



OPEN ACCESS

EDITED BY

Didier Fraix-Burnet,
UMR5274 Institut de Planétologie et
d'Astrophysique de Grenoble (IPAG), France

REVIEWED BY

Lior Shamir,
Kansas State University, United States
Hongyi Bai,
Heilongjiang University, China

*CORRESPONDENCE

Utsav Akhauri,
utsav.akhauri@epfl.ch

SPECIALTY SECTION

This article was submitted to Astrostatistics,
a section of the journal Frontiers in
Astronomy and Space Sciences

RECEIVED 22 July 2022

ACCEPTED 31 October 2022

PUBLISHED 18 November 2022

CITATION

Akhauri U, Starck JL, Jablonka P, Courbin F
and Michalewicz K (2022), Deep
learning-based galaxy image
deconvolution.
Front. Astron. Space Sci. 9:1001043.
doi: 10.3389/fspas.2022.1001043

COPYRIGHT

© 2022 Akhauri, Starck, Jablonka, Courbin
and Michalewicz. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License \(CC
BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in
other forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Deep learning-based galaxy image deconvolution

Utsav Akhauri^{1*}, Jean-Luc Starck², Pascale Jablonka¹,
Frédéric Courbin¹ and Kevin Michalewicz¹

¹Laboratory of Astrophysics, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, ²AIM CEA CNRS Université Paris-Saclay Université de Paris, Gif-sur-Yvette, France

With the onset of large-scale astronomical surveys capturing millions of images, there is an increasing need to develop fast and accurate deconvolution algorithms that generalize well to different images. A powerful and accessible deconvolution method would allow for the reconstruction of a cleaner estimation of the sky. The deconvolved images would be helpful to perform photometric measurements to help make progress in the fields of galaxy formation and evolution. We propose a new deconvolution method based on the Learnlet transform. Eventually, we investigate and compare the performance of different Unet architectures and Learnlet for image deconvolution in the astrophysical domain by following a two-step approach: a Tikhonov deconvolution with a closed-form solution, followed by post-processing with a neural network. To generate our training dataset, we extract HST cutouts from the CANDELS survey in the F606W filter (V-band) and corrupt these images to simulate their blurred-noisy versions. Our numerical results based on these simulations show a detailed comparison between the considered methods for different noise levels.

KEYWORDS

deconvolution, denoising, image processing, deep learning, inverse problem, regularization

1 Introduction

In the upcoming decade, large telescopes such as the Vera C. Rubin Observatory (LSST) and Euclid will offer a broader view of the universe by capturing several images. Both these telescopes would be covering a huge part of the sky, thus giving us access to a wide variety of objects, and capturing images in various frequency bands. Each frequency band captures certain unique information, which is beneficial for tracing the constituent components of galaxies. In addition, these surveys aim to go deeper—thus capturing fainter objects, and to higher redshifts—thus capturing very distant objects. A huge variety of galaxy images would help us better understand their origin and evolution. From an astrophysical point of view, the ultimate aim is to translate the information carried by these images and derive physical inferences (for e.g.: obtaining the metallicity, morphology, or flux of different galaxies). Unfortunately, imperfections are generated by every image acquisition system. Usually, a blurring effect is introduced in the images, which is modeled by a Point Spread Function (PSF) and is considered to be space-invariant. Moreover, the sensor variations introduce noise in the images,

which is usually additive, white, and Gaussian. As such, there is a dire need to develop fast and accurate deconvolution algorithms that generalize well to different images. In addition, bad pixels or cosmic ray hits also need to be taken into account. Generally, they are relatively easily identifiable, and a good solution would be to apply an in-painting method to replace the bad pixel values by reasonable ones. In this paper, we assume that this pre-processing step has been carried out, and that we can directly deal with the clean images. A powerful and accessible deconvolution method would allow for the reconstruction of a cleaner estimation of the sky. Deconvolution is useful for a broad range of applications, such as galaxy morphology studies, substructure identification in galaxies, bulge/disk separation, etc. Deconvolution is also very important for comparing two images at different resolutions. Hence, technical development in the field of image processing is essential to answering the fundamental questions in astrophysics.

By using the least-squares method, one can partially reconstruct the image; however, the solution oscillates while solving the equation since the problem is ill-conditioned. Broadly speaking, it belongs to the family of ill-posed problems that could alternatively be handled using regularization. Due to these effects, simply minimizing the Mean Squared Error (MSE) between the observation and the reconstruction does not lead us to proper convergence. To tackle this, constraints associated with the signal's energy, such as its derivatives (Bertero and Boccacci, 1998), total variation (Rudin et al., 1992; Chambolle et al., 2010) or sparsity (Starck et al., 2015; Farrens et al., 2017), could be added. These routinely used methods are apt for solutions optimizing the MSE. For such inverse problems, sparse wavelet regularization using the ℓ_0 or ℓ_1 norm has remained the routinely used approach for astrophysical image deconvolution. It has led to striking results, such as an improvement in resolution by a factor of four in the Cygnus-A radio image reconstruction compared to the standard CLEAN algorithm (Garsden et al., 2015). Sparsity could be considered a weak prior on the distribution of the wavelet coefficients of the solution. This comes from the fact that most images could be represented in a more compressed manner in the wavelet domain. The advancement in deep learning over the recent past has presented encouraging outcomes in the field of deconvolution (Xu et al., 2014). Within the astrophysics community, deep learning approaches have been introduced to carry out model fitting that could be considered as a parametric deconvolution (Tuccillo et al., 2018). Particularly, Unets (Ronneberger et al., 2015) have gathered tremendous attention due to their effectual performance, which can be attributed to their highly non-linear processing and the availability of large training datasets. Based on Unet, Sureau et al. (2020) developed the Tikhonet method for deconvolving galaxy images in the optical domain. Tikhonet is a two-step deep learning-approach, the first being a Tikhonov deconvolution, i.e., with a standard quadratic regularization, followed by a neural network denoising

using a 4-scale *XDense Unet*. Giving the Tikhonov deconvolution as an input to the network, the training aims at minimizing the MSE between the reconstruction and the ground truth image. It was shown that Tikhonet surpassed sparse regularization-based methods in terms of the MSE and a shape criterion, where a measure of the galaxy ellipticity was used to encode its shape (Sureau et al., 2020).

While the multiple instances of non-linearity are inherent to artificial neural networks, the idea of learning with training data could even be extended to methods that exploit sparsity. Recently, Ramzi et al. (2021) presented a novel architecture, termed Learnlet, that preserves the properties of sparsity-based approaches (for example, linear decomposition and reconstruction steps, good generalization properties, exact reconstruction) while simultaneously capturing the prowess of neural networks. The Learnlet decomposition aims at learning a filter bank in a denoising setting with backpropagation and gradient descent. Learnlets were originally proposed for denoising and have been shown to have better generalization properties than Unets (Ramzi et al., 2021), while Unets showed better performances. In this paper, we investigate if the Learnlet neural network could also be a good alternative to Unet for deconvolution. We propose a new deconvolution approach based on the Learnlet decomposition, using the same two-step approach as in Sureau et al. (2020), but by substituting the Unet denoiser by Learnlet. We compare our results on images extracted from the CANDELS survey (Grogin et al., 2011; Koekemoer et al., 2011). In Section 2, we introduce the deconvolution problem and the deep learning-methods that have been developed to tackle them. In Section 3, we discuss the concept of Learnlet decomposition as introduced by Ramzi et al. (2021), and extend the idea to use the network for image deconvolution. We detail out the process of generating our dataset and performing the numerical experiments in Section 4, and in Section 5, we demonstrate the results obtained for these experiments. Finally, in Section 6, we conclude our work.

2 Deep learning-based deconvolution

2.1 The deconvolution problem

Let $\mathbf{y} \in \mathbb{R}^{n \times n}$ be the observed image and $\mathbf{h} \in \mathbb{R}^{n \times n}$ be the PSF. The observed image can be modelled as

$$\mathbf{y} = \mathbf{h} * \mathbf{x}_t + \eta \quad (1)$$

where $\mathbf{x}_t \in \mathbb{R}^{n \times n}$ denotes the target image, $*$ denotes the convolution operation, and $\eta \in \mathbb{R}^{n \times n}$ denotes additive Gaussian noise. Our aim is to recover the ground truth image \mathbf{x}_t . One could partly recover an estimate $\hat{\mathbf{x}}$ of \mathbf{x}_t by the help of the least-squares method. However, the solution would oscillate because Eq. 1

is ill-conditioned. More broadly, it falls under the category of an ill-posed problem, which could be handled by regularization (Bertero and Boccacci, 1998).

Convolutional Neural Network architectures like Unet (Ronneberger et al., 2015) have shown to be very efficient in image noise removal, as seen in Gurrola-Ramos et al. (2021). The Unet was originally developed for biomedical image segmentation. Since then, it has been found to be relevant to many other imaging problems, not just segmentation. Meanwhile, denoising methods based on wavelets are no longer the state-of-the-art. However, they have theoretical guarantees [as seen in Donoho (1995)]. Unets consist of a multi-scale approach similar to wavelets, which allows the signal to be analyzed at multiple resolutions. The prime difference comes from the usage of non-linearities at various steps. Wavelets include only a single non-linearity step (wavelet shrinkage) when used for denoising. Contrastingly, Unets rely on several ReLU and max-pooling steps. The denoising analysis in Unets becomes very complicated due to such chained non-linearities. Particularly, it is tough to understand how a network that is trained on one kind of noise could be used for other kinds of noise. Moreover, a few works (Gottschling et al., 2020) even demonstrate that deep learning-based methods are unable to recover features that their classical counterparts can.

For a ground truth image $\mathbf{x} \in \mathbb{R}^{n \times n}$, let $\tilde{\mathbf{x}} = \mathbf{x} + \epsilon$ be its corrupted version generated by adding Gaussian noise $\epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_{n \times n})$ with a known covariance $\mathbf{C}_{n \times n}$. For a denoiser \mathbf{N}_θ , the general optimization problem can be represented as:

$$\operatorname{argmin} \mathbb{E}_{\tilde{\mathbf{x}}} [\|\mathbf{x} - \mathbf{N}_\theta(\tilde{\mathbf{x}})\|_2^2] \quad (2)$$

2.2 The Tikhonet solution

Tikhonet is a two-step deep learning-based deconvolution method. The first step involves deconvolving the input image with a Tikhonov filter based on a quadratic regularization. Let $\mathbf{H} \in \mathbb{R}^{n^2 \times n^2}$ be the circulant matrix associated with the convolution operator \mathbf{h} . Then, the Tikhonov solution of Eq. 1 can be written as:

$$\hat{\mathbf{x}} = (\mathbf{H}^T \mathbf{H} + \lambda \mathbf{\Gamma}^T \mathbf{\Gamma})^{-1} \mathbf{H}^T \mathbf{y}, \quad (3)$$

where $\mathbf{\Gamma} \in \mathbb{R}^{n^2 \times n^2}$ corresponds to the linear Tikhonov filter which is set to a Laplacian high-pass filter to penalize high frequencies, and $\lambda \in \mathbb{R}_+$ corresponds to the regularization weight, whose appropriate value for every image is found by Stein's Unbiased Risk Estimate (SURE) minimization (Sureau et al., 2020). The Tikhonov step results in a deconvolved image that contains correlated additive noise which can be removed in a following step by a four-scale *XDense Unet*. The training is aimed to make the network learn the mapping from the Tikhonov output $\hat{\mathbf{x}}$ to the ground truth image \mathbf{x}_t while minimizing the MSE (Eq. 2).

The model architecture heavily impacts the denoising performance. It should have some multi-scale processing in order to capture distant correlations, pointing to the usage of a Unet like layout (Ronneberger et al., 2015) shown in Figure 1, which has already found success in solving inverse problems (Jin et al., 2017). Ideally, one should also aim to bring down the trainable parameter count. Considering these factors, the *XDense Unet* inherits the global layout from (Jin et al., 2017), but with these alterations (Sureau et al., 2020):

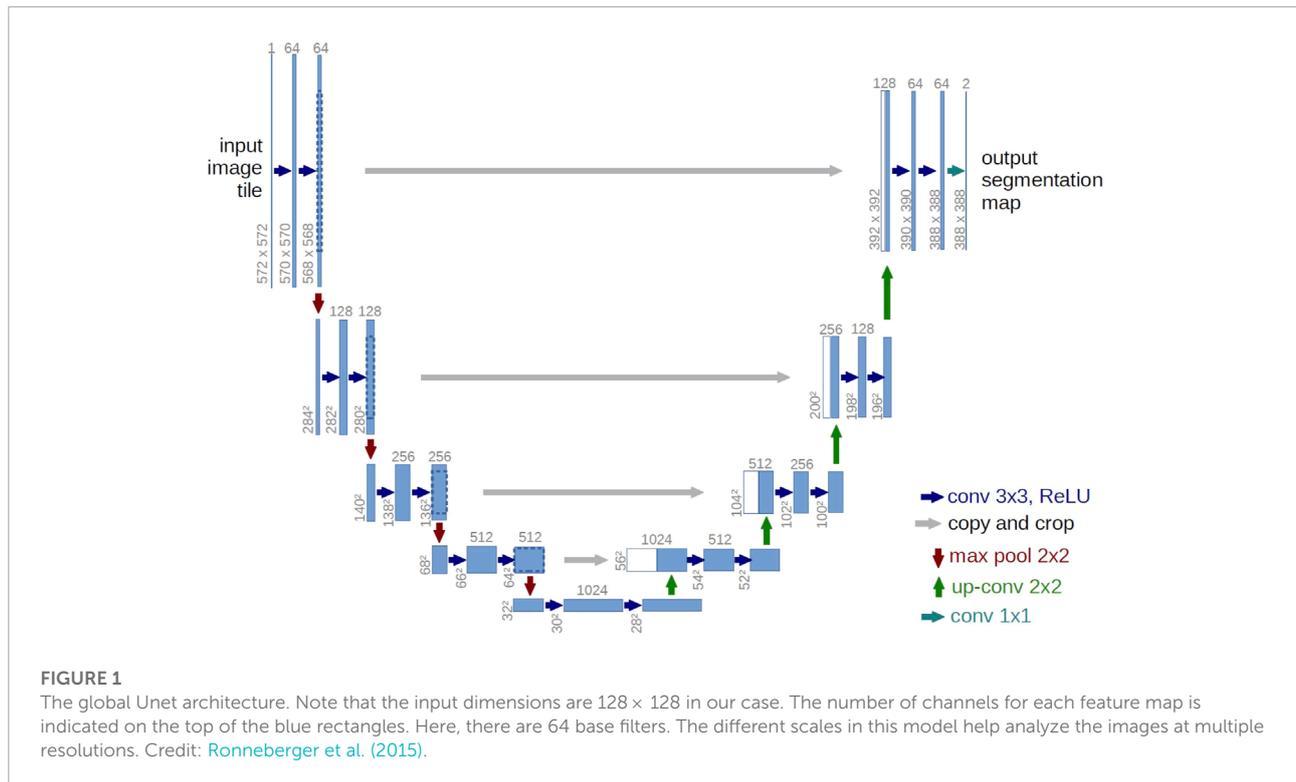
- At each scale, the convolutional layers were replaced by **dense blocks** introduced by Huang et al. (2017), which decrease the parameter count by concatenating feature maps from the previous layers to the current layer's input, and were claimed to help preserve information, enable reusing features, and limit vanishing gradients.
- 2D convolutions were substituted by **2D separable convolutions** (Chollet, 2017), which helped decrease the parameter count in the model by assuming that correlations across feature maps and spatial correlations can be independently captured.
- The max-pooling step was changed to **average pooling**, since it led to over-segmentation of the final estimates.
- The **skip connection** from the input to the output layer (Jin et al., 2017) was **removed**, since it degraded the network performance particularly at low Signal-to-Noise ratio (SNR).

The first two alterations notably bring down the parameter count per scale of the Unet, thus increasing the potential number of scales for a given number of trainable parameters. Additionally, we remove the biases from the convolutional layers, as they have been shown to lead to a low generalisation capacity (Mohan et al., 2020). In all, the *XDense Unet* model has 184,301 trainable parameters.

2.3 Advantages and drawbacks of Tikhonet for astronomical image deconvolution

Tikhonet presents many advantages:

- It provides better results than standard techniques usually used in astrophysics (Sureau et al., 2020).
- It is an extremely fast method working for both optical and radio image deconvolution, therefore well-adapted to the forthcoming big data challenges.
- It allows to easily take into account the spatial variation of the PSF. Indeed, we know that the PSF in large surveys such as the Canada-France Imaging Survey (CFIS) or



Euclid varies spatially. It was shown by [Sureau et al. \(2020\)](#) that Tikhonet can handle such data by deconvolving each galaxy in the field independently and using the PSF corresponding to the center of the galaxy. The PSF field is generally not known, but can be reconstructed from observed stars present in the image ([Liaudat et al., 2021](#); [Liaudat et al., 2022](#)).

- Many galaxies share very similar morphologies. Having a learning process allows to capture these morphologies and improve the deconvolution.

However, it also has a few drawbacks:

- Generalization properties are not so good. It was shown that Unet does not perform well on image morphologies or noise levels that were not included in the training dataset ([Ramzi et al., 2021](#)).
- It is not clear how Unets deal with images containing noise with non-white statistical properties (for e.g.: Poisson, non-stationary Gaussian noise), while sparsity-based techniques can easily consider different kinds of noise ([Starck et al., 2015](#)).
- There is no theory to support the method.

In a denoising framework, the Learnlet decomposition has shown interesting properties that could also be considered for deconvolution.

2.4 Using a deeper Unet

In addition to the compact *XDense Unet*, we also tested the performance of *Unet-64*, a deeper Unet as used in [Ramzi et al. \(2021\)](#) that has 31,023,940 trainable parameters (around 170 times more parameters than *XDense Unet*). The architecture is the same as shown in [Figure 1](#). As in the case of Tikhonet, we used the bias-free Unet architecture, as these biases lead to a low generalisation capacity ([Mohan et al., 2020](#)). The full Python implementation of the code is publicly available ¹.

3 Learnlet deconvolution

3.1 Learnlets

Denoising methods based on wavelets are no longer the state-of-the-art, but have theoretical guarantees and are the baseline for other approaches. In cases where guarantees are desired, like medical applications, they are the suitable candidates. Recently, [Ramzi et al. \(2021\)](#) presented a new network architecture, called Learnlet, which exploits one of the most desirable usefulnesses of deep learning: using gradient descent to improve the expressive

¹ <https://github.com/zaccharieramzi/understanding-Unets>.

power of wavelets, while preserving some interesting wavelet properties like exact reconstruction.

In the presence of White Gaussian Noise (WGN), if Σ denotes the set of possible values for the noise standard deviation σ , m the number of scales, and $\theta = (\theta_s, \theta_t, \theta_a) \in \Theta_m$ a given set of parameters, then Learnlets are defined as the following function f_θ from $(\mathbb{R}^{n \times n} \times \Sigma)$ to $\mathbb{R}^{n \times n}$:

$$f_\theta(\tilde{\mathbf{x}}, \sigma) = \mathbf{S}_{\theta_s} \left(\mathbf{T}_{\theta_t} \left(\mathbf{A}_{\theta_a}(\tilde{\mathbf{x}}), \sigma \right) \right) \quad (4)$$

where

1. \mathbf{A}_{θ_a} is the analysis function, which is equivalent to the wavelet transform with learned filters:

$$\mathbf{A}_{\theta_a}(\tilde{\mathbf{x}}) = \left(\left(\mathbf{F}_{\theta_a^{(i)}} * \mathbf{g}(\tilde{\mathbf{h}}^{i-1}(\tilde{\mathbf{x}})) \right)_{i=1}^m, \tilde{\mathbf{h}}^m(\tilde{\mathbf{x}}) \right) \quad (5)$$

where:

- $\mathbf{F}_{\theta_a^{(i)}}$ is the filter bank at scale i . $\theta_a^{(i)}$ are the convolution kernels all of the same square size.
- $\tilde{\mathbf{h}}: \mathbf{y} \mapsto \tilde{\mathbf{u}}(\mathbf{h} * \mathbf{y})$ is a low-pass filtering (\mathbf{h}) followed by decimation ($\tilde{\mathbf{u}}$).
- \mathbf{g} is a high-pass filter given by: $\mathbf{g}(\mathbf{y}) = \mathbf{y} - \mathbf{u}(\tilde{\mathbf{h}}(\mathbf{y}))$, where \mathbf{u} is the upsampling operation performed using a bicubic interpolator.

Only $\mathbf{F}_{\theta_a^{(i)}}$ filters are learned, while keeping the high and low pass filters (\mathbf{g}, \mathbf{h}) fixed.

2. \mathbf{T}_{θ_t} is the thresholding function. Ramzi et al. (2021) chose the soft-thresholding operation since it offers more stability (Abrial et al., 2007). Owing to the linear nature of the analysis operator, one could straightforwardly apply the thresholding strategy to any other type of noise, including non-stationary Gaussian noise. In case of white Gaussian noise with variance σ^2 , this function is described as:

$$\mathbf{T}_{\theta_t} \left(\left((\mathbf{d}_i)_{i=1}^m, \mathbf{c} \right), \sigma \right) = \left(\left(\left(t_{ij}(\mathbf{d}_{ij}, \sigma) \right)_{i=1}^m \right)_{i=1}^m, \mathbf{c} \right) \quad (6)$$

where $t_{ij}(\mathbf{d}, \sigma) = \hat{\sigma}_{ij} ST \left(\frac{1}{\hat{\sigma}_{ij}} \mathbf{d}_{ij}, \theta_T^{(ij)} \sigma \right)$, with:

- $\mathbf{d}_{ij} \in \mathbb{R}^{\frac{n}{2^{i-1}} \times \frac{n}{2^{i-1}}}$ the output of the j^{th} filter of i^{th} scale.
- $\hat{\sigma}_{ij}$ the estimated standard deviation of \mathbf{d}_{ij} when white Gaussian noise with variance 1 is input to the transform.
- $\theta_T^{(ij)}$ the level of thresholding applied to the j^{th} analysis filter at scale i .
- $ST(\mathbf{d}, \mathbf{s})$ the soft-thresholding function applied point-wise on \mathbf{d} with threshold \mathbf{s} : $ST(\mathbf{d}, \mathbf{s}) = \text{sign}(\mathbf{d}) \max(|\mathbf{d}| - \mathbf{s}, 0)$.

3. \mathbf{S}_{θ_s} is the synthesis function, which is equivalent to the wavelet reconstruction operator with learned filters and is linear. It is described recursively as:

$$\mathbf{S}_{\theta_s} \left(\left((\mathbf{d}_i)_{i=1}^m, \mathbf{c} \right) \right) = \mathbf{S}_{\theta_s}^{(m-1)} \left(\left((\mathbf{d}_i)_{i=1}^{m-1}, \mathbf{u}(\mathbf{c}) + \mathbf{F}_{\theta_s^{(m)}} * \mathbf{d}_m \right) \right) \quad (7)$$

where $\mathbf{S}_{\theta_s}(\emptyset, \mathbf{c}) = \mathbf{c}$ and:

- $\mathbf{F}_{\theta_s^{(i)}}$ is the filter bank at scale i . $\theta_s^{(i)}$ are the convolution kernels all of the same square size.
- \mathbf{u} is an upsampling operation performed by a bicubic interpolator.

To make Learnlets behave as closely as possible to wavelets and hence more intuitive, Ramzi et al. (2021) constrain the thresholding levels to lie within $[0, 5]$ and force the analysis filters to have a unit norm. Figure 2 shows a schematic of the Learnlet architecture. The full Python implementation of the code is publicly available².

3.2 Deconvolution with learnlets

Similar to Tikhonet deconvolution, in Learnlet deconvolution, we find the closed-form solution $\hat{\mathbf{x}}$ as given in 3. $\hat{\mathbf{x}}$ is then post-processed by a Learnlet network, which is also trained on pairs of Tikhonov outputs $\hat{\mathbf{x}}$ and ground truth images \mathbf{x}_t to minimize the MSE (Eq. 2). The following model parameters were chosen for the Learnlet architecture, which amounted to 372,000 trainable network parameters (Ramzi et al., 2021):

- $m = 5$ scales.
- 256 learnable analysis filters plus 1 fixed identity analysis filter $\mathbf{F}_{\theta_a^{(i)}}$ of dimensions 3×3 .
- 257 learnable synthesis filters $\mathbf{F}_{\theta_s^{(i)}}$ of dimensions 5×5 .

Additionally, Learnlets also require the standard deviation of the noise σ_{noise} as an input to the model. An interesting thing to note here is that after the Tikhonov deconvolution step, the nature of the noise changes from White Gaussian (uncorrelated) to Colored Gaussian (correlated). Even though Ramzi et al. (2021) developed Learnlets for White Gaussian Noise (WGN) removal, we observe that the algorithm is able to adapt its thresholding step (Eq. 6) and efficiently recover the solution. This is supported by the results shown in Section 5. Note that there also exist methods to estimate the noise standard deviation in an image with an accuracy of around 1% (Starck et al., 2015).

In all of the deconvolution methods -

- The Tikhonov deconvolution step remains the same, while the network that post-processes the Tikhonov output $\hat{\mathbf{x}}$ differs.
- The success lies on the training accuracy between the Tikhonov deconvolution $\hat{\mathbf{x}}$ and the ground truth image \mathbf{x}_t .
- The network used for post-processing acts as a denoiser. The efficiency and accuracy of the denoising step governs the

² <https://github.com/zaccharieramzi/understanding-Unets>.

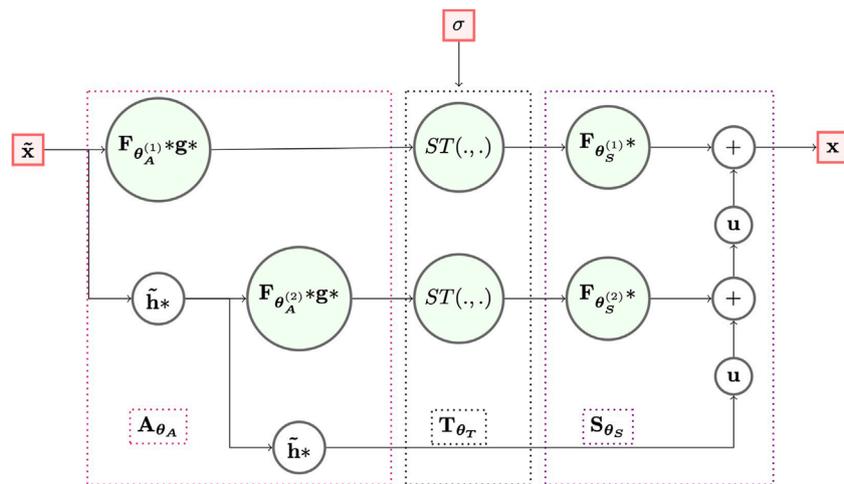


FIGURE 2

Schematic representation of the Learnlet model, with $m = 2$ scales. Inputs/outputs are represented by the red nodes. The light-green nodes denote functions with learnable parameters. Credit: Ramzi et al. (2021).

reconstructed output, which should ideally be as close to the target image \mathbf{x}_t as possible.

4 Dataset and experiments

The code was implemented in Python 3.6, using TensorFlow 2.2 (Abadi et al., 2015) for model design. All simulations and trainings were performed on the Yggdrasil supercomputer based at the University of Geneva, using a single Titan RTX Turing GPU with 24 GB RAM for each job.

4.1 Dataset generation

4.1.1 Ground truths—CANDELS

The Cosmic Assembly Near-IR Deep Extragalactic Legacy Survey, CANDELS (Grogin et al., 2011; Koekemoer et al., 2011), was conceived to record the first third of galactic evolution from $z = 8$ to 1.5 via deep imaging of more than 250,000 galaxies. The entire survey consists of five different image mosaics (GOODS-N, GOODS-S, EGS, UDS, COSMOS), each covering different regions of the sky. To generate our training dataset, using Python, we extracted cutout windows of dimensions 128×128 pixels from the CANDELS FITS image mosaics in the $F606W$ filter (V-band) by centering them at the object centroid, without any dynamic range-scaling. For larger objects, we adaptively increased the cutout window dimensions to completely enclose them. In that case, if the window dimensions exceeded 128×128 pixels, we downsampled the

images to 128×128 . To select good galaxy candidates and exclude point-sized objects, we use the following filtering criteria:

- $MAG_AUTO < 26$ (AB magnitude in SExtractor “AUTO” aperture).
- $Flux_Radius_{80} > 10$ (80% enclosed flux radius in pixels).
- $FWHM > 10$ (full width at half maximum in pixels).

In all, we end up with 22,317 ground-truth images.

4.1.2 Simulations

All the extracted ground truth images are convolved with a Gaussian PSF with an FWHM of 15 pixels, which visually blurs the images such that individual small-scale structures are lost. As seen from Eq. 3, since the algorithm takes the PSF as an input parameter and performs Tikhonov deconvolution in the first step, the method would work with any other kind of blurring kernel. The actual performance of the method is mainly dependent on the neural network training, which corresponds to the denoising step. After convolution, we add white Gaussian noise with a standard deviation σ_{noise} having a value such that the faintest object in our dataset has a peak SNR close to 1 and is hence barely visible. For this value of σ_{noise} , we observe a range of SNR values depending on the magnitude of the galaxy. Eventually, the σ_{noise} values are also required as inputs to the Learnlet architecture. Finally, the batch of noisy simulations and their corresponding ground truth images is randomly split into Train-Test subsets in the ratio 0.9:0.1

4.2 Pre-processing

Each image $\mathbf{x}_{(i)}$ was normalised by subtracting its own mean $\mu_{(i)}$ and scaling within the $[-1, 1]$ range as follows:

$$\frac{\mathbf{x}_{(i)} - \mu_{(i)}}{\max[\mathbf{x}_{(i)}^t - \mu_{(i)}^t]} \quad (8)$$

where $\mathbf{x}_{(i)}^t$ and $\mu_{(i)}^t$ denote the i^{th} target image and its corresponding mean. The denominator ensures that all target images have a maximum pixel intensity of 1, and that all noisy images are scaled with respect to their corresponding targets such that the ratios of the peak intensities between them remain the same before and after normalization. Since the flux of the target image is spread out by the PSF as a result of convolution, the noisy images have a lower peak intensity than their corresponding target images. One of the main goals of astrophysical deconvolution is to recover this drop in peak intensity. During training, we augment our dataset by performing random rotations in multiples of 90° , translations and flips along horizontal and vertical axes. Data augmentation is also beneficial for making neural networks invariant to rotations and translations.

4.3 Training

After performing a Tikhonov deconvolution as described by Eq. 3, the networks were trained to learn the mapping from the Tikhonov outputs $\hat{\mathbf{x}}$ to the ground truth images \mathbf{x}_t

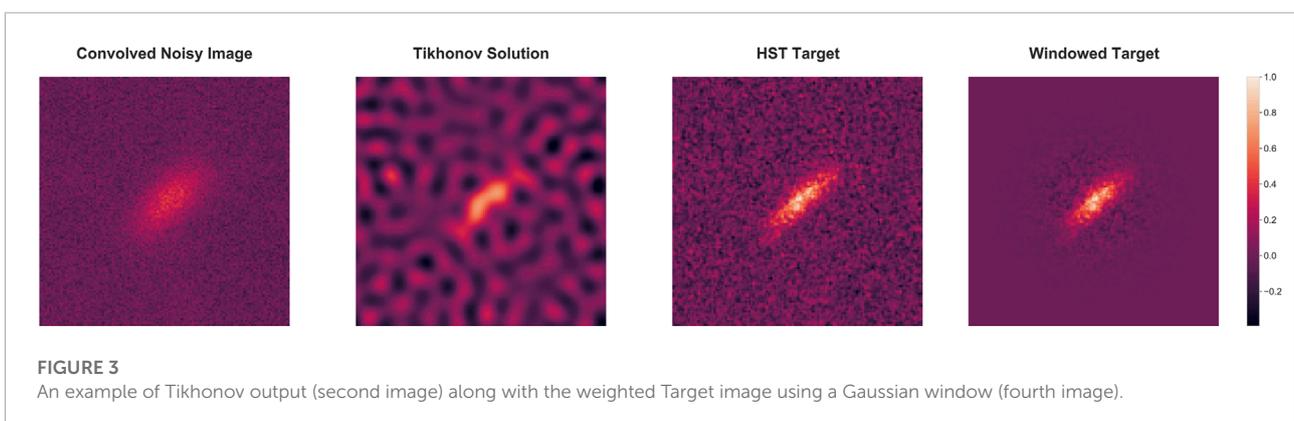
while minimizing the MSE (Eq. 2). The images were fed to the networks in mini-batches of size 32. All networks were trained until convergence was achieved. We trained the networks with an Adam optimizer (Kingma and Ba, 2014) and an initial learning rate of 10^{-3} , which was then decreased by half every 25 epochs until it reached a minimum of 10^{-5} . The way we generate our dataset (as described in Section 4.1) results in images with varying noise levels, which ensures that the training is not biased by a certain noise level. Table 1 shows a performance comparison of the three methods.

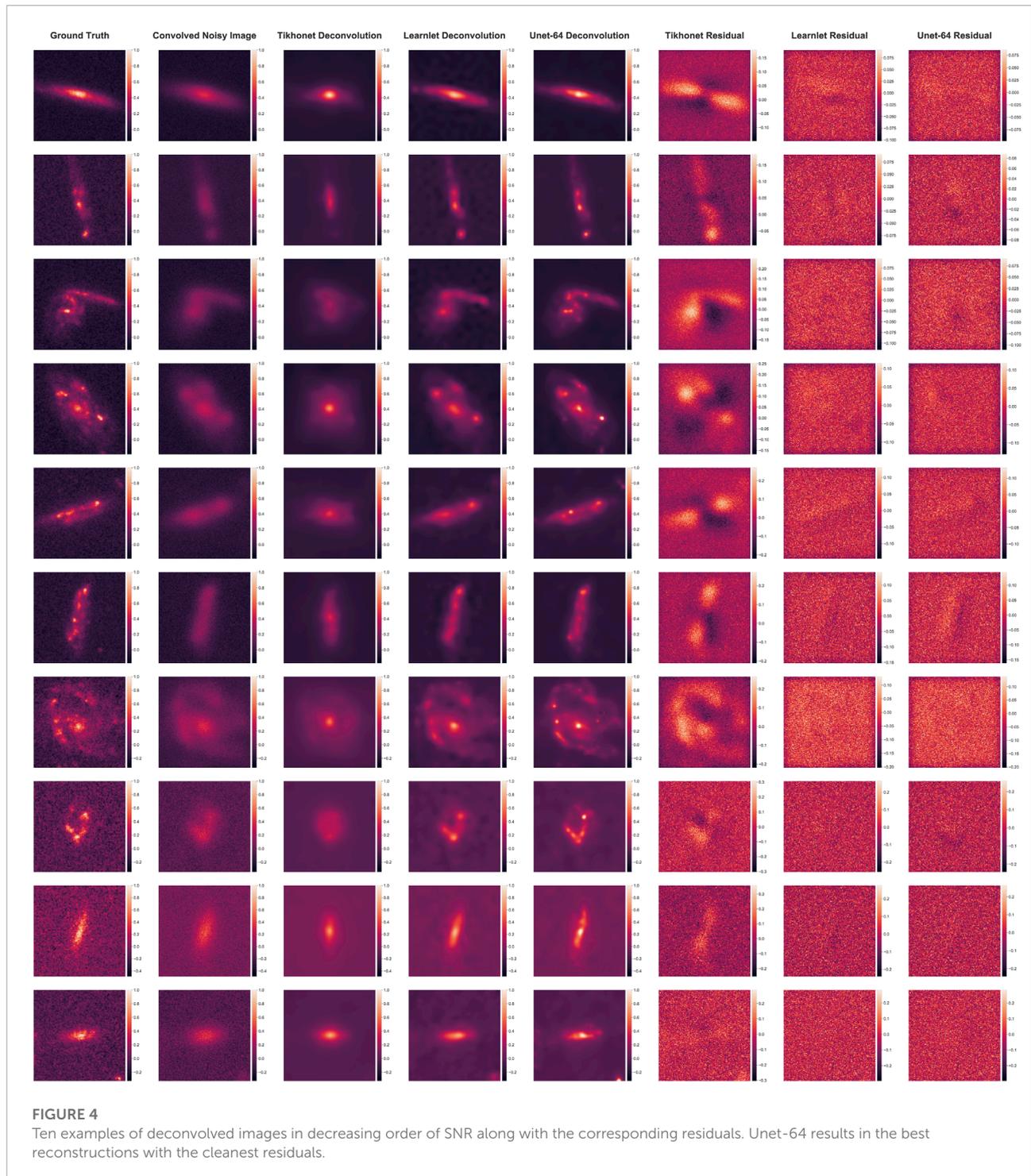
5 Results

For evaluation, our test dataset contains images with varying noise levels (4.1.2). To perform a quantitative comparison, we use the Normalised Mean Squared Error (NMSE) and the Structural Similarity Index Measure (SSIM), two commonly used metrics to quantify image-reconstruction quality. To measure the NMSE and SSIM, we weighted all images by a Gaussian window centered around the galaxy. The FWHM for each object was obtained from its catalog and used for the Gaussian weighting, thus ensuring that the window would only encircle the object and discard the noise present in the background of the target images. An illustration of the same along with the noisy Tikhonov output $\hat{\mathbf{x}}$ is shown in Figure 3. To visually analyze the final deconvolution results, in Figure 4, we show ten examples of reconstructed galaxies with different morphologies in decreasing order of SNR. We observe that Tikhonet is mostly able to reconstruct pixels around the central region. Consequently, it is able to deconvolve

TABLE 1 Performance comparison of the three deconvolution methods. The runtime per image was calculated on the same GPU on which the networks were trained.

Method	No. of parameters	Epochs	Training time (hrs.)	Runtime per image (ms)
Tikhonet	184,301	40	3.65	4.02
Learnlet	372,000	150	9.81	30.8
Unet-64	31,023,940	500	16.9	26.3

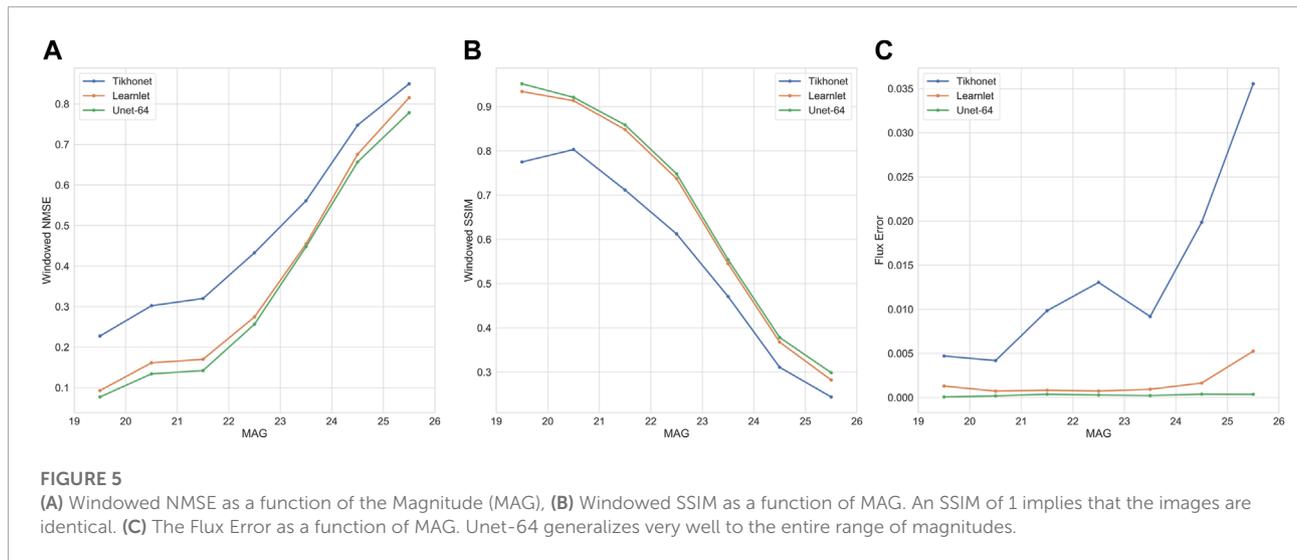




small objects but is unable generalize well to extended objects. Learnlet and Unet-64, on the other hand, are also able capture small-scale structures and non-central peaks present in the images, while well-preserving the global shape simultaneously. Moreover, visually, their performance is consistent on both high

and low SNR images, with an even further improvement for Unet-64. We also compute and show the residual maps for these three deconvolution methods. The residuals are defined as:

$$\text{Residual} = \mathbf{y} - \mathbf{h} * \mathbf{N}_{\theta}(\hat{\mathbf{x}}) \tag{9}$$



where \mathbf{y} corresponds to the noisy image, \mathbf{h} is the PSF, $\hat{\mathbf{x}}$ is the noisy Tikhonov input, and \mathbf{N}_θ is the network model.

Ideally, one should see pure noise in the residuals. We can see that there is still some structure present in the Tikhonet residuals, which comes as a result of incomplete reconstruction. The residuals are more structured in case of extended objects as compared to smaller ones, which implies that Tikhonet struggles to recover extended objects. Contrastingly, the Learnlet and Unet-64 residuals contain minimal structure with no orientational preference, which indicates that the reconstruction is better regardless of the size of the galaxy.

In **Figure 5**, we show the trends obtained for windowed NMSE and SSIM on our test dataset images binned according to their magnitudes. All networks perform better on low magnitude (or high SNR) images, and there is a decrease in performance with an increase in the magnitude (or decrease in SNR). Furthermore, for the range of magnitudes and metrics considered, Unet-64 outperforms all other methods, followed by Learnlet. The mean NMSE improvement is 11.4% when going from Tikhonet to Learnlet, and 2.2% when going from Learnlet to Unet-64. Similarly, the mean SSIM improvements are 10% and 1.2% respectively. We also compare the three networks based on their ability to preserve the image flux, which is essentially the sum of all pixel intensities. As seen in **Figure 5**, although the errors are very low for all the networks, Unet-64 performs the best in terms of flux preservation, followed by Learnlet and Tikhonet. More specifically for Tikhonet, the error steeply increases with an increase in magnitude. On the other hand, Unet-64 and Learnlet have flatter curves and generalize well to all magnitudes, with Unet-64 generalizing the best. Based on the metrics considered, we finally conclude that Unet-64 deconvolution surpasses Learnlet and Tikhonet deconvolution.

6 Conclusion

We proposed a new deconvolution method based on the Learnlet transform. Consequently, we compared the performance of Tikhonet, Learnlet, and Unet-64 for image deconvolution in the astrophysical domain by adapting a two-step approach involving a Tikhonov deconvolution followed by post-processing with a denoiser. Visually, we observed that Unet-64 and Learnlet are able to capture the small-scale structures in the images in addition to the global shape, while simultaneously well-preserving the flux. Since the networks were evaluated on a range of noise levels, we could conclude that Unet-64 and Learnlet generalize well unlike Tikhonet, with Unet-64 having the best performance. These observations are further supported by the quantitative metrics used for comparison, where Unet-64 outperforms Learnlet and Tikhonet. As seen in **Table 1**, although Tikhonet has the smallest runtime, it does not perform well on extended objects. Furthermore, Unet-64 has a smaller runtime per image compared to Learnlet, making it more efficient. This makes it an ideal candidate to be used for astrophysical image deconvolution. The strategy we proposed in this paper for building the training dataset seems very efficient and could easily be applicable to Euclid images by using the deep field survey to build a training dataset that could eventually be used for the wide field survey. More experiments will however be required to check more in details the generalization properties.

It would also be interesting to compare all the networks with the shape constraint investigated in [Nammour et al. \(2022\)](#). For Learnlets, as in the case of wavelets, various different kinds of noise could be considered on a single model with an undecimated implementation by adjusting the thresholding function according to the noise. One could enhance the

Unet-based methods by substituting the Unet with a sophisticated denoiser like the Deep Iterative Down-Up CNN (DIDN) (Yu et al., 2019), and adopt a similar strategy for Learnlets by integrating them as elementary units of the DIDN model. On a slightly different note, it would also be interesting to test the concept of unrolled networks (Monga et al., 2021) and neural augmentation (Behrens et al., 2021) in order to mimic iterative deconvolution algorithms with theoretical guarantees. A powerful and accessible deconvolution method would allow for the reconstruction of a cleaner estimation of the sky with less data than classical methods, which is essential for optimising the observing time and the amount of data required to attain a given image reconstruction fidelity (Nammour et al., 2022). On a broader scope, these deconvolution methods could also be applied to other fields.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: <https://archive.stsci.edu/prepds/candels/>.

Author contributions

UA, J-LS, PJ, and FC contributed to the conception and design of the study. UA, PJ, and FC contributed to the dataset generation. PJ and FC gave valuable inputs to envision and describe the astrophysical problem. J-LS gave valuable inputs to

make progress in the technical aspects of the project. UA and KM ran the training scripts. UA produced the results and figures from the experiments. UA wrote the first draft of the manuscript. UA, J-LS, and KM wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Acknowledgments

This work is based on observations taken by the CANDELS Multi-Cycle Treasury Program with the NASA/ESA HST, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., et al. (2015). *TensorFlow: Large-scale machine learning on heterogeneous systems*. Savannah, GA: Software available from tensorflow.org.
- Abrial, P., Moudden, Y., Starck, J.-L., Afeyan, B., Bobin, J., Nguyen, M. K., et al. (2007). Morphological component analysis and inpainting on the sphere: Application in physics and astrophysics. *J. Fourier Anal. Appl.* 13, 729–748. doi:10.1007/s00041-006-6908-x
- Behrens, F., Sauder, J., and Jung, P. (2021). “Neurally augmented {alista},” in International Conference on Learning Representations.
- Bertero, M., and Boccacci, P. (1998). *Introduction to inverse problems in imaging*. CRC Press.
- Chambolle, A., Caselles, V., Cremers, D., Novaga, M., and Pock, T. (2010). An introduction to total variation for image analysis. *Theor. Found. Numer. methods sparse recovery*. Boca Raton, 9, 227.
- Chollet, F. (2017). “Xception: Deep learning with depthwise separable convolutions,” in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, July 21–26, 2017. doi:10.1109/CVPR.2017.195
- Donoho, D. L. (1995). De-noising by soft-thresholding. *IEEE Trans. Inf. Theory*. 41, 613–627. doi:10.1109/18.382009
- Farrens, S., Mboula, F. M. N., and Starck, J.-L. (2017). Space variant deconvolution of galaxy survey images. *Astron. Astrophys.* 601, A66. doi:10.1051/0004-6361/201629709
- Garsden, H., Girard, J. N., Starck, J. L., Corbel, S., Tasse, C., Woiselle, A., et al. (2015). LOFAR sparse image reconstruction. *Astron. Astrophys.* 575, A90. doi:10.1051/0004-6361/201424504
- Gottschling, N. M., Antun, V., Adcock, B., and Hansen, A. C. (2020). *The troublesome kernel: Why deep learning for inverse problems is typically unstable*. arXiv. arXiv preprint arXiv:2001.01258.
- Grogin, N. A., Kocevski, D. D., Faber, S. M., Ferguson, H. C., Koekemoer, A. M., Riess, A. G., et al. (2011). CanDELS: The cosmic assembly near-infrared deep extragalactic legacy survey. *Astrophys. J. Suppl. Ser.* 197, 35. doi:10.1088/0067-0049/197/2/35
- Gurrola-Ramos, J., Dalmau, O., and Alarcón, T. E. (2021). A residual dense u-net neural network for image denoising. *IEEE Access*. 9, 31742–31754. doi:10.1109/ACCESS.2021.3061062
- Huang, G., Liu, Z., Maaten, L. V. D., and Weinberger, K. Q. (2017). “Densely connected convolutional networks,” in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, July 21–26, 2017. doi:10.1109/CVPR.2017.243

- Jin, K. H., McCann, M. T., Froustey, E., and Unser, M. (2017). Deep convolutional neural network for inverse problems in imaging. *IEEE Trans. Image Process.* 26, 4509–4522. doi:10.1109/TIP.2017.2713099
- Kingma, D. P., and Ba, J. (2014). *Adam: A method for stochastic optimization*. San Diego: arXiv.
- Koekemoer, A. M., Faber, S. M., Ferguson, H. C., Grogan, N. A., Kocevski, D. D., Koo, D. C., et al. (2011). Candels: The cosmic assembly near-infrared deep extragalactic legacy survey — The hubble space telescope observations, imaging data products, and mosaics. *Astrophys. J. Suppl. Ser.* 197, 36. doi:10.1088/0067-0049/197/2/36
- Liaudat, T., Starck, J.-L., Kilbinger, M., and Frugier, P.-A. (2022). *Rethinking data-driven point spread function modeling with a differentiable optical model*. doi:10.48550/ARXIV.2203.04908
- Liaudat, T., Starck, J.-L., Kilbinger, M., and Frugier, P.-A. (2021). *Rethinking the modeling of the instrumental response of telescopes with a differentiable optical model*. Virtual conference, Formerly: Vancouver: arXiv. doi:10.48550/ARXIV.2111.12541
- Mohan, S., Kadkhodaie, Z., Simoncelli, E. P., and Fernandez-Granda, C. (2020). “Robust and interpretable blind image denoising via bias-free convolutional neural networks,” in International Conference on Learning Representations, Virtual Conference, Formerly: Addis Ababa ETHIOPIA, April 26–May 1, 2020.
- Monga, V., Li, Y., and Eldar, Y. C. (2021). Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing. *IEEE Signal Process. Mag.* 38, 18–44. doi:10.1109/MSP.2020.3016905
- Nammour, F., Akhaury, U., Girard, J. N., Lanusse, F., Sureau, F., Ali, C. B., et al. (2022). Shapenet: Shape constraint for galaxy image deconvolution. *Astron. Astrophys.* 663, A69. doi:10.1051/0004-6361/202142626
- Ramzi, Z., Michalewicz, K., Starck, J.-L., Moreau, T., and Ciuciu, P. (2021). *Wavelets in the deep learning era*. Amsterdam: HAL. Working paper or preprint.
- Ronneberger, O., Fischer, P., and Brox, T. (2015). *U-net: Convolutional networks for biomedical image segmentation*. Cham: Springer International Publishing. CoRR abs/1505.04597.
- Rudin, L. I., Osher, S., and Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Phys. D. nonlinear Phenom.* 60, 259–268. doi:10.1016/0167-2789(92)90242-f
- Starck, J.-L., Murtagh, F., and Fadili, J. (2015). *Sparse image and signal processing: Wavelets and related geometric multiscale analysis*. Cambridge: Cambridge University Press.
- Sureau, F., Lechat, A., and Starck, J.-L. (2020). Deep learning for a space-variant deconvolution in galaxy surveys. *Astron. Astrophys.* 641, A67. doi:10.1051/0004-6361/201937039
- Tuccillo, D., Huertas-Company, M., Decencière, E., Velasco-Forero, S., Sánchez, H. D., and Dimauro, P. (2018). Deep learning for galaxy surface brightness profile fitting. *Mon. Not. R. Astron. Soc.* 475, 894–909. doi:10.1093/mnras/stx3186
- Xu, L., Ren, J. S., Liu, C., and Jia, J. (2014). “Deep convolutional neural network for image deconvolution,” in *Advances in neural information processing systems*. Long Beach, CA. Editors Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K. Q. Weinberger (Curran Associates, Inc.), 27.
- Yu, S., Park, B., and Jeong, J. (2019). “Deep iterative down-up cnn for image denoising,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Long Beach, CA, USA, June 16–20, 2019