d'encodage 2D sur le signal ST multivariée plutôt que considérer chaque ST univariée indépendamment, ce qui permet de modéliser les interactions entre les variables sur lesquelles la ST multivariée est définie.

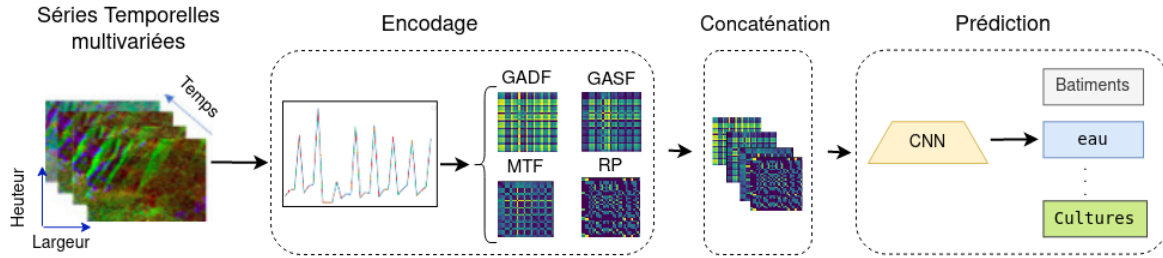


Figure 1. Aperçu complet du méthode proposée

2.1 Encodage d'images

Cette partie permet d'encoder les STIS multivariées en image 2D à l'aide de quatre techniques d'encodage. Pour cela, les données brutes des STIS sont aplaties. Les

valeurs spectro-temporelles des STIS sont combinées pour créer une représentation d'image 2D, permettant ainsi une reconnaissance visuelle et une classification plus efficaces que celles fournies par un signal 1D de la ST seul. Cette technique conduit à des représentations plus résilientes aux bruits et aux valeurs aberrantes dans les données pour plusieurs raisons. Tout d'abord, lorsqu'une méthode de lissage ou d'aplanissement est appliqué aux données, cela permet d'atténuer les variations rapides ou imprévues dans les valeurs, ce qui peut réduire l'impact des valeurs aberrantes. Ensuite, les ST sont souvent caractérisées par une corrélation temporelle entre les observations, ce qui signifie que les observations voisines sont généralement similaires. Ainsi, même si une valeur aberrante apparaît dans une observation individuelle, cela peut avoir moins d'impact sur la tendance générale de la STIS en raison de la similarité des observations voisines. La Figure 2 illustre la principale différence entre les techniques d'encodage de STIS multivariées standards et notre méthode.

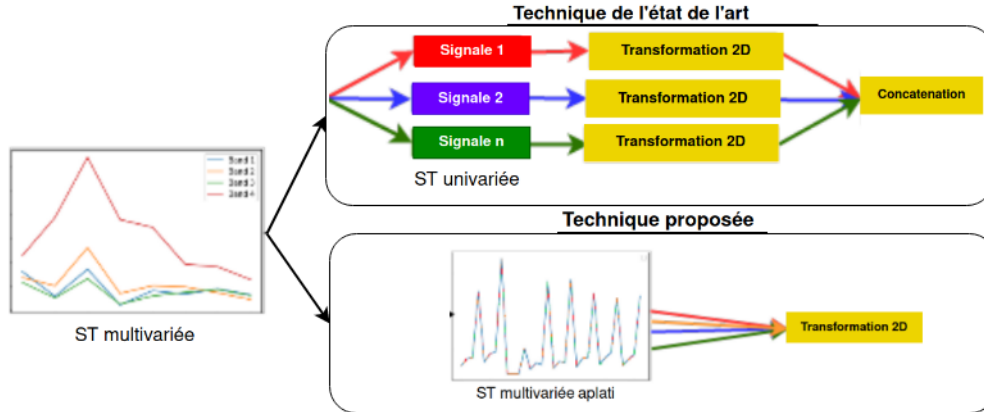


Figure 2. Aperçu de technique d'encodage de l'état de l'art et le technique proposé par Encodeep

2.2 Concaténation d'images

Après avoir appliqué des transformations sur les données ST, nous avons obtenu quatre images représentant des matrices de dimension $N \times N$. En combinant les sorties obtenues, nous avons construit une image finale qui sert de nouvelle représentation STIS sous la forme d'un tenseur $X \in R^{N \times N \times 4}$. La forme de ce tenseur dépend de différents paramètres tels que la largeur et la hauteur de l'image, le nombre de bandes de l'image STIS, ainsi que le nombre de techniques d'encodage 2D considérées. En utilisant différentes techniques d'encodage 2D complémentaire, nous

sommes en mesure d'obtenir une représentation plus complète des ST contenant des informations riches qui peuvent être exploitées pour la classification d'image.

2.3 Classification avec CNN

Suite à une étude comparative menée entre plusieurs modèles d'apprentissage en profondeur tels que DenseNet, MobileNet, ResNet et VGG afin de déterminer le modèle le plus précis pour l'analyse d'informations d'image obtenue à partir des ST encodées en 2D [6], nous avons choisi d'utiliser l'architecture ResNet-50, car elle s'est avérée être la plus précise pour la tâche de classification d'images [6].

Le modèle ResNet-50 [9] est particulièrement intéressant en raison de son utilisation de connexions résiduelles. Les connexions résiduelles sont des connexions qui permettent au gradient (la mesure de la pente de la courbe d'erreur de prédiction) de refluer vers une couche intermédiaire du réseau. Cela signifie que lors de l'entraînement du modèle, les informations peuvent être réutilisées à plusieurs reprises à différents niveaux du réseau, plutôt que d'être simplement transmises en avant. En conséquence, cela permet une meilleure évolution du modèle pendant l'entraînement et une amélioration de la précision de la prédiction finale du modèle. Ainsi, l'utilisation de connexions résiduelles dans le modèle ResNet-50 est un facteur clé de sa performance en tant que classificateur d'image.

3 Expérimentation

3.1 Matériels

Les évaluations expérimentales ont été menées sur un site d'études couvrant de grandes zones spatiales avec des caractéristiques de couverture terrestre contrastées et associé à des STIS Sentinel-2. Les données STIS ont été collectées entre Janvier et Décembre 2016. La ST est composée de 23 images multi-spectrales, chacune ayant une taille de $5\,578 \times 5\,396$ pixels correspondant à $3\,010\text{ Km}^2$. Quatre bandes radiométriques (Bleu (B2), vert (B3), rouge (B4) et proche-infrarouge (B8) avec une résolution spatiale de 10 mètres ont été considérées et acquises via la plateforme THEIA¹. Des scènes avec moins de 50 % de pixels nuageux ont été considérées au niveau 2A (réflectance du sommet de la canopée). Les valeurs des pixels nuageux ont été interpolées en utilisant une méthode d'interpolation temporelle à partir des dates sans nuages des jours précédents et suivants. Les données ont été obtenues à partir

Les données de Vérité Terrain (VT) ont été collectées à partir de deux sources, notamment: le Registre Parcellaire Graphique (RPG) fourni par l'Agence française

1. <http://theia.cnes.fr>

de services et de paiement et de La base de données topographiques (BD-TOPO)² de l’Institut géographique national (IGN) français, qui comprend les entités géographiques et administratives du territoire national. L’encodage a été effectué sur le signal de la STIS multivariée aplati. En fait, la forme des données a été transformée de $(x, 23, 4)$ à $(x, 92)$ pour créer des images de dimensions $(92, 92)$. Nous avons par la suite re-dimensionné ces images encodées en des images de taille 32×32 pour pouvoir les utiliser dans le classifieur choisi.

3.2 Protocol expérimental

Afin de garantir une évaluation impartiale, nous avons pris soin d’attribuer exclusivement à une seule partition de données (entraînement, validation ou test) les pixels appartenant à un même polygone de référence. Les proportions d’objets pour chacune de ces partitions étaient respectivement de 50%, 20% et 30%. Pour les données STIS, nous avons normalisé les valeurs de ST par bande dans l’intervalle $[0, 1]$. Concernant l’entraînement des modèles basés sur des réseaux de neurones, nous avons utilisé l’optimiseur ADAM sur 100 époques, avec un taux d’apprentissage de $1e-4$. Pour le classificateur RF, nous avons ajusté sur l’ensemble 20, 40, 60, 80, 100 avec une profondeur maximale de 100, 200, 300, 400 et 500 pour optimiser le modèle. Nous avons finalement sélectionné le modèle offrant la meilleure performance sur l’ensemble de validation pour effectuer la classification sur l’ensemble du test.

La performance de classification est évaluée en utilisant les métriques de précision global (Accuracy), de la F-mesure et de Kappa. Toutes les mesures ont été calculées sur cinq répartitions aléatoires du jeu de données. Toutes les expériences ont été menées sur une station de travail équipée de 256 Go de RAM, d’un processeur Intel (R) Xeon (R) CPU E5-2667 v4 @ 3,20 GHz et de quatre GPU TITAN X

3.3 Détails sur la configuration des expériences

En évaluant les informations de ST à la fois originales et transformées, nous comparons notre approche à plusieurs techniques de l’état de l’art. Nous avons effectué une comparaison avec:

- Classification sur la ST multivariée, où le signal d’origine est utilisé directement en entrée pour les techniques de classification de ST.
- STUV: classification sur la ST transformée [7,13], où le codage est effectué sur chaque signal univarié. Ensuite, les résultats ont été agrégés pour reconstruire une seule image qui se réfère à la ST multivariée initiale.

2. <https://geoservices.ign.fr/documentation/donnees/vecteur/bdtopo>

- STMV: classification avec la méthode proposée, où le codage a été effectué sur l'ensemble de la ST multivariée et sur des images 2D redimensionnées en forme des images de 32×32 pixels.

Pour évaluer l'efficacité de notre proposition, nous avons réalisé une série d'expérimentations supplémentaires en portant une attention particulière à la combinaison des techniques d'encodage utilisées. Pour la combinaison des deux techniques GASF et GADF dans notre cas, nous avons utilisé le terme `EncodeepCmb2`. Quant à la combinaison des trois méthodes GASF, GADF et MTF, nous l'avons appelée `EncodeepCmb3`. Nous avons ensuite sélectionné plusieurs méthodes concurrentes ayant des caractéristiques différentes et complémentaires afin d'évaluer la qualité de Encodeep sur le site d'étude, telles que:

- Les Forêts Aléatoires (RF) [10]. Ce classifieur est couramment utilisé pour sa capacité à traiter la classification de données STIS.
- Le CNN unidimensionnel (CNN-1D) [1]. Cette méthode classe les ST multivariés en utilisant un CNN-1D où les opérateurs de convolution sont déployés sur la dimension temporelle des données STIS.
- Fully Connected Gated Recurrent Unit (FCGRU) [8]. Dans cette approche, les auteurs effectuent une analyse STIS à travers une cellule GRU modifiée dans laquelle le réseau neuronal récurrent a été enrichi grâce à un mécanisme d'attention personnalisé.
- Le CNN Temporel (TempCNN) [11]. C'est une technique basée sur des convolutions 1D pour extraire des caractéristiques temporelles dans des séquences de données. Cette approche permet au CNN Temporel de capturer des motifs complexes dans les données temporelles et de les utiliser pour des tâches telles que la classification

3.4 Résultats

Le Tableau 1 présente les résultats moyens pour le site d'étude de la Dordogne. Les meilleurs résultats sur le signal STIS et les STIS multivariées transformées en 2D sont présentés en gras, tandis que les extraits de cartes d'utilisation des sols sont reportés dans la Figure 3.

D'un point de vu général, l'utilisation de Encodeep a amélioré les performances de classification par rapport aux approches classiques utilisant la transformation des ST univariées ou des signaux ST bruts, selon les résultats rapportés dans Table 1. En particulier, nous avons observé une augmentation de l'exactitude de classification de plus de 5% et une augmentation significative du F1-score. Les résultats montrent également que la combinaison de Encodeep avec le modèle ResNet-50 est une approche efficace pour la classification des données STIS. Nous avons également comparé notre approche avec d'autres méthodes de classification couramment

Table 1. Performances réalisées par le modèle ResNet-50 en utilisant le signal ST original, le signal ST univarié transformé et Encodeep.

Type de données	Transform.	Modèle	Précision	F1-score	Kappa
Séries temporelles		TempCNN	86.43 ± 1.74	86.42 ± 1.86	84.14 ± 2.10
		CNN-1D	86.73 ± 1.58	86.74 ± 1.70	84.42 ± 2.30
		RF	86.00 ± 1.95	86.14 ± 1.75	84.08 ± 3.48
		FCGRU	86.54 ± 1.94	86.55 ± 1.82	84.20 ± 1.12
STUV transformée en 2D	GASF	Resnet-50	81.78 ± 1.33	81.45 ± 1.44	76.78 ± 1.41
	GADF		82.17 ± 1.24	82.11 ± 1.28	77.22 ± 1.20
	MTF		80.84 ± 1.39	80.96 ± 1.34	76.11 ± 1.37
	RP		82.48 ± 1.42	82.86 ± 1.37	78.72 ± 1.13
	Encodeep _{Cmb2}		82.65 ± 1.48	82.48 ± 1.36	78.85 ± 1.34
	Encodeep _{Cmb3}		82.77 ± 1.33	82.91 ± 1.35	79.02 ± 1.36
	Encodeep _{STUV}		83.17 ± 1.48	82.98 ± 1.41	78.46 ± 1.44
STMV transformée en 2D	GASF	Resnet-50	84.45 ± 1.88	84.67 ± 1.70	78.67 ± 1.81
	GADF		85.67 ± 1.81	85.52 ± 1.20	79.90 ± 1.97
	MTF		81.07 ± 1.84	79.68 ± 1.97	76.51 ± 1.66
	RP		85.92 ± 1.80	85.94 ± 1.34	80.01 ± 1.87
	Encodeep _{Cmb2}		86.11 ± 1.66	85.96 ± 1.27	80.16 ± 1.45
	Encodeep _{Cmb3}		87.06 ± 1.34	87.75 ± 1.19	82.14 ± 1.32
	Encodeep		90.22 ± 1.05	90.26 ± 1.14	85.86 ± 1.20

utilisées pour les données STIS, dont les résultats montrent que notre méthode a surpassé ces méthodes pour toutes les mesures d'évaluation.

Plus en détail, nous avons également étudié l'efficacité de différentes techniques d'encodage, en ce qui concerne la technique RP s'est avérée la plus performante avec une précision de 85,92% grâce à sa capacité de détecter des motifs récurrents cachés. En revanche, GADF a fourni une précision un peu dégradée. Ces résultats sont obtenus à cause de sa pauvre capacité à modéliser la connexion temporelle entre les observations à une échelle de temps différent. Aussi, la méthode MTF a obtenu la plus faible performance avec une précision de 81%, d'où cette méthode semble insuffisante pour soutenir les tâches de classification liées aux données STIS.

Les extraits de cartes obtenus à partir de Encodeep dans la Figure 3 montrent clairement son efficacité dans la diminution de l'effet de

4 Conclusion & Perspectives

Une nouvelle méthode de classification de STIS multivariées a été proposée pour cartographier l'occupation des sols à partir des SITS multivariées. Cette méthode utilise une combinaison des techniques d'encodage pour convertir les données en images 2D, qui sont ensuite analysées à l'aide d'un réseau de neurones convolutifs. Nos

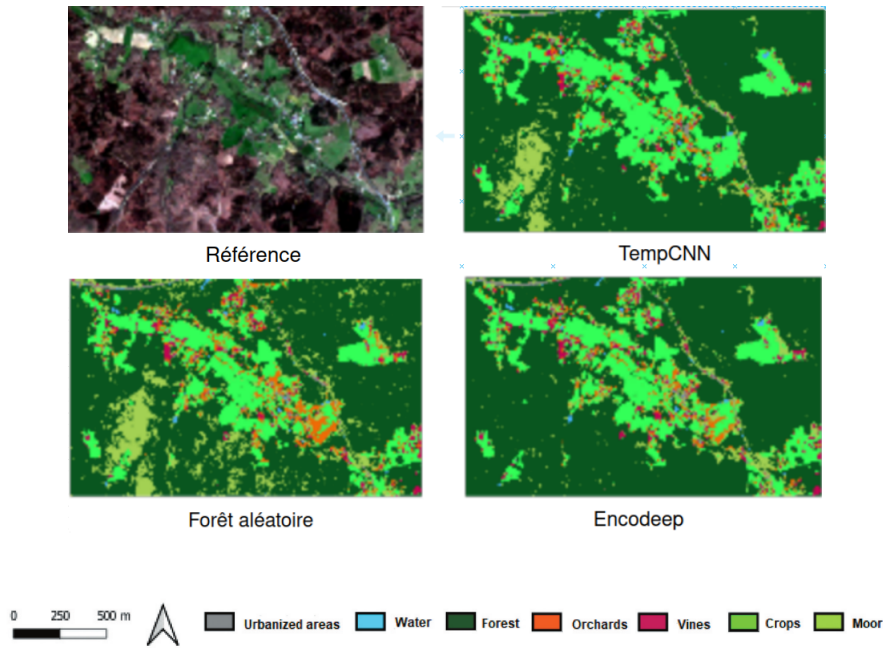


Figure 3. Extraits de carte obtenus à partir de l’application des différentes méthodes concurrentes sur le site d’étude de la Dordogne.

résultats expérimentaux ont montré que cette approche est efficace sur les données STIS multivariées complexes, en modélisant l’interaction inter-bandes des données multispectrales. Les travaux futurs porteront sur l’évaluation de la méthode avec des réseaux neuronaux plus récents et sur l’extension de la méthode à d’autres sites d’étude.

References

1. A. Abidi, A. Ben Abbes, Y. J. E. Gbodjo, D. Ienco, and I. R. Farah. Combining pixel- and object-level information for land-cover mapping using time-series of Sentinel-2 satellite data. *Remote Sensing Letters*, 2022.
2. Jorge Belaire-Franch and Dulce Contreras. Recurrence plots in nonlinear time series analysis: Free software. *Journal of Statistical Software*, 2002.
3. Aybuke Civrizoglu Buz. A Novel Approach and Application of Time Series to Image Transformation Methods on Classification of Underwater Objects. *Gazi Journal of Engineering Sciences*, 2021.
4. Xing Cai, Huimin Zhao, Shifan Shang, Yongquan Zhou, Wu Deng, Huayue Chen, and Wuquan Deng. An improved quantum-inspired cooperative co-evolution algorithm with multi-strategy and its application. *Expert Systems with Applications*, 2021.
5. Wu Deng, Junjie Xu, Xiao-Zhi Gao, and Huimin Zhao. An enhanced msqde algorithm with novel multiple strategies for global optimization problems. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2022.

6. Danielle Dias, Allan Pinto, Ulisses Dias, Rubens Lamparelli, Gueric Le, and Ricardo S Torres. Image-Based Time Series Representations for Pixelwise Eucalyptus Region Classification : A Comparative Study. *IEEE Geoscience and Remote Sensing Letters*, 2020.
7. Danielle Dias, Allan Pinto, Ulisses Dias, Rubens Lamparelli, Gueric Le Maire, and Ricardo S Torres. A multi-representational fusion of time series for pixelwise classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2020.
8. Yawogan Jean Eudes Gbodjo, Dino Ienco, Louise Leroux, Roberto Interdonato, Raffaele Gaetano, and Babacar Ndao. Object-based multi-temporal and multi-source land cover mapping leveraging hierarchical class relationships. *Remote. Sens.*, 2020.
9. Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016.
10. J. Inglada, A. Vincent, M. Arias, B. Tardy, D. Morin, and I. Rodes. Operational High Resolution Land Cover Map Production at the Country Scale Using Satellite Image Time Series. *Remote Sensing*, 2017.
11. C. Pelletier, G. I. Webb, and F. Petitjean. Temporal convolutional neural network for the classification of satellite image time series. *Remote Sensing*, 2019.
12. Zhiguang Wang and Tim Oates. Imaging time-series to improve classification and imputation. *IJCAI International Joint Conference on Artificial Intelligence*, 2015.
13. Chao lung Yang, Chen yi Yang, Zhi xuan Chen, Nai wei Lo, and Senior Member. Multivariate Time Series Data Transformation for Convolutional Neural Network. *Proceedings of the 2019 IEEE/SICE International Symposium on System Integration, SII 2019*, 2019.