



**HAL**  
open science

## Filtrage des Données Radar par Deep Learning

Pierre Lepetit, Laurent Barthès, Cécile Mallet, Camille Ly, Yvon Lemaître, Nicolas Viltard

► **To cite this version:**

Pierre Lepetit, Laurent Barthès, Cécile Mallet, Camille Ly, Yvon Lemaître, et al.. Filtrage des Données Radar par Deep Learning. Conférence ENVIREM 2019, Jul 2019, Palaiseau, France. <insu-02360403>

**HAL Id: insu-02360403**

**<https://insu.hal.science/insu-02360403v1>**

Submitted on 12 Nov 2019

**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.



HAL Authorization

# Filtrage des Données Radar par Deep Learning

Pierre Lepetit \* ‡, Laurent Barthès ‡, Cécile Mallet ‡, Camille Ly ‡, Yvon Lemaître‡, Nicolas Viltard‡

\*Météo-France

† LATMOS, Guyancourt

Emails:

prénom.nom@latmos.ipsl.fr

***Index Terms***—Radar météorologique, restauration d’images, Full Convolutional Network, supervision faible

## INTRODUCTION

Les données brutes issues des radars météorologiques fixes peuvent comporter des signaux parasites. Ainsi le relief et les constructions élevées, la turbulence atmosphérique, les oiseaux et les insectes génèrent des motifs qui compliquent l’interprétation de l’image et faussent les cumuls de précipitations. Ces motifs sont caractérisés par une signature polarimétrique, par une forme, ou par une dynamique particulière. De nombreuses techniques ont été développées pour les détecter dans l’image et pour les en soustraire. Nous utiliserons les termes de segmentation et de restauration pour désigner ces deux tâches.

Le problème de la segmentation des signaux parasites a été abordé par seuillage simple, arbres de décision [1], réseaux de neurones [2], [3] et logique floue [4], [5], [6], [7]. La catégorie du radar (généralement bande S ou bande C) et les signaux parasites ciblés varient, mais la plupart de ces méthodes ont en commun le caractère local de la prédiction: la classe est décidée à partir des mesures associées au pixel, ou, éventuellement aux pixels adjacents lorsque des indices de textures [5] ou des gradients [6] sont utilisées.

Le problème de la restauration est généralement abordé en deux étapes : la segmentation des signaux parasites permet de définir un masque binaire. L’étape de restauration à proprement parler consiste alors à compléter les régions masquées (non blind inpainting). Elle peut impliquer plusieurs angles d’élévation et des procédés locaux d’interpolation, mais à notre connaissance, ni les corrélations longue portée ni les valeurs des variables polarimétriques associées aux pixels masqués ne sont exploitées. Elles pourraient pourtant être utiles lorsque les signaux parasites et les échos de précipitation sont mêlés. La correction de ces “échos mêlés” est d’ailleurs encore un challenge [5]. Dans notre étude, ce problème fait l’objet d’une attention particulière.

Récemment, des tâches de segmentation et de restauration ont été abordées à l’aide de réseaux à couches de convolution particuliers, les FCN (“Fully Convolutional Network”) [8], [9]. Avec ces modèles, la prédiction d’une classe ou d’une intensité dépend d’un large voisinage, à travers des règles de décision sélectionnées par apprentissage. Si les performances de ces réseaux ont d’abord été établies dans le cadre d’un

apprentissage supervisé, elles ont dernièrement été étendues à des contextes moins contraignant en matière de labellisation.

Dans cet article, nous présentons des méthodes faiblement supervisées de segmentation binaire et de restauration de la réflectivité horizontale à partir de FCN. C’est la première fois que les FCN sont utilisés dans ce contexte. Pour éviter d’avoir à nettoyer “à la main ” des images de réflectivité, nous nous sommes placés dans le cadre étudié dans [10], où deux bases sont disponibles. La première contient les images à nettoyer, la deuxième ne contient que des signaux parasites. Dans notre cas, les signaux parasites sont collectés en dehors des périodes de précipitations. La méthode consiste à les superposer aux images à nettoyer. Nous entraînons les réseaux à détecter ou à soustraire les signaux parasites additionnels en espérant qu’ils traitent de la même façon les signaux parasites déjà présents dans l’image. Ces signaux parasites “initiaux” sont assimilés à un problème de cible bruitée contre lequel nous pouvons employer des stratégies déjà éprouvées. Les méthodes sont d’abord validées sur des données de synthèse, pour lesquelles ce bruit de labellisation est maîtrisable. Puis elles sont évaluées sur les données réelles à travers des études de cas et une analyse des cumuls.

## *Nettoyage des données des radars météorologiques polarimétriques*

Au début des années 2000, la segmentation des échos radar est déjà dominée par les méthodes statistiques, avec d’un côté des réseaux de neurones peu profonds [2], [3] et de l’autre des approches par logique floue [6], [5]. Dans les deux cas, le traitement est très local, ou basé sur des indices de textures intuitifs, et repose sur une collection de pixels classés à la main par des experts (approche supervisée).

Plus récemment, la classification a été subordonnée à une analyse plus globale du contexte par Deep Learning: Wang et al. [11] commencent par entraîner un réseau de neurones à couches de convolution (CNN) sur une tâche de segmentation qui n’implique que quelques grandes classes d’échos. Ils raffinent ensuite la prédiction par logique floue. Les auteurs remarquent une amélioration de la détection des signaux parasites, et l’attribuent à la prise en compte d’un large voisinage. L’estimation des précipitations par l’observation satellitaire fait face aux mêmes problèmes de classification. Les réseaux de neurone profonds ont donné des résultats intéressants sur le problème basique de la classification pluie/non pluie dans un

cadre supervisé [12], [13]. Cependant, dans ces travaux, les architectures utilisées ne correspondent pas l'état de l'art sur des problèmes de segmentation ou de restauration, même dans un cadre supervisé.

Quelques études sortent du cadre de l'apprentissage supervisé : Bechini et al. [14], Besic et al. [15] ont adapté et développé une méthode de clustering non-supervisée pour la classification des hydrométéores. Ils ajoutent une contrainte chargée d'assurer la cohérence spatiale des types d'hydrométéores prédits. Par contre, la question de la restauration n'est pas abordée.

#### *Segmentation et restauration avec les FCN*

Depuis 2015, les FCN (Full Convolutional neural Networks) offrent une alternative performante aux CNNs simples sur de nombreuses tâches. Ces FCN ne possèdent pas de couches complètement connectées. Grâce à des couches de déconvolution, la taille de la sortie et celle de l'entrée peuvent être rendues identiques. Ces réseaux sont entre autres employés pour des tâches de segmentation [8], [9], [16] et de restauration aveugle <sup>1</sup> [9], [17], [18]. Dans [19], la restauration est abordée en parallèle de la segmentation des motifs à gommer, suivant une approche multi-tâche. En restauration d'image, les FCN font l'état de l'art sur des problèmes d'école ("text removal", "watermark removal", etc) et commencent à être utilisés en télédétection. Par exemple, Xiao et al. [20] nettoient par FCN les motifs parasites (bandes parallèles) sur des images satellites IR.

Dans les études citées précédemment, l'entraînement est parfaitement supervisé. Pour la restauration, on dispose dans ce cadre de deux bases, l'une constituée de motifs ( $z_i$ ), par exemples des textes de couleurs et police variés, "watermark", etc ; l'autre constituée d'images propres ( $y_i$ ). Le réseau est alors entraîné sur des paires entrée-cible ( $x_i = y_i \oplus z_i, y_i$ ), où l'opérateur  $\oplus$  désigne une superposition (matting). Nous conservons ces notations génériques par la suite.

#### *Restauration faiblement supervisée*

Pour la restauration, des méthodes par Deep Learning ont été récemment développées pour le cas où les images propres ne sont pas disponibles. Pajot et De Bezenac [21] supposent seulement que le processus de corruption  $F_\theta : y_i \rightarrow x_i$ , où  $\theta$  est une variable aléatoire latente, est simulable. Halperin et al. [10] travaillent dans le cas particulier où le procédé de corruption est une superposition : seuls sont disponibles des images parasitées  $y_i \oplus z_i$  et des motifs parasites  $z'_i$  de même nature que les  $z_i$ . Les  $y_i$  sont estimés itérativement selon un procédé sensible à l'initialisation (Neural Egg Separation) : à chaque pas, un FCN est entraîné à isoler les  $z'_i$  des  $\tilde{y}_i^k \oplus z'_i$  où les  $\tilde{y}_i^k$  sont les estimées successives des  $y_i$ .

Comme dans [10], nous supposons que le processus de corruption correspond à une superposition de signaux parasites  $z_i$  indépendants des  $y_i$  et que ces signaux peuvent être échantillonnés en dehors des épisodes de pluie. Notre

<sup>1</sup>"blind inpainting". Sur ces tâches, il s'agit de restaurer l'image sans savoir quels pixels ont été dégradés

algorithme ressemble à la première étape du NES, dans sa forme la plus naïve. Nous entraînons un FCN à restaurer des images corrompues sur les paires entrée-cible de la forme  $([y_i \oplus z_i] \oplus z'_i, [y_i \oplus z_i])$ . Mais l'angle d'attaque est différent : dans ces paires, la cible  $y'_i = y_i \oplus z_i$  est corrompue par les signaux parasites initiaux. Nous avons donc développé une méthode adaptée au cas de cibles bruitées. Nous abordons en parallèle le problème de la segmentation binaire. De nouveau, les images à nettoyer sont enrichies en signaux parasites additionnels. Le support de ces signaux parasites additionnels définit les deux classes : un pixel est "corrompu" ou "non corrompu". Les régions de l'image contenues dans le support des signaux parasites initiaux sont donc associées au mauvais label : il s'agit encore d'un problème de bruit de labellisation.

#### *Cibles bruitées. Quel impact? Comment améliorer la robustesse?*

Un problème de restauration avec bruit de cible a récemment été étudié par Lehtinen et al. [18]. A partir d'une formalisation du problème d'apprentissage, ils adaptent la fonction de coût à la forme du bruit et vérifient sur des FCN que les performances sont identiques au cas supervisé, voire meilleures. Mais dans cette étude, les bruits sur la cibles et ceux sur les entrées sont indépendants et leur conclusion ne peut pas être étendue au cas où les entrées contiennent le même bruit que les cibles.

L'analyse et la correction des effets du bruit de labellisation sur les algorithmes de classification ont été plus étudiés [22]. Sur des tâches de classification et de segmentation où les labels sont intervertis aléatoirement ("flip noise"), les CNN se sont montrés relativement robustes [23], [24], [25]. Rolnick et al. [24] montrent par exemple que pourvu que la taille des bases soit assez grande, un "flip noise" important a des effets limités sur la justesse malgré une proportion de labels erronés très défavorable.

Petit et al. [26] étudient les performances d'un FCN sur une tâche de segmentation binaire, dans un cadre semi-supervisé. Ils comparent deux approches. Dans la première, ils utilisent toutes les images disponibles et des régions entières peuvent être mal labellisées, comme le sont nos signaux parasites initiaux. Les auteurs montrent que l'effet du bruit de labellisation sur les performances est significatif. L'autre approche, itérative, leur permet de recoller aux scores obtenus sans bruit de labellisation. Mais elle est initialisée sur une sous-base d'images proprement labellisées.

Des méthodes moins exigeantes existent. Les unes intègrent une étape de détection des erreurs de labellisation [27], [28], [29], etc, les autres incorporent un modèle de bruit [23], [30], etc.

Dans [27], sur une tâche de classification, un premier apprentissage portant sur toutes les images permet de repérer les données suspectes, et de les rejeter. Un deuxième apprentissage est ensuite conduit, sur la base nettoyée. Avec cette approche, le risque est de rejeter les cas difficiles, importants pour l'apprentissage, avec les cas de labellisation erronée. Cependant, comme elle porte ses fruits sur des données radar

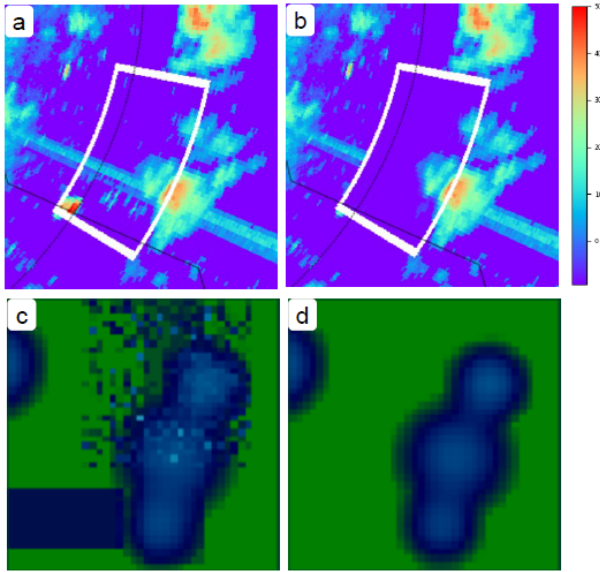


Fig. 1. a. : Réflectivité horizontale en dBZ à  $0.4^\circ$  d'élévation. Secteur de Sainte-Assise (77). 9/11/2017, 11h40 UTC. La partie de l'image dans le cadre blanc est restaurée en b. où l'écho fixe associé au radio-émetteur de Sainte-Assise (en bas à gauche du cadre) et l'interférence principale ont été correctement gommés. c. : exemple d'image radar de synthèse. Les échos de précipitation sont représentés par des disques, les signaux parasites par des rectangles plein ou bruités. d. : restauration de c.

de synthèse, nous l'avons adaptée aux cas de la segmentation binaire et de la restauration.

## LES DONNÉES

### Les données brutes

La méthode est mise en place sur les données brutes du radar météorologique de Trappes (78, France) associées à l'angle d'élévation le plus bas ( $0.4^\circ$ ). Ces données présentent des échos de sol permanents mais d'intensité variable, des échos de propagation anormales, des échos de ciel clair, etc. La figure 1.a. présente quelques exemples de ces signaux parasites : échos de sol et interférence.

L'entraînement des réseaux porte sur environ dix mois d'archives au pas de temps 5 minutes. En entrée, nous utilisons les quatre canaux polarimétriques disponibles : réflectivité horizontale ( $Z_h$ ), autocorrelation en 0 ( $\rho_0$ ), gradient radial de la phase différentielle ( $K_{dp}$ ) et réflectivité différentielle ( $Z_{dr}$ ). Nous avons exclu de l'analyse le rond central (0 – 15 kms), trop bruité, et les parties éloignées de l'image radar (à partir de 150 kms).

### Préparation des bases

Nous avons préparé deux bases, l'une,  $B_p$ , ne comportant que des signaux parasites (Base "parasites"), l'autre  $B_r$ , comportant d'éventuels signaux parasites mêlés à des échos de précipitations (Base "pluie"). Pour extraire ces images, nous avons utilisé les données pluviomètres de la région parisienne. Nous isolons des épisodes de pluie des intervalles de non-pluie sur un critère développé dans [31]. Les images radar de la base

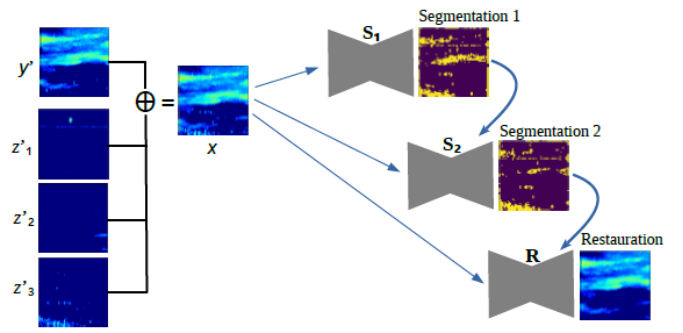


Fig. 2. Les  $z'_j$  issus de  $B_p$  sont superposés à  $y'$ , prise dans  $B_r$ , pour former l'entrée  $x$ . D'abord,  $S_1$  est entraînée à segmenter les signaux parasites contenus dans les  $z'_j$ . Dans un second temps, les pixels signalés par  $S_1$  comme mal labellisés sont masqués pour apprendre  $S_2$ . Pour la restauration, c'est  $S_2$  qui sert de masque pour entraîner  $R$ .

"parasites" (resp. base "pluies") sont extraites de la couronne située entre 15 kms et 150 kms pendant les intervalles de non-pluie (resp. pendant les épisodes de pluie). Sans autre traitement, les images extraites ne présentent souvent aucun signal significatif. Pour améliorer les bases, ces images ont été regroupées par clustering (K-means) et exclues des bases. Après cette étape, les deux bases contiennent environ  $10^5$  images de  $128 \times 128$  pixels.

### Données de synthèse

Pour évaluer et régler l'algorithme, nous avons utilisé des images de simulation (voir figure 1.c). Les pseudo-échos de précipitations, discoïdes, et les signaux parasites, rectangulaires, sont générés aléatoirement et les centres des formes sont répartis suivant une loi uniforme. Les signaux parasites sont de deux formes différentes (pleine et bruitée) de dimensions et d'intensité variable. Les données de synthèse permettent de contrôler le bruit de cible. Les briques successives de notre algorithme ont été validées sur ces données.

## MÉTHODE

Une fois les bases mises en place, une opération de superposition est définie de façon à ce que les signaux parasites additionnels  $z'_i$  ne puissent pas être distingués des signaux parasites initiaux  $z_i$ . Lorsque les signaux parasites additionnels sont majoritaires, l'entraînement d'un premier FCN sur une tâche de segmentation devient envisageable. Suivent deux étapes successives d'entraînement pour affiner la segmentation et apprendre la restauration. La figure 2 présente un aperçu de la méthode. L'ajout de signaux parasites modifie la distribution des entrées. Pour contenir les effets négatif de cette modifie, la sortie est post-traitée suivant une méthode d'ensemble.

### Opération de superposition et définition des entrées

Les images de  $B_r$  sont notées  $y'_i$ , celles de  $B_p$ ,  $z'_i$ . Elles contiennent les quatre canaux polarimétriques. Pour deux cibles  $R$  et  $P$  situées à la même distance du radar, les puissances du signal rétropropagées s'additionnent :

$$E^{S_1} = E \setminus F_{noisy\ label}^{S_1} \quad (3)$$

$$Z_h = Z_h(P) + Z_h(R)$$

$$Z_v = Z_v(P) + Z_v(R)$$

$$\begin{aligned} Z_{dr} &= Z_v/Z_h \\ &= 1/Z_h[Z_h(P) \times Z_{dr}(P) + Z_v(R) \times Z_{dr}(R)] \\ &= m_h(Z_{dr}(P), (Z_{dr}(R))) \end{aligned}$$

Où  $m_h$  désigne la moyenne pondérée par les réflectivités horizontales. Pour  $K_{dp}$  et l'autocorrélation en zéro, la composition des signaux fait intervenir des données indisponibles. Par défaut, nous interpolons avec  $m_h$  entre les situations où l'un des signal est de puissance nulle :  $K_{dp} = m_h(K_{dp}(R), K_{dp}(P))$  et  $\rho_0 = m_h(\rho_0(R), \rho_0(P))$ . Nous supposons ainsi que les images de  $B_r$  sont décomposables en:  $y'_i = y_i \oplus z_i$  où  $y_i$  contient les échos de précipitation,  $z_i$  les signaux parasites et  $\oplus$  correspond à une somme sur la première composante et à  $m_h$  sur les trois autres.

Nous supposons encore que les signaux parasites initiaux contenus dans les  $z_i$ , peuvent être considérés comme indépendants des échos de pluie  $y_i$ , et distribués comme les  $z'_i$ , si bien qu'en formant les entrées  $x_i = y'_i \oplus z'_i$ , les signaux parasites contenus dans les  $z'_i$  et ceux contenus dans les  $z_i$  sont indiscernables. Ajouter plusieurs images de  $B_p$  permet de s'assurer que les signaux parasites additionnels sont plus fréquents que ceux déjà contenus dans l'image. Mais des signaux parasites trop nombreux ne seraient plus identifiables. D'après expérience sur les données de synthèse, l'ajout de trois images  $z'_{i,1}, z'_{i,2}, z'_{i,3}$  prises dans  $B_p$  est un bon compromis. Ainsi, les entrées sont définies par :

$$x_i = y'_i \oplus \left[ \bigoplus_{j=1}^3 z'_{i,j} \right] \quad (1)$$

### Segmentation et restauration

La méthode comprend trois phases qui consistent à entraîner des architectures U-Net [32] sur des tâches de segmentation et de restauration. L'initialisation des poids [33], la méthode d'optimisation [34], le taux d'apprentissage ( $10^{-4}$ ) sont standards. Le premier réseau est entraîné à la segmentation et la fonction associée est notée  $S_1$ . La classification est binaire : les pixels cibles sont classés "corrompus" partout où  $z'_i > 0$ . La fonction de coût est définie par l'entropie croisée :

$$L_1 = - \sum_{k \in E} \ln(P_{S_1}^k(C^k)) \quad (2)$$

où  $E$  désigne l'ensemble des pixels de la cible,  $C^k$ , la classe associée au pixel  $k$  et  $P_{S_1}^k(C^k)$ , la probabilité associée à  $C^k$  en sortie de  $S_1$ .

Ensuite, un deuxième U-Net  $S_2$  est initialisé avec les poids de  $S_1$  et entraîné avec une fonction de coût modifiée. Suivant [27], le poids des pixels soupçonnés d'être mal labellisés (ceux contenus dans  $F_{noisy\ label}^{S_1}$  — voir équation (3)) est réduit. Sur les données de synthèse, c'est un poids nul qui offre les meilleurs résultats. Cela conduit, pour définir  $L_2$ , à remplacer  $E$  par :

Où:

$$\begin{aligned} F_{noisy\ label}^{S_1} &= \{k \in E \mid P_{S_1}^k(\text{"corrompu"}) > 0.5 \\ &\quad \text{et } C^k = \text{"non corrompu"}\} \end{aligned}$$

Enfin, un troisième U-Net, noté  $R$ , est entraîné sur une tâche de restauration. Les entrées sont inchangées, mais les cibles sont définies par les premières composantes des  $y'$  (i.e., la réflectivité horizontale), et la dernière couche du réseau, modifiée en conséquence, génère un champ scalaire, noté  $\tilde{y}$ . Pour la fonction de coût, l'erreur absolue moyenne est pondérée en utilisant cette fois les sorties de  $S_2$  :

$$L_3 = \sum_{k \in E^{S_2}} |\tilde{y}^k - y'^k| \quad (4)$$

En phase de test, pour éviter de bruite les sorties sur les régions qui n'ont pas été corrompues par des signaux parasites, les sorties  $\tilde{y}$  sont conservées selon les prédictions de  $S_2$  :

$$\tilde{y}^k := \begin{cases} \tilde{y}^k & \text{si } P_k^{S_2}(\text{"corrompu"}) > 0.5 \\ y'^k & \text{sinon} \end{cases}$$

### Post-traitement

L'ajout d'échos parasites présente un défaut commun aux méthodes auto-supervisées : la distribution des entrées est modifiée (cf. [35]). Plutôt que d'appliquer directement l'algorithme sur les images de la base "pluies", nous répétons  $N$  fois l'ajout suivant une méthode d'ensemble naïve. A partir d'une entrée  $x_{test}$ , nous formons les entrées:

$x_{test,l} = x_{test} \oplus \bigoplus_{j=1}^3 z'_{l,j}$ ,  $l = 1, \dots, N$  (en pratique, on prend  $N = 16$ ). Pour calculer la sortie post-traitée  $\tilde{y}_{p.t.}$  au pixel  $k$ , on ne tient compte que des entrées  $x_{test,l}$  qui n'ont pas été corrompues en  $k$  :

$$\tilde{y}_{p.t.}^k = \frac{1}{|F_k|} \left[ \sum_{l \in F_k} \tilde{y}_l^k \right] \quad (5)$$

où  $F_k = \{l \mid \forall j, z'_{l,j} = 0\}$  et  $\tilde{y}_l^k$  correspond à  $S_2(x_{test,l})^k$  ou de  $R(x_{test,l})^k$  selon la tâche.

## RÉSULTATS

### Etude du secteur de Sainte-Assise

Les réflectivités moyennes révèlent la présence d'échos parasites fixes. Parmi eux, les interférences avec des signaux parasites émis par une source fixe et les échos de sols dont l'intensité peut varier selon l'indice de réfractivité de l'atmosphère et l'humidité des sols.

Le secteur de Sainte-Assise (77), délimité par le contour blanc de la figure 1.a, présente plusieurs types d'échos fixes. Des pylônes de plus de deux cents mètres de haut y génèrent un écho fixe massif important, visible en bas à gauche de la figure 3.a. Dans ce secteur sont aussi visibles deux "bandes d'interférence", la première massive et très fréquente (cf

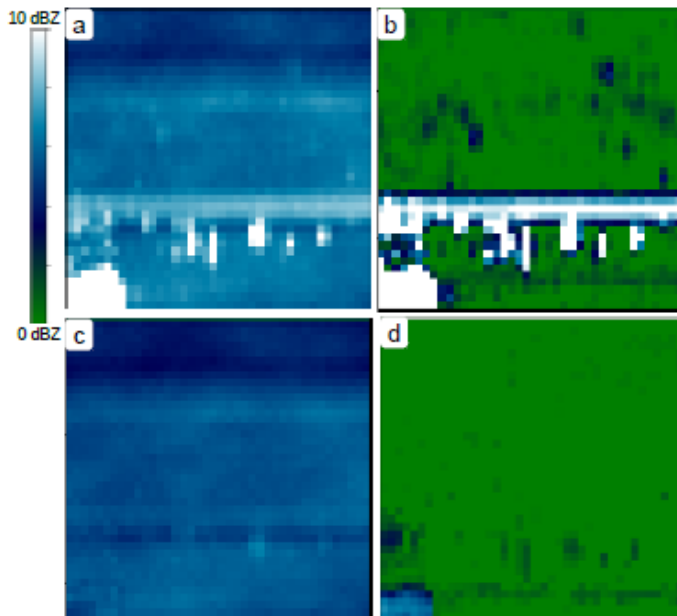


Fig. 3. Etude du secteur de Sainte-Assise. a. : moyenne des réflectivité en entrée. c. : moyenne des réflectivité après restauration. b. : différence entre a. et b. (échelle non respectée). d. : cas particulier où l'interférence secondaire est visible

figure 1.a). La seconde est plus ténue et occasionnelle ; des échos de sol intermittents peuvent aussi se manifester (cf. figure 3.d).

La figure 3.c montre la réflectivité moyenne après restauration. Le traitement conserve les effets de l'atténuation (bandes plus sombres) mais gomme correctement les traces des signaux parasites fixes, fréquents ou non.

Sur ce secteur de Sainte-Assise, l'efficacité de la méthode a été évaluée sur une base d'un millier d'images. L'écho de sol principal et les interférences sont presque systématiquement détectés (respectivement 99 % sur 896 cas et 97 % sur 190 cas). Les échos de sols isolés des précipitations sont correctement gommés, comme sur la figure 1.b, et ceux mêlés aux échos de pluie le sont dans 71 % des cas. Les interférences ne sont correctement gommées que dans 74 % des cas (traces résiduelles).

#### Comparaison avec le typage de Météo-France

Notre algorithme a été comparé avec celui de Météo-France dans une situation de convection en air chaud (figure 4.a). Sur l'image de réflectivité, des échos d'une intensité de quelques dBZ forment une auréole d'une cinquantaine de kilomètres autour du radar. D'après l'analyse des données Doppler et des séquences des radars ROXI et BASTA situés au SIRTA, il s'agit d'échos de ciel clair.

Au delà d'un disque de 10 kms de rayon, ces signaux sont en partie confondus avec des échos de précipitation par l'algorithme de typage de Météo-France basé sur une approche par logique floue (figure 4.b). Dans la même zone, notre algorithme les soustrait entièrement (figure 4.c.).

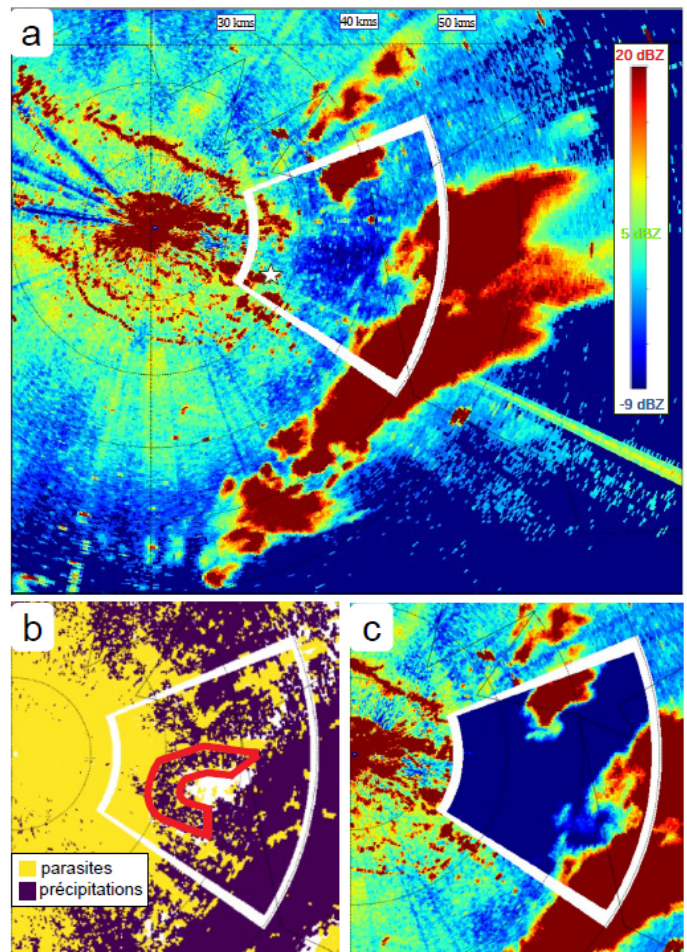


Fig. 4. Comparaison avec le typage Météo-France. Cas de précipitations convectives le 15/08/2017 à 20 h 20. a.:Réflectivité horizontale à 0.4°. L'étoile blanche correspond au site du SIRTA. b. Classification pluie/parasites issue de l'algorithme de typage de Météo-France. Les zones blanches correspondent aux pixels à -9dBZ. Dans le contour en forme de C, le typage est erroné. c.: restauration par notre algorithme

#### CONCLUSION

Nous avons proposé un algorithme de segmentation et de restauration des images radar. Cet algorithme est basé sur les FCN (Full Convulsive neural Networks), des architectures qui font l'état de l'art sur des problématiques de segmentation sémantique et de restauration d'image, dans un cadre supervisé comme en situation de supervision faible. Dans cette étude, nous avons évité le recours à l'expertise humaine, en échantillonnant les signaux parasites entre les épisodes de pluie. Ces signaux sont superposés aux images à nettoyer, et les FCN sont entraînés à les détecter et les éliminer. Les signaux parasites initialement présents dans l'image sont alors assimilés à un bruit de cible. Une méthode d'apprentissage simple et adaptée favorise un nettoyage complet des images, signaux parasites superposés et initiaux confondus.

Malgré des hypothèses relativement fortes sur la distribution et la composition de signaux parasites avec les échos de précipitation, les résultats sont prometteurs. L'algorithme a

été évalué sur le site du radio-émetteur de Sainte-Assise par des cumuls et une étude de cas, puis comparé au typage opérationnel de Météo-France. En dehors de son caractère objectif, cette méthode relativement générique pourra être testée sur les données issues de radars embarqués. Des alternatives récentes (cf. [10], [21]) pourront aussi lui être comparées.

#### REMERCIEMENTS

Les données brutes et les typages nous ont été transmis par le Centre de Météorologie Radar de Météo-France. Nous remercions en particulier Clotilde Augros, Valérie Vogt et Nicolas Gaussiat pour la transmission des données et les retours constructifs.

#### REFERENCES

- [1] A. V. Ryzhkov and D. S. Zrnica, "Polarimetric rainfall estimation in the presence of anomalous propagation," *Journal of Atmospheric and Oceanic Technology*, vol. 15, no. 6, pp. 1320–1330, 1998.
- [2] R. Cornelius, R. Gagon, and F. Pratte, "Optimization of wsr-88d clutter processing and ap clutter mitigation," *Final Report*, 1995.
- [3] V. Lakshmanan, K. Hondl, G. Stumpf, and T. Smith, "Quality control of weather radar data using texture features and a neural network," in *Preprints, 31st Radar Conference*. Citeseer, 2003, pp. 522–525.
- [4] M. Steiner and J. A. Smith, "Use of three-dimensional reflectivity structure for automated detection and removal of nonprecipitating echoes in radar data," *Journal of Atmospheric and Oceanic Technology*, vol. 19, no. 5, pp. 673–686, 2002.
- [5] J. J. Gourley, P. Tabary, and J. Parent du Chatelet, "A fuzzy logic algorithm for the separation of precipitating from nonprecipitating echoes using polarimetric radar observations," *Journal of Atmospheric and Oceanic Technology*, vol. 24, no. 8, pp. 1439–1451, 2007.
- [6] Y.-H. Cho, G. W. Lee, K.-E. Kim, and I. Zawadzki, "Identification and removal of ground echoes and anomalous propagation using the characteristics of radar echoes," *Journal of Atmospheric and Oceanic Technology*, vol. 23, no. 9, pp. 1206–1222, 2006.
- [7] V. Chandrasekar, R. Keränen, S. Lim, and D. Moisseev, "Recent advances in classification of observations from dual polarization weather radars," *Atmospheric Research*, vol. 119, pp. 97–111, 2013.
- [8] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.
- [9] X.-J. Mao, C. Shen, and Y.-B. Yang, "Image restoration using convolutional auto-encoders with symmetric skip connections," *arXiv preprint arXiv:1606.08921*, 2016.
- [10] T. Halperin, A. Ephrat, and Y. Hoshen, "Neural separation of observed and unobserved distributions," *arXiv preprint arXiv:1811.12739*, 2018.
- [11] H. Wang, Y. Ran, Y. Deng, and X. Wang, "Study on deep-learning-based identification of hydrometeors observed by dual polarization doppler weather radars," *EURASIP Journal on Wireless Communications and Networking*, vol. 2017, no. 1, p. 173, 2017.
- [12] Y. Tao, X. Gao, K. Hsu, S. Sorooshian, and A. Ihler, "A deep neural network modeling framework to reduce bias in satellite precipitation products," *Journal of Hydrometeorology*, vol. 17, no. 3, pp. 931–945, 2016.
- [13] Y. Tao, X. Gao, A. Ihler, S. Sorooshian, and K. Hsu, "Precipitation identification with bispectral satellite information using deep learning approaches," *Journal of Hydrometeorology*, vol. 18, no. 5, pp. 1271–1283, 2017.
- [14] R. Bechini and V. Chandrasekar, "A semisupervised robust hydrometeor classification method for dual-polarization radar applications," *Journal of Atmospheric and Oceanic Technology*, vol. 32, no. 1, pp. 22–47, 2015.
- [15] N. Besic, J. Grazioli, M. Gabella, U. Germann, A. Berne *et al.*, "Hydrometeor classification through statistical clustering of polarimetric radar measurements: a semi-supervised approach," *Atmospheric Measurement Techniques*, vol. 9, no. 9, 2016.
- [16] P. Zhang, W. Liu, H. Wang, Y. Lei, and H. Lu, "Deep gated attention networks for large-scale street-level scene segmentation," *Pattern Recognition*, vol. 88, pp. 702–714, 2019.
- [17] Y. P. Liu, J. shan Pan, and Z. Su, "Deep blind image inpainting," *CoRR*, vol. abs/1712.09078, 2017.
- [18] J. Lehtinen, J. Munkberg, J. Hasselgren, S. Laine, T. Karras, M. Aittala, and T. Aila, "Noise2noise: Learning image restoration without clean data," *arXiv preprint arXiv:1803.04189*, 2018.
- [19] A. Hertz, S. Fogel, R. Hanocka, R. Giryes, and D. Cohen-Or, "Blind visual motif removal from a single image," *arXiv preprint arXiv:1904.02756*, 2019.
- [20] P. Xiao, Y. Guo, and P. Zhuang, "Removing stripe noise from infrared cloud images via deep convolutional networks," *IEEE Photonics Journal*, vol. 10, no. 4, pp. 1–14, 2018.
- [21] A. Pajot, E. de Bezenac, and P. Gallinari, "Unsupervised adversarial image reconstruction," 2018.
- [22] B. Frény and M. Verleysen, "Classification in the presence of label noise: a survey," *IEEE transactions on neural networks and learning systems*, vol. 25, no. 5, pp. 845–869, 2013.
- [23] S. Sukhbaatar, J. Bruna, M. Paluri, L. Bourdev, and R. Fergus, "Training convolutional networks with noisy labels," *arXiv preprint arXiv:1406.2080*, 2014.
- [24] D. Rolnick, A. Veit, S. Belongie, and N. Shavit, "Deep learning is robust to massive label noise," *arXiv preprint arXiv:1705.10694*, 2017.
- [25] N. Heller, J. Dean, and N. Papanikolopoulos, "Imperfect segmentation labels: How much do they matter?" in *Intravascular Imaging and Computer Assisted Stenting and Large-Scale Annotation of Biomedical Data and Expert Label Synthesis*. Springer, 2018, pp. 112–120.
- [26] O. Petit, N. Thome, A. Charnoz, A. Hostettler, and L. Soler, "Handling missing annotations for semantic segmentation with deep convnets," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Springer, 2018, pp. 20–28.
- [27] C. E. Brodley and M. A. Friedl, "Identifying mislabeled training data," *Journal of artificial intelligence research*, vol. 11, pp. 131–167, 1999.
- [28] S. Reed, H. Lee, D. Anguelov, C. Szegedy, D. Erhan, and A. Rabinovich, "Training deep neural networks on noisy labels with bootstrapping," *arXiv preprint arXiv:1412.6596*, 2014.
- [29] Y. Wang, W. Liu, X. Ma, J. Bailey, H. Zha, L. Song, and S.-T. Xia, "Iterative learning with open-set noisy labels," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 8688–8696.
- [30] J. Li, Y. Wong, Q. Zhao, and M. Kankanhalli, "Learning to learn from noisy labeled data," 12 2018.
- [31] M. D. Dilmi, "Méthode de classification des séries temporelles: application à un réseau de pluviomètres," Ph.D. dissertation, 2019.
- [32] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.
- [33] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, 2010, pp. 249–256.
- [34] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [35] J. Donahue, P. Krähenbühl, and T. Darrell, "Adversarial feature learning," *arXiv preprint arXiv:1605.09782*, 2016.