

# Apprentissage profond appliqué à la reconstruction interférométrique de transitoires radio

Benjamin Naoto CHICHE<sup>1,2</sup>, Julien N. GIRARD<sup>3</sup>, Joana FRONTERA-PONS<sup>1,4</sup>, Arnaud WOISELLE<sup>2</sup>, Jean-Luc STARCK<sup>1</sup>

<sup>1</sup>AIM, CEA, CNRS, Université Paris-Saclay, Université Paris Cité, F-91191 Gif-sur-Yvette, France

<sup>2</sup>Safran Electronics & Defense, F-91344 Massy, France

<sup>3</sup>LESIA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Université de Paris Cité, 92195 Meudon, France

benjamin.chiche@safrangroup.com, julien.girard@obspm.fr  
joana.frontera-pons@cea.fr, arnaud.woiselle@safrangroup.com  
<https://orcid.org/0000-0003-2177-7794>

**Résumé** – L’avènement des radio-interféromètres nouvelle génération tels que LOFAR, MeerKAT et ASKAP permet d’étudier de plus en plus des sources radio “transitoires”, traduisant des mécanismes physiques d’accélération aux hautes énergies (*e.g.* pulsars, “fast radio bursts” et objets magnétisés du système solaire). La recherche de telles sources dans des énormes bases de données représente un nouveau défi nécessitant des algorithmes de traitement du signal performants. Dans cet étude, on formule la recherche des sources transitoires comme un problème inverse de déconvolution de série temporelle d’images et on propose une nouvelle architecture de réseau de neurones pour résoudre ce problème. On montre expérimentalement sur des données simulées que la méthode proposée est plus performante que CLEAN, l’algorithme de déconvolution le plus utilisé en radio interférométrie.

**Abstract** – Next generation radio facilities like LOFAR, MeerKAT and ASKAP allow for advanced studies of transient radio sources, that feature physical mechanisms involving high energy accelerations (pulsars, “fast radio bursts” and solar-system magnetised objects). Search for such sources in large datasets is a new challenge that requires competitive signal reconstruction algorithms. In our study, we formulate the search for transient sources as a deconvolution inverse problem of image time series and we propose a new neural network architecture to solve this problem. We empirically show on simulated data that the proposed method outperforms CLEAN, the most popular deconvolution algorithm in radio interferometry.

## 1 Introduction et problème

Les radio-interféromètres multi-antennes nouvelle génération (tels que les précurseurs de SKA : LOFAR [8], MeerKAT et ASKAP [2] permettent d’atteindre des résolutions angulaire, temporelle et spectrale inégalées et de très bonnes sensibilités instantanées. Ils nous permettent notamment d’étudier des sources radio “transitoires”, associées à des mécanismes physiques d’accélération aux hautes énergies (*e.g.* “fast radio bursts”, pulsars et objets magnétisés du système solaire). Par ailleurs, la recherche de telles sources dans des énormes bases de données représente un nouveau défi qui nécessite des algorithmes de traitement du signal performants.

Les radio-interféromètres permettent la formation d’image par synthèse d’ouverture par corrélation de signaux de chaque paire d’antennes. Le problème mal posé à résoudre est la déconvolution du ciel observé par la réponse instrumentale (Point Spread Function - PSF) de l’instrument :

$$y = h * x + \eta \quad (1)$$

avec  $x$  l’image vérité terrain,  $y$  est l’image corrompue (*dirty image*) et  $h$  l’opérateur de mesure, la PSF, représentant l’opérateur de mesure de l’instrument et  $\eta$ , le bruit additif. Dans le cas

d’un interféromètre, la mesure et le problème se représentent dans l’espace de Fourier (espace des visibilitées) :

$$V_y = MV + \epsilon \quad (2)$$

avec  $V_y$  est l’ensemble des visibilitées mesurées,  $V$ , la vraie visibilité (une transformée de Fourier du ciel),  $M$  l’opérateur d’échantillonnage de l’interféromètre,  $\epsilon$  le bruit additif gaussien lié à la mesure des échantillons (On note  $\sigma_\epsilon$  la déviation standard du bruit  $\epsilon$  dans l’espace de Fourier). Par transformée de Fourier, l’équation (2) se ramène à l’équation (1) liant le problème de convolution du ciel par la PSF à l’échantillonnage de la transformée de Fourier du ciel.

La résolution du problème d’imagerie statique a fait l’objet de développement à partir de l’algorithme CLEAN [5]. Les améliorations suivantes prennent en compte la morphologie des sources [1], leur dépendance spectrale [7] ou leur représentation parcimonieuse [3, 6]. Il existe cependant peu de méthodes tenant compte de la structure temporelle des sources.

Pour permettre l’imagerie robuste de sources “transitoires”, autrement que par résolution de l’équation (1) trame par trame, on adapte le problème d’imagerie intégrant la dépendance temporelle du ciel et de l’instrument. Les différentes quantités et

les opérateurs deviennent alors toutes dépendantes du temps :

$$y_t = h_t * x_t + \eta_t, \quad t \in I = \{t_0, \dots, t_{T-1}\} \quad (3)$$

avec  $\eta_t$  de niveau  $\sigma_t = \|M_t\|_2 \cdot \sigma_\epsilon$ . En empilant  $\{y_t\}_{t \in I}$ ,  $\{x_t\}_{t \in I}$  et  $\{h_t\}_{t \in I}$  dans la dimension temporelle, on obtient respectivement un cube vérité terrain (*dirty cube*) et un cube de PSF, notés respectivement  $Y$ ,  $X$  and  $H$ .  $I$  est l'ensemble ordonné des indices temporels sur une durée d'observation. Dans le problème mono-image (1),  $h$  dépend de l'intégration en temps et en fréquence d'une observation et  $x$  est le ciel supposé statique. Pour le problème d'imagerie dynamique, on admet une variation du ciel ainsi que celle de la réponse de l'interféromètre. Le masque  $M_t$  échantillonne la transformée du ciel aux différents instants, rendant possible la capture de sources dépendantes du temps. Le cube  $X$  est essentiellement constitué de sources ponctuelles statiques et d'amplitude constante dans le temps et des sources ponctuelles transitoires qui apparaissent et disparaissent dans le temps. Leurs localisations dans le ciel sont supposées constantes dans le temps. On suppose que chaque source a une résolution angulaire inférieure à la taille angulaire de la PSF, et peut être représentée par un point. Les méthodes fondées sur CLEAN deviennent rapidement inadaptées dès que la dépendance du ciel se combine à celle de l'instrument. Il convient d'avoir une méthode capable d'apprendre, d'une part, le comportement temporel du télescope, et d'autre part, de celui du ciel. Nous présentons ici une méthode fondée sur l'apprentissage profond.

## 2 Méthode

Dans le cadre du problème de déconvolution de série temporelle d'images (3), les méthodes de déconvolution d'image statique [5, 1, 7, 3, 6] pourraient être appliquées image par image, de façon indépendante. Cela ne capturera pas la structure temporelle des amplitudes des sources qui évoluent, donc sera sous-optimale. Dans cette étude, nous proposons un algorithme qui exploite cette structure temporelle. Étant données des performances en restauration d'images et de vidéos de plus en plus élevées permises par l'apprentissage profond [11, 10], nous proposons de résoudre le problème (3) en utilisant un réseau de neurones.

On propose un réseau de neurones  $f$  qui réalise l'application suivante :  $\hat{X} = f(Y, H, \sigma_\epsilon; \theta)$ . Son entrée contient le cube dégradé, mais aussi les informations sur la dégradation ( $H, \sigma_\epsilon$ ), de façon non-aveugle, similairement à [11].  $\theta$  dénote l'ensemble de paramètres du réseau appris à partir des données d'entraînement  $\{Y_i, X_i, H_i, \sigma_{\epsilon,i}\}$ .

On découple  $f$  en deux modules, qui effectuent des traitements séquentiels en partant des données d'entrées. Le premier module est un réseau avec des convolutions 2D qui encode indépendamment et successivement chaque image du cube dégradé en des cartes d'activation intermédiaires. Chaque encodage prend en compte la PSF et le niveau du bruit qui ont été utilisés pour dégrader l'image. En effet, ces cartes tra-

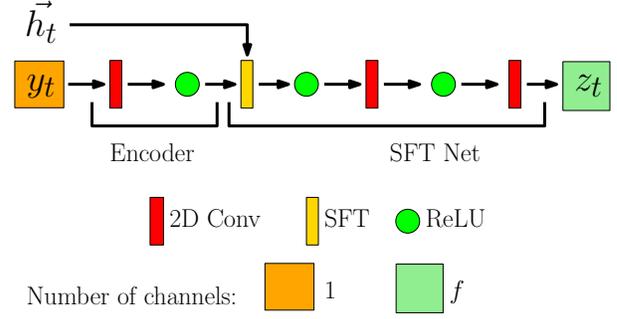


FIGURE 1 – 2D Net. Chaque couche de convolution produit  $f$  cartes d'activation.  $f = 32$  dans notre étude. La taille du noyau pour chaque couche de convolution est de  $3 \times 3$ .

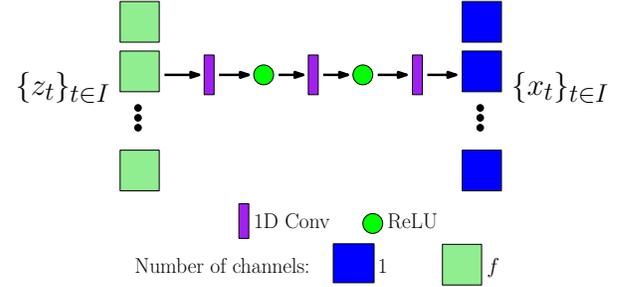


FIGURE 2 – 1D Net.  $\{z_t\}_{t \in I}$  ont des dimensions  $C \times T \times H \times W$ . Chaque couche de convolution produit  $f$  cartes d'activation, sauf la dernière qui produit des images avec un canal chacune. La taille du noyau pour chaque couche de convolution 1D est de 5.

duisent la déconvolution de cette image. On nomme ce module 2D Net. Après ces déconvolutions indépendantes, toutes ces cartes sont empilées dans la dimension temporelle et fournies au deuxième module. Ce dernier capture la structure temporelle au sein de ces cartes et restaure le cube. Le profil temporel d'une source a une continuité dans le temps, ce qui justifie le choix de l'architecture. Comme les sources ponctuelles ne se déplacent pas dans le ciel, dans ce deuxième module des convolutions 1D dans la dimension temporelle sont adaptées pour extraire cette structure temporelle. On nomme ce module 1D Net. En réunissant ces deux modules, on nomme notre réseau 2D-1D Net.

L'architecture du 2D Net est résumé par la Fig. 1. L'ensemble des cartes d'activation intermédiaires est noté  $z_t$ . La PSF  $h_t$  et le niveau du bruit qui ont été utilisés pour la dégradation de l'instant  $t$  sont encodés dans le vecteur  $\vec{h}_t$ . Il est construit de la façon suivante. D'abord, chaque  $h_t$ , originellement de taille  $l \times l$  ( $l = 256$  dans notre travail) est rogné centralement avec une taille  $r \times r$  ( $r = 96$  dans notre travail). Ensuite, cette PSF rognée est projetée dans un espace linéaire de dimension  $b$  à l'aide d'une matrice de composantes principales  $P \in \mathbb{R}^{b \times r^2}$ .  $P$  est apprise sur l'ensemble des cubes de PSF d'entraînement  $\{H_i\}$ . On note  $h_t^b$  cette PSF projetée. Cette PSF est finalement concaténée avec  $\sigma_t = \|M_t\|_2 \cdot \sigma_\epsilon$ ,

ce qui donne  $\vec{h}_t$ . Ce dernier est utilisé pour la déconvolution à l'aide de la couche *Spatial Feature Transform (SFT)*, introduite par [9] puis utilisée dans [4]. Elle permet de moduler les cartes d'activation en tenant compte de  $\vec{h}_t$ . Pour ce faire, elle applique une transformation affine aux cartes d'activation conditionnée par les cartes de dégradation obtenues à partir de  $\vec{h}_t$ .

L'architecture du module 1D Net est résumée par la Fig. 2. Les  $\{z_t\}_{t \in I}$  sont empilés dans la dimension temporelle, ce qui donne un tenseur de dimensions  $C \times T \times H \times W$ . Ce dernier est fourni à l'ensemble de trois couches de convolution 1D et des fonctions d'activation ReLU entrelacées. La convolution 1D est implémentée avec un noyau de taille 5, ce qui donne un champ temporel récepteur de  $5 + 4 + 4 = 13$  instants temporels, qui est suffisamment grande au vue des temps caractéristiques des phénomènes transitoires. Finalement, l'Algorithme 1 résume le 2D-1D Net.

---

**Algorithm 1** 2D-1D Net.  $C_r$  désigne l'opération qui croppe centralement.

---

**Input :**  $(Y, H, \sigma_\epsilon) = (\{y_t\}_{t \in I}, \{h_t\}_{t \in I}, \sigma_\epsilon)$

**Output :**  $\hat{X}$

```

1:  $S \leftarrow$  an empty list
2: for all  $t \in [0, T - 1]$  do
3:    $h_t^b \leftarrow PC_r h_t$ 
4:    $\sigma_t \leftarrow \|M_t\|_2 \cdot \sigma_\epsilon = \|\mathcal{F}(h_t)\|_2 \cdot \sigma_\epsilon$ 
5:    $\vec{h}_t \leftarrow \text{Concat}(h_t^b, \sigma_t)$ 
6:    $f_t \leftarrow \text{Encoder}(y_t)$ 
7:    $z_t \leftarrow \text{SFT Net}(f_t, \vec{h}_t)$ 
8:    $S.append(z_t)$ 
9: end for
10:  $Z \leftarrow \text{Stack}(S)$ 
11:  $\hat{X} \leftarrow \text{1D Net}(Z)$ 
12: return  $\hat{X}$ 

```

---

## 3 Expériences

### 3.1 Données

On génère des bases de données disjointes d'entraînement, de validation et de test. La base d'entraînement contient des cubes vérité terrain et des cubes de PSF. Les bases de validation et de test contiennent chacune des cubes vérité terrain et un cube de PSF. Les *dirty cubes* sont générés en se basant sur le problème (3) à partir des cubes vérité terrain.

Concernant les cubes de PSFs, on simule la réponse interférométrique de l'instrument MeerKAT, composé de 64 antennes et travaillant dans la bande radio L (environ 1420 MHz). La fréquence d'observation est fixée à 1420 MHz. En se basant sur ces paramètres, on génère des cubes de PSF. Chaque cube comprend 32 PSF intégrées successivement sur des intervalles successifs de 15 minutes, couvrant une durée d'observation totale de 8h typique d'observation journalières. Chaque PSF a une taille de  $256 \times 256$ . En reprenant la notation du problème (3), on a  $T = 32$  et  $t_{i+1} - t_i = 15$  minutes. On

génère 435 cubes de PSF d'entraînement, un cube de PSF de validation et un cube de PSF de test. La matrice de composantes principales  $P$  est apprise sur toutes les PSFs d'entraînement et le nombre de composantes  $b = 50$  explique 90 % de la variance totale.

Concernant les cubes de ciel vérité terrain, on suppose que les sources sont des points non résolus dans une image de ciel. On fixe la taille du pixel de cette image à  $1.5''$ , et on considère que dans une image de ciel de taille  $30' \times 30'$  (*i.e.*  $1200 \times 1200$  pixels) il y a au plus 30 sources constantes et 2 sources transitoires. On génère 39000 cubes de ciel d'entraînement de dimensions  $32 \times 256 \times 256$  en respectant cette distribution des sources dans le ciel. Ces données sont divisées en trois tiers contenant de zéro à deux sources transitoires. Les ensembles de validation et de test comportent chacun 66 cubes, dont la moitié ont une source transitoire, et l'autre deux.

L'amplitude d'une source constante est choisie aléatoirement entre 1 et 100. L'amplitude d'une source transitoire est une fonction discrète du temps  $A(t_i)$  sur  $I$  échantillonnée d'une fonction  $A(t)$  qui représente le profil temporel de la source. Pour chaque source, ce profil est choisi aléatoirement parmi trois modèles : *porte*, *gaussien* et *skew-normal*. Leurs paramètres sont sélectionnés aléatoirement et l'amplitude maximale de la source transitoire est choisie entre  $[50, 100]$ .

### 3.2 Entraînement et évaluation

On entraîne le 2D-1D Net sur les données d'entraînement à l'aide de la descente de gradient mini batch. La taille d'un batch est fixée à 4 exemples. Pour chaque exemple, un cube de PSF d'entraînement est aléatoirement échantillonné. Le niveau de bruit  $\sigma_\epsilon$  est échantillonné aléatoirement entre 0 et 6. On adopte l'erreur quadratique moyenne comme fonction de coût.

Étant donnée une source constante ou transitoire localisée à une certaine position dans un cube vérité terrain  $X$ , on peut définir d'abord un sous-cube obtenu en rognant spatialement le voisinage de la source avec une taille  $p \times p$  (avec  $p = 3$ ). On peut extraire de la même façon un sous-cube dans le cube estimé. On peut alors calculer l'erreur quadratique moyenne entre les deux sous-cubes. On note  $MSE_s$  pour la source  $s$  cette quantité. Elle quantifie la fidélité de la restauration du profil temporel de la source. On peut lui appliquer une fonction racine carrée puis une normalisation par la norme du sous-cube relatif au cube vérité terrain. On note  $NRMSE_s$  pour la source  $s$  cette quantité. On surveille l'entraînement en utilisant la moyenne de  $MSE_s$  sur les cubes de validation. Une PSF de validation est utilisée.

Pour mesurer la qualité du signal d'entrée, pour chaque source  $s$  dans le *dirty cube* d'entrée  $\{y_t\}_{t \in I}$  ayant pour coordonnées  $(i_s, j_s)$ , on définit le rapport signal sur bruit  $SNR_s$ , d'après la définition en radioastronomie, de la façon suivante :

$$SNR_s = \sum_{i,j \in \mathcal{D}_s} \frac{y_\tau[i,j]}{\sigma_b} \quad (4)$$

où  $\tau$  fait référence à la localisation du transitoire. Il s'agit de

l'instant temporel où l'amplitude de la source transitoire est maximale. Pour une source constante, on choisit  $\tau = 15$ .  $\mathcal{D}_s$  est le voisinage de taille  $3 \times 3$  centré sur  $(i_s, j_s)$ .  $\sigma_b$  est le niveau de bruit estimé sur  $y_\tau$  excluant  $\mathcal{D}_s$ .

En outre, pour mesurer la qualité de suppression du bruit de fond, pour un cube vérité terrain on exclue les sous-cubes de taille spatiale  $3 \times 3$  contenant chacun une source au centre. On effectue la même opération pour le cube estimé correspondant et on calcule la racine de l'erreur quadratique moyenne entre les deux. On note cette métrique  $\text{RMSE}_{\text{noise}}$ .

On compare notre réseau à la méthode qui déconvolue chaque image avec l'algorithme populaire CLEAN [5].

## 4 Résultats

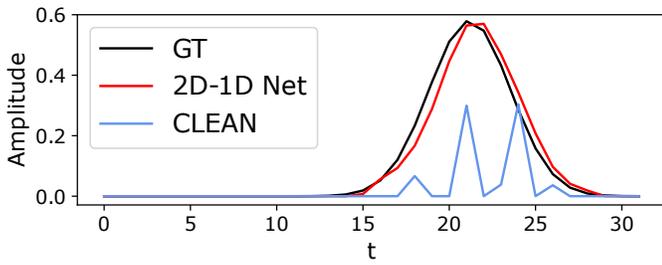


FIGURE 3 – Reconstruction d'un profil de source transitoire avec niveau de bruit  $\sigma_\epsilon = 3$ .

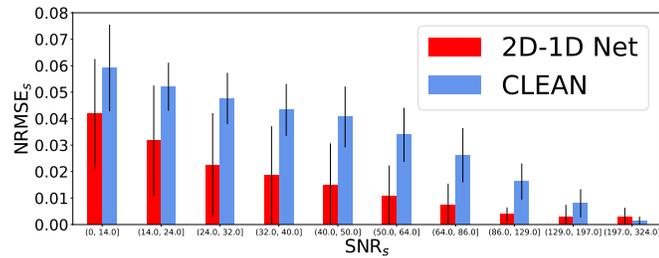


FIGURE 4 –  $\text{NRMSE}_s$  moyennés sur des sources appartenant à la même intervalle délimitée par des déciles en  $\text{SNR}_s$ . La barre verticale atteste de l'écart-type.

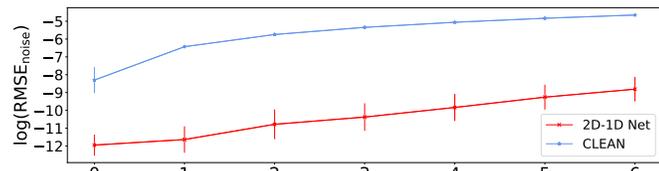


FIGURE 5 –  $\log(\text{RMSE}_{\text{noise}})$  moyenné sur les cubes de ciel test à différents niveau de bruit. La barre vertical indique l'écart type.

On utilise un cube de PSF de test et on effectue des inférences sur les cubes de ciel de test avec des niveaux de bruit :  $\sigma_\epsilon \in$

$\{0, 1, 2, 3, 4, 5, 6\}$ . La Fig. 3 compare les reconstructions d'un profil de source transitoire par CLEAN et le réseau 2D-1D Net, au niveau de bruit  $\sigma_\epsilon = 3$ . On observe que la reconstruction par la méthode proposée est plus fidèle. En particulier, malgré la présence de bruit, le début et la fin de l'évènement transitoire sont mieux reconstruits. En agrégeant les résultats de tous les inférences avec les différents niveaux de bruits, la Fig. 4 montre les  $\text{NRMSE}_s$  moyennés sur des sources appartenant à la même intervalle délimitée par des déciles en  $\text{SNR}_s$ . On observe que notre 2D-1D Net donne des meilleurs résultats que CLEAN en moyenne, hormis quand le SNR est très élevé (plus élevé que 197). Ce cas correspond, en effet, au cas où le niveau de bruit  $\sigma_\epsilon$  est nul et CLEAN est en effet optimal.

Dans la Fig. 5 on compare les  $\text{RMSE}_{\text{noise}}$  entre CLEAN et notre réseau, moyenné sur tous les cubes de ciel test à différents niveaux de bruit. On observe que pour tout niveau de bruit, notre réseau supprime de façon meilleure le bruit de fond.

## 5 Conclusions

Cette étude s'est intéressée à l'imagerie dynamique par synthèse d'ouverture à partir de données interférométriques en radio. On a cherché à résoudre le problème inverse de déconvolution pour faciliter la recherche des sources radio transitoires. Dans cette optique, on a proposé une nouvelle architecture de réseau de neurones qu'on a nommé 2D-1D Net. On a expérimentalement montré sur des données simulées que la méthode proposée est plus performante que CLEAN, l'algorithme de déconvolution le plus utilisé en radio interférométrie.

## Références

- [1] F. Abrantes, C. Lopes, T. Rodrigues, I. Gil, L. Witt, J. Grimalt, and I. Harris. Proxy calibration for instrumental data set : Implications for paleoceanographic reconstructions. *Geochemistry, Geophysics, Geosystems*, 10(9) :Q09U07, Sept. 2009.
- [2] P. E. Dewdney, P. J. Hall, R. T. Schilizzi, and T. J. L. W. Lazio. The Square Kilometre Array. *IEEE Proceedings*, 97(8) :1482–1496, Aug. 2009.
- [3] H. Garnden, J. Girard, J.-L. Starck, S. Corbel, C. Tasse, A. Woiselle, J. McKean, A. S. Van Amesfoort, J. Anderson, I. Avruch, et al. Lofar sparse image reconstruction. *Astronomy & Astrophysics*, 575 :A90, 2015.
- [4] J. Gu, H. Lu, W. Zuo, and C. Dong. Blind super-resolution with iterative kernel correction. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1604–1613, 2019.
- [5] J. Högbom. Aperture synthesis with a non-regular distribution of interferometer baselines. *Astronomy and Astrophysics Supplement Series*, 15 :417, 1974.
- [6] F. Nammour, U. Akhaury, J. N. Girard, F. Lanusse, F. Sureau, C. Ben Ali, and J. L. Starck. ShapeNet : Shape Constraint for Galaxy Image Deconvolution. *in press*, page arXiv :2203.07412, Mar. 2022.
- [7] U. Rau and T. J. Cornwell. A multi-scale multi-frequency deconvolution algorithm for synthesis imaging in radio interferometry. *Astronomy & Astrophysics*, 532 :A71, 2011.
- [8] M. P. van Haarlem, M. W. Wise, A. Gunst, G. Heald, J. P. McKean, J. W. Hessels, A. G. de Bruyn, R. Nijboer, J. Swinbank, R. Fallows, et al. Lofar : The low-frequency array. *Astronomy & Astrophysics*, 556 :A2, 2013.
- [9] X. Wang, K. Yu, C. Dong, and C. Change Loy. Recovering realistic texture in image super-resolution by deep spatial feature transform. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 606–615, 2018.
- [10] L. Xu, J. S. Ren, C. Liu, and J. Jia. Deep convolutional neural network for image deconvolution. In Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 27. Curran Associates, Inc., 2014.
- [11] K. Zhang, W. Zuo, and L. Zhang. Learning a single convolutional super-resolution network for multiple degradations. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3262–3271, 2018.