

Open Research Online

The Open University's repository of research publications and other research outputs

Super-resolving *imaging*: a proof of concept using Deep Neural Networks

Journal Item

How to cite:

Lauritsen, Lyngé; Dickinson, Hugh; Bromley, Jane; Serjeant, Stephen; Lim, Chen-Fatt; Gao, Zhen-Kai and Wang, Wei-Hao (2021). Super-resolving Herschel imaging: a proof of concept using Deep Neural Networks. *Monthly Notices of the Royal Astronomical Society*, 507(1) pp. 1546–1556.

For guidance on citations see [FAQs](#).

© 2021 The Authors.



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Version: Accepted Manuscript

Link(s) to article on publisher's website:

<http://dx.doi.org/doi:10.1093/mnras/stab2195>

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's [data policy](#) on reuse of materials please consult the policies page.

oro.open.ac.uk

Super-resolving *Herschel* imaging: a proof of concept using Deep Neural Networks

Lynge Lauritsen^{1*}, Hugh Dickinson¹, Jane Bromley², Stephen Serjeant¹, Chen-Fatt Lim^{3,4},
Zhen-Kai Gao^{4,5}, Wei-Hao Wang⁴

¹*School of Physical Sciences, Faculty of Science, Technology, Engineering & Mathematics, The Open University, Walton Hall, Kents Hill, Milton Keynes, MK7 6AA, United Kingdom*

²*School of Computing & Communications, Faculty of Science, Technology, Engineering & Mathematics, The Open University, Walton Hall, Kents Hill, Milton Keynes, MK7 6AA, United Kingdom*

³*Graduate Institute of Astrophysics, National Taiwan University, Taipei 10617, Taiwan*

⁴*Academia Sinica Institute of Astronomy and Astrophysics (ASIAA), No. 1, Section 4, Roosevelt Road, Taipei 10617, Taiwan*

⁵*Graduate Institute of Astronomy, National Central University, Taoyuan 32001, Taiwan*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

Wide-field sub-millimetre surveys have driven many major advances in galaxy evolution in the past decade, but without extensive follow-up observations the coarse angular resolution of these surveys limits the science exploitation. This has driven the development of various analytical deconvolution methods. In the last half a decade Generative Adversarial Networks have been used to attempt deconvolutions on optical data. Here we present an autoencoder with a novel loss function to overcome this problem in the sub-millimetre wavelength range. This approach is successfully demonstrated on *Herschel* SPIRE 500 μm COSMOS data, with the super-resolving target being the JCMT SCUBA-2 450 μm observations of the same field. We reproduce the JCMT SCUBA-2 images with high fidelity using this autoencoder. This is quantified through the point source fluxes and positions, the completeness and the purity.

Key words: software: data analysis – submillimetre: galaxies – methods: data analysis

1 INTRODUCTION

All astronomical imaging has an intrinsic angular resolution limit, whether due to seeing, diffraction, instrumental effects or (in the case of an interferometer) the longest available baselines. In single-dish diffraction-limited imaging, there is formally no signal on Fourier scales smaller than the diffraction limit. This is because Fraunhofer diffraction is mathematically equivalent¹ to a Fourier transform, so the large-scale boundary of the telescope aperture also implies there is no image information smaller than some angular scale. For interferometers, the Fourier plane has incomplete coverage especially approaching the smallest angular scales.

However, there are often strong scientific drivers for improving angular resolution. Among many advantages, higher angular resolution affords the possibility of more reliable multi-wavelength cross-identifications (e.g. Franco et al. 2018; Dudzevičiūtė et al. 2020), improved deblending of nearby sources (e.g. Hodge et al. 2013; Simpson et al. 2015), and fainter fundamental confusion limits. For example, Geach et al. (2013) used the better angular resolution of

the James Clerk Maxwell Telescope (JCMT) SCUBA-2 450 μm data compared to the *Herschel* SPIRE instrument to probe sub-millimetre (sub-mm) number counts with fluxes below 20 mJy where source confusion becomes problematic in *Herschel* SPIRE data (Oliver et al. 2010; Valiante et al. 2016). Further, Geach et al. (2013) also resolved a larger part of the Cosmic Infrared Background than that possible using *Herschel* SPIRE. There has therefore been a great deal of interest in developing algorithms for recovering or estimating some of the missing Fourier data on smaller angular scales (see Starck et al. (2002) and references therein for a detailed discussion), including approaches that exploit abundant multi-wavelength data where that exists (e.g. Hurley et al. 2017; Jin et al. 2018).

One domain where angular resolution gains are particularly advantageous is sub-mm astronomy. Wide-field extragalactic surveys have proved transformative for e.g. nearby galaxies (e.g. Clark et al. 2018), galaxy evolution (e.g. Lutz 2014; Hayward et al. 2013; Geach et al. 2017) and strong gravitational lensing (e.g. Negrello et al. 2010). Sub-mm galaxies can also be used to trace possible protoclusters through overdensities (Ma et al. 2015; Lewis et al. 2018; Greenslade et al. 2018). Much of this progress has been driven by surveys with the SPIRE instrument (Griffin et al. 2010) on the ESA *Herschel*²

* E-mail: lynge.lauritsen@open.ac.uk

¹ The Fourier transform of the aperture gives the amplitude pattern of a point source, e.g. a 1D top-hat aperture yields a sinc function. Incident energy is proportional to amplitude squared, e.g. a top-hat aperture yields sinc^2 . For a 2D circular aperture this is $\text{sinc}^2(|\mathbf{r}|)$, i.e. an Airy function.

² *Herschel* is an ESA space observatory with science instruments provided by European-led Principal Investigator consortia and with important participation from NASA.

mission (Pilbratt et al. 2010), but at moderately high redshifts (e.g. $z \gtrsim 4$) the detections tend to be dominated by the longest wavelength band ($500 \mu\text{m}$) where the diffraction-limited point spread function (PSF) has a full width half maximum (FWHM) of $36''$. Higher resolution mapping is possible with ground-based facilities such as SCUBA-2 (Holland et al. 2013) and the Atacama Large Millimeter Array (ALMA), but the mapping efficiencies are far lower and it is not feasible to map the entire *Herschel* SPIRE extragalactic survey fields with sub-mm ground-based facilities to comparable depths. Furthermore, the abundant multi-wavelength data available for multi-wavelength prior based deconvolution work in the deeper *Herschel* fields (e.g. Oliver et al. 2012) does not exist at equivalent depths for all wider-area *Herschel* surveys (e.g. Eales et al. 2010).

In the past half-decade the use of machine learning, and in particular Convolutional Neural Networks (CNNs), has gained popularity as a potential solution to image deconvolution (Schawinski et al. 2017; Jia et al. 2021; Moriwaki et al. 2021). These CNNs all use a Generative Adversarial Neural Network (GAN) for their CNN based image restoration. A GAN consists of two neural networks called the generator, and the discriminator. The generator is trained to generate an image that looks “realistic” (according to some relevant quantitative metric), while the discriminator will try to determine if a given image is real or generated (Goodfellow et al. 2014). As they are trained concurrently with competing objectives the performance of the generator will depend on the detailed characteristics of the discriminator. This paper presents an alternative approach using an autoencoder³ with a specially designed loss function. Similar networks have been used to enhance and remove noise from astronomical images at other wavelengths (e.g. Vojtekova et al. 2020).

Architecturally, an autoencoder contains an encoding CNN which extracts a relatively small number of scalar-valued features from input images. The values of these features are referred to collectively as an *embedding* of the input image. A second, decoding CNN is then used to generate an image with the same dimensions as the input images, using only the information encoded by the embedding. The objective of the decoding network is to produce an image that closely matches a given target image that is associated with the corresponding input. During training, the encoding network learns to construct an embedding which optimally represents the features of the input image that are required for the decoding network to generate a close match to the corresponding target (Goodfellow et al. 2016). This paper will show that a simple autoencoder network can be used to super-resolve *Herschel* SPIRE data, and achieve angular resolution comparable to that of JCMT. This super-resolved data can then be used to determine the sky locations and fluxes of previously lower-resolution observations of sub-mm galaxies.

In §2 two different training sets are discussed, in §3 the network architecture, and the loss function is described, in §4 the network performance on both observed, and simulated data is presented, and finally §5, and §6 will discuss and summarise the network performance.

2 TRAINING DATA

The autoencoder presented in this paper is a *supervised* machine learning algorithm. Supervised learning requires the use of a training dataset with known truth values. Two separate training sets were

³ The architecture of an autoencoder is similar to that of many GANs Schawinski et al. (2017); Jia et al. (2021); Moriwaki et al. (2021) but does not use a discriminator network output as part of its objective or loss functions.

used to train the network presented in this paper: (i) a simulated training set, made using images generated by a modified version of the Empirical Galaxy Generator (EGG) software (Schreiber et al. 2017) as both the target and input images, and (ii) using the JCMT SCUBA-2 $450 \mu\text{m}$ maps from the STUDIES project (Wang et al. 2017) as target examples, with the *Herschel* SPIRE maps for the COSMOS field as input images (Levenson et al. 2010; Viero et al. 2013). Table 1 shows the FWHM, confusion limit and pre-interpolation pixel scales of the instruments whose data are used in this paper.

2.1 Simulated Data

There are very few large astronomical fields that have been surveyed by both *Herschel* SPIRE and JCMT SCUBA-2. Accordingly, simulations must be used to create a large representative dataset of images to train the network. The simulated dataset was generated using a version of The Empirical Galaxy Generator (EGG) software (see Schreiber et al. (2017) for an in-depth discussion on the workings of EGG), that was modified to avoid simulating galaxies with negligible infrared fluxes using an empirically determined bolometric luminosity threshold imposed within the code. This was done to improve efficiency and objects with negligible FIR luminosity were not simulated. This modification altered the number counts outside the FIR range, but reproduced realistic number counts in the FIR range at a lower computational cost. We verified that it made no discernible difference to the output images, while saving considerable computation time. EGG is designed to generate a mock survey catalog with realistic multi-wavelength galaxy number counts, using an empirical calibration, and with realistic galaxy clustering. To reduce the number of simulated galaxies, dependent on the depth of the simulated image, the original EGG code uses either a stellar mass cutoff on simulated galaxies, or a UVJ-diagram based selection criteria designed around optical galaxies. The modification in this paper uses the estimated star formation rate (SFR) calculated a bolometric infrared luminosity using the same empirically calibrated formulas already used in the EGG code. EGG uses the SkyMaker (Bertin & Fouqué 2010) code to generate survey images from the mock catalogues. The EGG software was used to create a training data set representing the redshift range $0.1 \leq z \leq 6$. The EGG code generated a number of co-spatial 20 deg^2 images for the 4 bands used in this paper with three *Herschel* SPIRE bands and one JCMT SCUBA-2 band. Each set of 4 images was cut into non-overlapping subregions covering $424 \times 424 \text{ arcsec}^2$, providing a total of 2373 images to train on. 10% of the generated images were reserved for use as a test set.

2.2 Observational Data

A smaller training sample was derived from the small area of overlap within the COSMOS field between the JCMT SCUBA-2 STUDIES large program, and the *Herschel* SPIRE maps. The small area of overlap was divided into 144 overlapping images offset from each other in RA and Dec in steps of 12 arcsec. The 12 arcsec offset steps correspond with the pixel scale of the *Herschel* SPIRE $500 \mu\text{m}$ images and are therefore the smallest possible increment consistent with the lowest resolution images. Each set of images was then flipped and/or rotated to augment the data, to produce 8 images in total at each shifted position. This procedure resulted in an overall dataset containing 1156 images. The 144 non-flipped, non-rotated original images were removed from the training set to be used as a test set, leaving 1008 images for training and ensuring that the images used for testing differed as much as possible from the training set.

Table 1. Characteristics of the *Herschel* SPIRE and JCMT SCUBA-2 instruments. Confusion limits and PSF widths are from [Nguyen et al. \(2010\)](#), [Wang et al. \(2017\)](#) and [Dempsey et al. \(2013\)](#).

Characteristic	Herschel SPIRE			JCMT SCUBA-2
Wavelength	250 μm	350 μm	500 μm	450 μm
PSF FWHM	18.1"	24.9"	36.6"	7.9"
Confusion noise (σ , mJy/beam)	5.8 ± 0.3	6.3 ± 0.4	6.8 ± 0.4	1
Pixel scale	6"	8.33"	12"	1"

3 NETWORK ARCHITECTURE

The generator network of the auto-encoder is based on that used by the GalaxyGAN code ([Schawinski et al. 2017](#)). The CNN presented in this paper differs from GalaxyGAN, and other previous works in two significant ways: (i) the use of a more computationally expensive loss function, that is better designed to extract the individual features of interest, and (ii) no discriminator network is used.

The network processes the two training sets independently, in succession. Each epoch⁴ begins by training on the entire simulated training set, before training on the observed data set three times in succession. The aim was to have the network learn the key structural features of the sub-mm images on the simulated data before using the observed data to fine-tune the network to handle any small differences between observed and simulated images. Each training set was randomised before each run through the data. Due to simulation differences in the flux distribution the observed data were renormalised before training on them to ensure a comparable flux distribution to the simulations.

3.1 Autoencoder

The architecture presented in this paper uses a U-net configuration ([Ronneberger et al. 2015](#)). The outputs from each convolutional layer in the encoder network are concatenated with the inputs of their corresponding layer in the decoder network. This helps to prevent the overall network output from diverging substantially from its input. It takes as its input the three *Herschel* SPIRE bands (250 μm , 350 μm and 500 μm) images and is trained to produce an output image that closely matches a target image consistent with the single JCMT SCUBA-2 450 μm band. All activation functions in the CNN are LeakyReLU:

$$f(x) = \begin{cases} \alpha x, & x < 0 \\ x, & x \geq 0 \end{cases} \quad (1)$$

except for the final layer where a sigmoid function is used:

$$f(x) = \frac{1}{1 + e^{-x}}. \quad (2)$$

LeakyReLU was chosen as the activation functions over the ReLU function, as the zero-gradient nature of the ReLU function at $x < 0$ can cause "dead neurons" in the network. The sigmoid function in the final layer ensures a well constrained output range with continuous coverage. Batch normalisation is included after each convolutional layer to regularise their outputs, which enhances the overall stability

⁴ The term *epoch* refers to a complete pass over the combined observed and simulated training data sets.

of the network and its predictive performance on unseen input data ([Ioffe & Szegedy 2015](#)). For similar reasons, dropout layers are used to randomly disable training 50% of the kernel weights in the first three layers of the decoder network ([Srivastava et al. 2014](#)).

The architecture of the CNN is described in table 2 and a schematic shown in fig. 1. Using this architecture requires that the pixel dimensions of the input and output images match. However, the *Herschel* SPIRE 250 μm , 350 μm , and 500 μm image pixel scales are 6", 8.33", and 12" respectively, while the JCMT SCUBA-2 images have a pixel scale of 1". Accordingly, since the input and output images represent equal areas on the sky, a 2-D linear interpolation routine from the SciPy Python package ([Virtanen et al. 2020](#)) was used to subsample the input *Herschel* images.

3.2 Loss Function

The common approach for deconvolution when designing the loss function for GANs (e.g. [Schawinski et al. 2017](#); [Moriwaki et al. 2021](#)) is to combine the loss L_{disc} from a discriminator network with a simple $L_{1-\text{loss}}$ between the encoder-decoder network output $y_{\text{predicted}}$ and a target image y_{true} .

$$\begin{aligned} L &= L_{1-\text{loss}} + L_{\text{disc}} \\ &= \sum |y_{\text{true}} - y_{\text{predicted}}| + L_{\text{disc}} \end{aligned} \quad (3)$$

The approach in this paper is different. While the CNN presented retains the $L_{1-\text{loss}}$ as part of the overall loss function, it is not the main component. The main goal of the CNN in this paper is to super-resolve the sub-mm telescope PSF. To achieve this, a novel, custom loss function was designed to better target the data features of interest. In particular, this multifaceted loss function focuses on the differences between the fluxes of any point sources that are identified in corresponding pairs of generated and target images.

The loss computation uses the PHOTUTILS Python package ([Bradley et al. 2020](#)) to identify point sources within the target or generated images and extract fluxes from circular apertures with 10 arcsec radius, centred on the identified source locations.

The loss is computed by comparing the fluxes extracted from the generated and target images, but the details of the computation are different when training on the simulated and observed training sets.

3.2.1 Training on simulated data

When the network is training on simulated data, the locations of sources with signal-to-noise ratios of $S/N > 3$ are derived using the target image only. Fluxes are extracted from both the target and generated images using apertures corresponding to the target image locations.

The loss is computed as the sum over all apertures of the absolute

Table 2. Autoencoder network architecture. Layers 1-8 comprise the encoder, while layers 9-16 comprise the decoder. The output of the encoder network is a $2 \times 2 \times 512$ element tensor embedding of the input image, which is used as the input to the decoder network. The outputs from each convolutional layer in the encoder network are concatenated with the inputs of their corresponding layer in the decoder network. All layers use convolutional kernels 4×4 pixel extent in the width and height dimensions. The output of layer 8 encodes the embedding for this auto-encoder network.

Layer	Input dimensions	Kernels	Part 1	Part 2	Part 3	Part 4	Activation	Output dimension
1	$424 \times 424 \times 3$	64	Conv.	BatchNorm.			LeakyReLU	$212 \times 212 \times 64$
2	$212 \times 212 \times 64$	128	Conv.	BatchNorm.			LeakyReLU	$106 \times 106 \times 128$
3	$106 \times 106 \times 128$	256	Conv.	BatchNorm.			LeakyReLU	$53 \times 53 \times 256$
4	$53 \times 53 \times 256$	512	Conv.	BatchNorm.			LeakyReLU	$27 \times 27 \times 512$
5	$27 \times 27 \times 512$	512	Conv.	BatchNorm.			LeakyReLU	$14 \times 14 \times 512$
6	$14 \times 14 \times 512$	512	Conv.	BatchNorm.			LeakyReLU	$7 \times 7 \times 512$
7	$7 \times 7 \times 512$	512	Conv.	BatchNorm.			LeakyReLU	$4 \times 4 \times 512$
8	$4 \times 4 \times 512$	512	Conv.	BatchNorm.			LeakyReLU	$2 \times 2 \times 512$
9	$2 \times 2 \times 512$	512	De-Conv.	BatchNorm.	Dropout	Concat 7	LeakyReLU	$4 \times 4 \times 1024$
10	$4 \times 4 \times 1024$	512	De-Conv.	BatchNorm.	Dropout	Concat 6	LeakyReLU	$7 \times 7 \times 1024$
11	$7 \times 7 \times 1024$	512	De-Conv.	BatchNorm.	Dropout	Concat 5	LeakyReLU	$14 \times 14 \times 1024$
12	$14 \times 14 \times 1024$	512	De-Conv.	BatchNorm.		Concat 4	LeakyReLU	$27 \times 27 \times 1024$
13	$27 \times 27 \times 1024$	256	De-Conv.	BatchNorm.		Concat 3	LeakyReLU	$53 \times 53 \times 512$
14	$53 \times 53 \times 512$	128	De-Conv.	BatchNorm.		Concat 2	LeakyReLU	$106 \times 106 \times 256$
15	$106 \times 106 \times 256$	64	De-Conv.	BatchNorm.		Concat 1	LeakyReLU	$256 \times 256 \times 128$
16	$256 \times 256 \times 128$	1	De-Conv.				Sigmoid	$424 \times 424 \times 1$

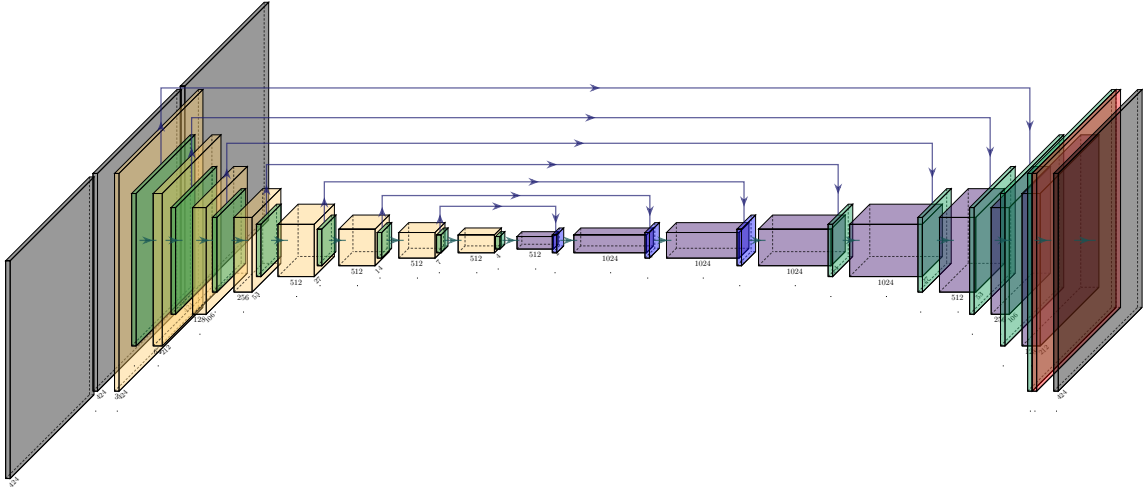


Figure 1. Schematic of the autoencoder used in this work. The yellow boxes represent convolutional layers. The purple boxes represent de-convolutional layers. The green boxes represent combined batch normalisation and LeakyReLU activation function layers. Blue boxes represent a sequence of batch normalisation, dropout and LeakyReLU activation layers. The red box represents the sigmoid activation function in the final layer, and the grey boxes are the input/output images.

difference between the extracted flux in the target and generated image.

$$L_{\text{train}}^{\text{sim}} = \sum_{i=1}^{N_s^{\text{target}}} |f_i^{\text{target}} - f_i^{\text{generated}}| \quad (4)$$

where N_s^{target} is the number of point sources that are identified in the target image, f_i^{target} is the flux extracted from the i th aperture in the target image and $f_i^{\text{generated}}$ is the flux extracted from the i th aperture in the generated image.

3.2.2 Training on observed data

When the network is training on observed data, the locations of source are derived for both the target and generated images. For the target image the source identification criterion remains $S/N > 3$, but for the generated image, this threshold is relaxed and all sources with $S/N > 1$ are identified. The disparity in S/N detection limits used originates in the different purposes of the loss function in the two cases. When detecting sources in the real data the purpose is to replicate the aperture flux in real galaxies. This necessitates that a reasonable lower limit has to be set on sources that are attempted reconstructed. The opposite holds true when detecting sources in the generated data. In this case it is just a important to identify spurious flux anomalies that does not correspond to actual sources, necessitating a lower S/N threshold. Four sets of fluxes are then extracted from both the target and generated images using both sets of apertures. Fluxes are extracted from the target images using apertures from the target and generated images, and vice-versa.

The loss is computed as

$$L_{\text{train}}^{\text{observed}} = \frac{1}{N_s^{\text{target}}} \sum_{i=1}^{N_s^{\text{target}}} |f_i^{\text{target}} - f_i^{\text{generated}}| + \sum_{j=1}^{N_s^{\text{generated}}} \left| \frac{f_j^{\text{target}} - f_j^{\text{generated}}}{f_j^{\text{generated}}} \right| \quad (5)$$

where $N_s^{\text{generated}}$ is the number of point sources that are identified in the generated image. The second term explicitly penalises spurious features that appear in the generated image. The normalisation is applied to account for the fact that most spurious features are faint with $1 < S/N < 3$ and would therefore have insufficient relative influence on the computed loss.

3.2.3 Common loss components

In addition to those based on the aperture flux differences, three common components also contribute to training loss functions for both the observational and simulated training datasets. The first is the reduced mean of the absolute per-pixel difference between the target and generated images⁵.

$$L_{\text{train}}^{L1} = \frac{1}{N_{\text{pix}}} \sum_{i=1}^{N_{\text{pix}}} |f_i^{\text{target}} - f_i^{\text{generated}}| \quad (6)$$

where N_{pix} is the number of pixels in either of the images. This component ensures that the loss includes some influence from the bulk of image pixels outside extracted apertures.

The second common loss component is the absolute difference between the mean pixel fluxes of the generated and target images. This loss component is designed to encourage the generated image to have an integrated flux similar to that of the target image.

Finally, the loss includes the absolute difference between the median pixel fluxes of the generated and target images. This component is intended to produce generated images that have a similar distribution of pixel intensities to the target images. Since the majority of image pixels are noise or background dominated, this tends to result in generated images with similar background properties to the targets.

4 RESULTS

The CNN presented here was trained and tested using both a pure simulation data set and on a data set combining simulated and observed data. In Fig. 2 the performance on a pure simulation data set is demonstrated. The performance of the network on the combined simulated and observed data can be seen in Fig. 3. It is clear that the target images for the simulated data contain more discernible sources than the real JCMT images do. This is likely due to the reduced noise in the simulated data. Figs. 2 and 3 show the *Herschel* bands, as they are provided to the network, post-normalisation. The network has no effective way of recreating the noise inherent in real observations. The median, and L_1 -loss components of the loss function, should drive the network to represent the mean noise level as a quasi-uniform background flux, the spatially varying nature of the real data noise will not be reproduced. The noise in the real background-subtracted JCMT image is distributed around zero. This drives the background level generated by the CNN to be very close to zero, but it can never be negative because the sigmoid activation of the output layer does not allow negative values. The enforced non-negativity of the super-resolved image pixel values also means that the distribution of background noise is highly non-Gaussian and the significance of any point sources in the super-resolved image, relative to the background level cannot be interpreted in a standard Gaussian framework. The right-hand panel of Fig. 4 shows an Eddington-like bias in the reconstructed fluxes at the faint end (Eddington 1913), caused by pre-selecting faint features in the reconstruction. The matching with the JCMT observations uses a lower threshold for features.

Fig. 5 shows the astrometric error on the predicted locations of all of extracted sources detected in the super-resolved image. The natural intuition from single-dish observations is that the positional uncertainty of a point source should be approximately $0.6\theta_{\text{FWHM}}/(S/N)$ where θ_{FWHM} is the beam full-width half maximum and S/N is the signal-to-noise ratio of that source (e.g. Ivison et al. 2007). However, in this case, the reconstructed map is not a single-dish observation, even though it resembles one. The astrometric uncertainty is a non-trivial product of the map reconstruction, and therefore it is something that must be determined directly from the comparison with the truth data. A further complicating factor is that the pixel scales of the originally used *Herschel* SPIRE images are 6", 8.33" and 12", and that of the JCMT images is 1", while the network uses images of the size 424×424 pixels. As 424 is not divisible by either of the *Herschel* SPIRE pixel scales, this will cause minor differences in the exact astrometric alignment of the images fed to the network causing small astrometric errors. Furthermore, the misalignment of the pixels in the *Herschel* SPIRE data, due to the individual pixel scales not being integer multiples of one another might cause additional issues. However, we find that the astrometric accuracy is often better than the pixel scale of the *Herschel* SPIRE 500 μm images, which is the closest equivalent image to the JCMT SCUBA-2 images.

Observed flux is a key characteristic of observed galaxies. A significant portion of the loss function was designed to target the recovery of this observable. Fig. 4 shows the relationship between the super-resolved fluxes and the target fluxes of all sources brighter than 10 times the background flux RMS in the super-resolved image, for all the test image pairs. To compute an absolute flux calibration for the super-resolved point-source fluxes, the JCMT SCUBA-2 450 μm and super-resolved images are both convolved with the *Herschel* 500 μm PSF. A mask is generated to isolate the brightest pixels in the *Herschel* 500 μm image and the pixel fluxes at the unmasked locations are compared with corresponding pixel fluxes in the two convolved images. Two linear scaling relations are found which map the pixel

⁵ This is often referred to as the L1 loss.

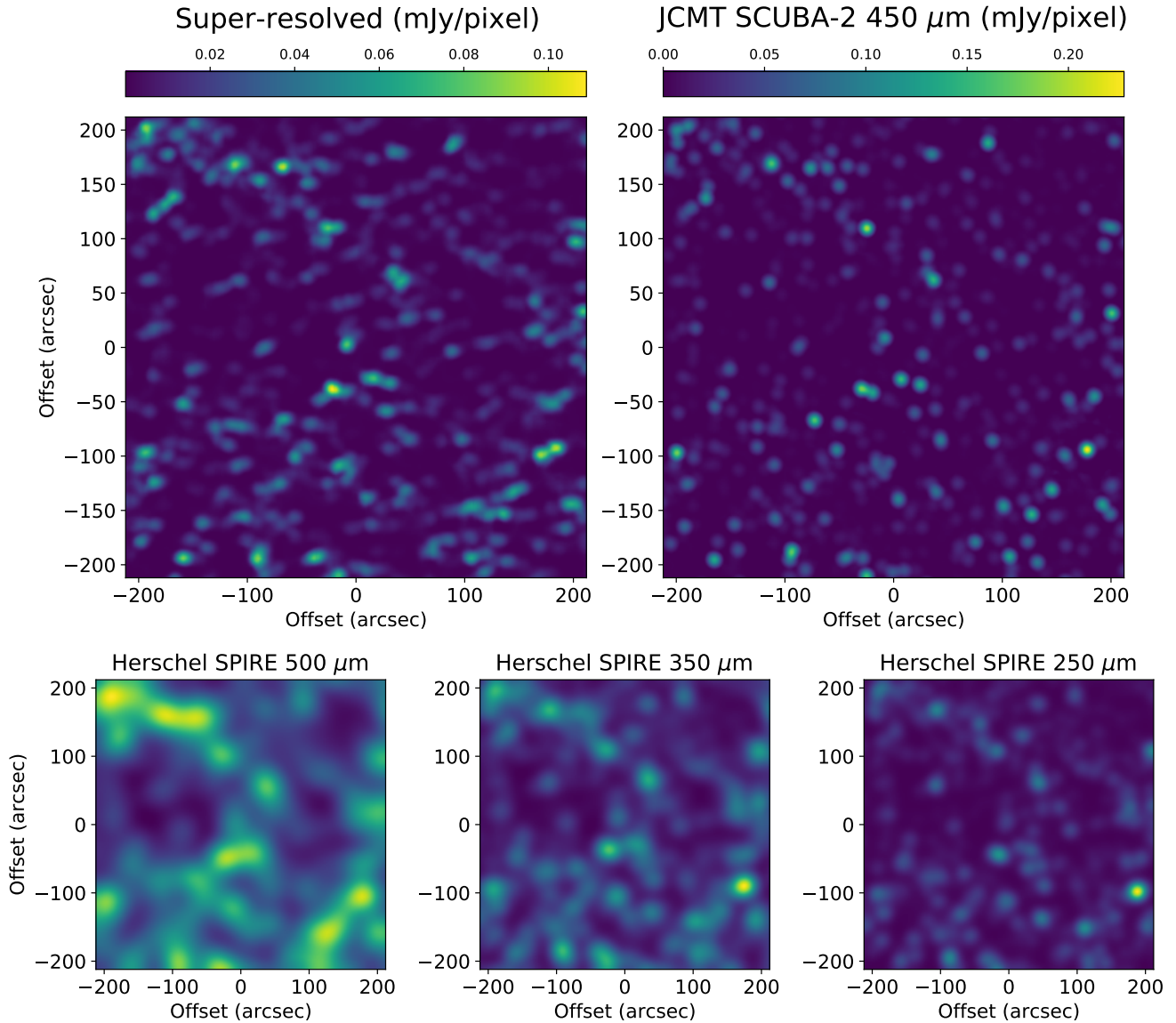


Figure 2. The performance of the CNN on simulated COSMOS data from *Herschel* SPIRE. Top left is the deconvolved image, top right is the actual JCMT SCUBA-2 450 μm image, and the bottom row are the *Herschel* SPIRE images. The *Herschel* SPIRE images shown here are processed identically to the network inputs. They are 2-D linearly interpolated, and linearly normalised to have pixel values between zero and unity. The simulated JCMT 450 μm image demonstrates the depth that originates in the high S/N possible with simulations in the lack of discernible noise.

fluxes in the *Herschel* 500 μm image to those in the convolved JCMT SCUBA-2 450 μm and super-resolved images. Finally, a direct calibration from the super-resolved image flux to the corresponding high-resolution is derived by concatenating these two linear mappings.

While the network does seem to slightly underestimate the calibrated flux for the simulated sources, the results for observational data show good promise. It is worth noting that due to the substantial overlap between the sky areas covered by the individual test images, many of the extracted fluxes correspond to the the same sub-mm galaxy, seen in a different image. Thus the source distribution might not be entirely representative. For bright sources identified in both the simulated and observational datasets, an approximately 1:1 correlation between the calibrated, super-resolved fluxes and their counterparts in the target images is evident, albeit with some scatter. This correlation implies that the fluxes of bright sources can be reli-

ably extracted from super-resolved images. Note that the custom loss function is designed to recover the total flux within a 10" aperture. The pull is defined as $g = |x - \mu|/\sigma$, with x being the expected value, μ being the mean value of the bin, and σ being the standard deviation. The pull has been calculated for the reconstructed source fluxes in table 3, where it is shown that the pull for sources between 9 and 24 mJy with only one exception varies between 0.11 and 0.65. A stacking of the reconstructed sources reveals that the reconstruction has a PSF profile very similar to that of the target data (see fig. 6). This is achieved with only the $L_{1-\text{loss}}$ part of the loss function trying to replicate the PSF shape.

The completeness (also known as recall) and purity (also known as reliability) of the reconstructed sources are shown in fig. 7. Completeness is defined as $TP/(TP + FN)$ where TP is the number of true positives and FN is the number of false negatives; purity is $TP/(TP + FP)$ where FP is the number of false positives. Comple-

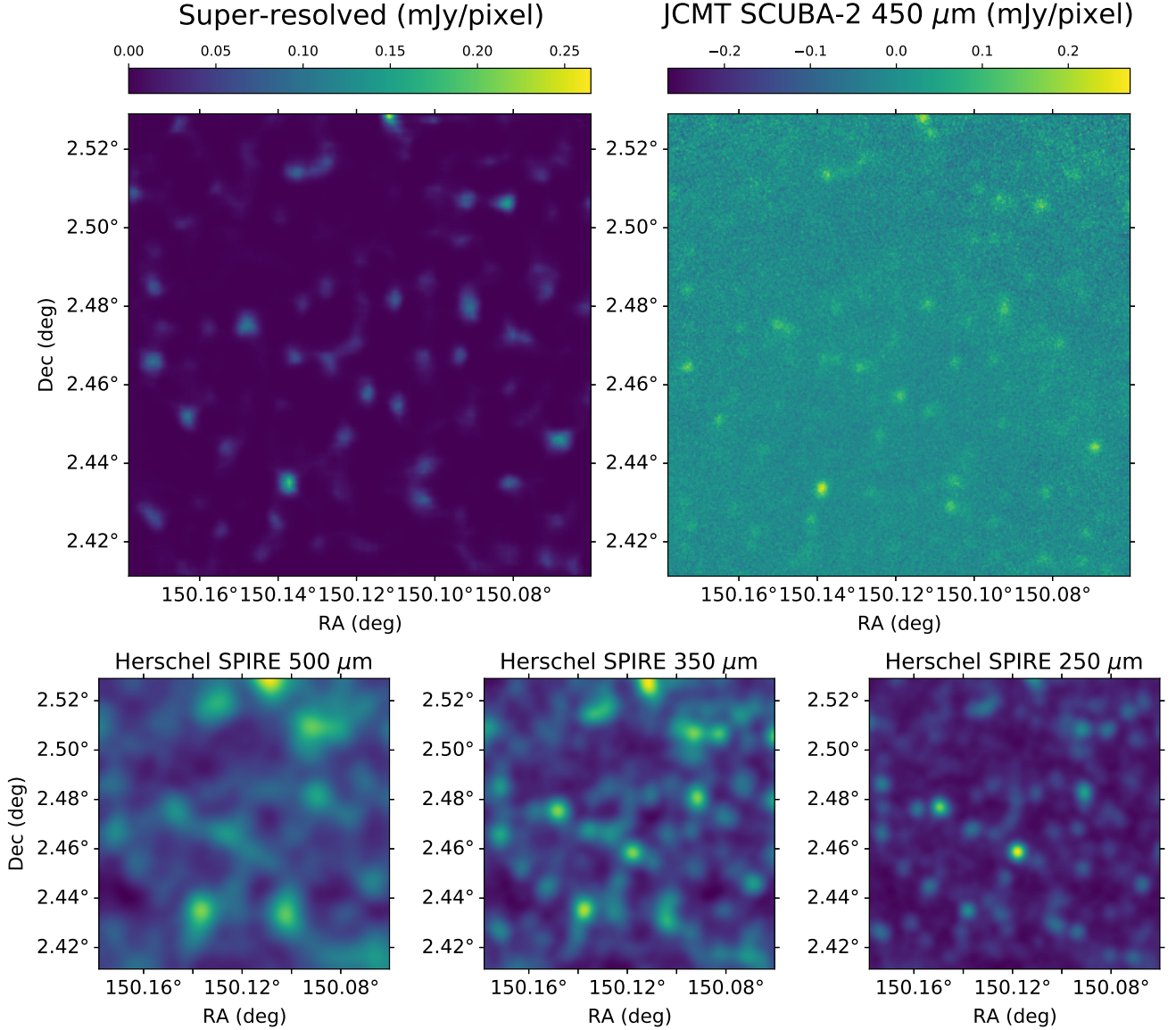


Figure 3. The performance of the CNN on real COSMOS data from *Herschel* SPIRE. Top left is the deconvolved image, top right is the actual JCMT SCUBA-2 450 μm image, and the bottom row are the *Herschel* SPIRE images. The *Herschel* SPIRE images shown here are processed identically to the network inputs. They are 2-D linearly interpolated, and linearly normalised to have pixel values between zero and unity. The JCMT 450 μm image shows the noise inherent in real observations, while the super-resolved image shows the power of an autoencoder in reconstructing the JCMT 450 μm image without the clear noise contribution.

ness is evaluated considering a set of “real” sources with $\text{SNR} \geq 5$ in the JCMT SCUBA-2 450 μm STUDIES survey maps. Sources that are detected in the generated maps are considered to be true positives if they fall within $10''$ of a real source and false negatives otherwise. On the other hand we evaluate purity by considering the set of all “potential” sources that are detected in the generated maps. Potential sources that fall within $10''$ of a real source are deemed to be true positives and all other potential sources are counted as false positives. The completeness is $> 95\%$ at sources brighter than 15 mJy, and above 60% at 10 mJy. The purity does not drop below 87% at any point. Note that our reconstruction is remarkably complete even below the formal 500 μm blank-field confusion limit for *Herschel* SPIRE (table 1).

5 DISCUSSION

While many comparable neural networks (e.g. Schawinski et al. 2017; Jia et al. 2021; Moriwaki et al. 2021) use the output of a discriminator network as part of their loss functions this paper adopts a different approach. Recall that the training objective of a discriminator component in a GAN is to effectively distinguish between images that have been artificially generated or processed and images that are genuine or pristine. However, in order to make this distinction, it may rely on features of the images that a human interpreter might consider unimportant. In this paper, the most important objective from a human perspective is for the neural network to recover the locations and fluxes of the genuine point sources in the target data. However, from the perspective of a CNN it may be that the data sets used in this paper (see Figs. 2 and 3) differ most significantly in their noise

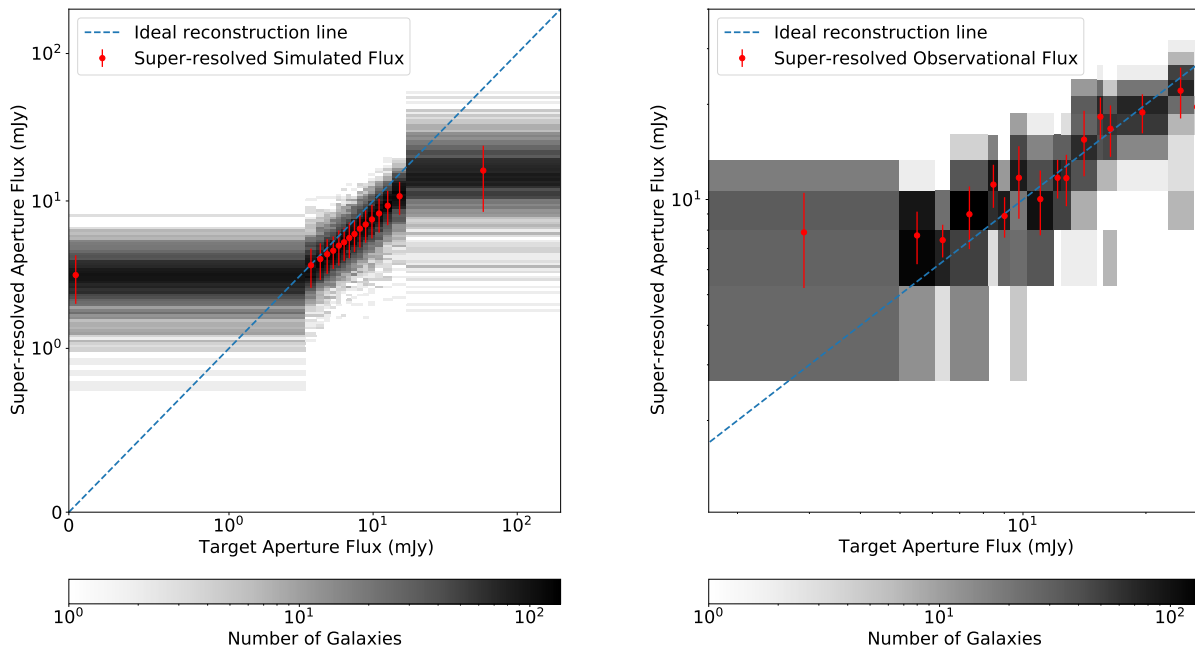


Figure 4. *Left panel:* Comparison between fluxes of point sources that are detected in super-resolved simulated images and fluxes extracted from spatially coincident point sources in simulated high resolution target images. Only sources with fluxes exceeding 10 times the background RMS in the super-resolved image are considered. Note that the distribution of background noise is highly non-Gaussian and the significance of any point sources in the super-resolved image, relative to the background level cannot be interpreted in a standard Gaussian framework. The flux is calculated using aperture photometry within a circular aperture of $10''$ centered on the source locations. The grey histogram shows the number of sub-mm galaxies detected in each bin of target versus super-resolved flux space. The red points and errors show the mean super-resolved aperture flux within each target flux bin and its associated standard deviation, respectively. The bins are defined to ensure equal numbers of galaxies in each bin, which results in the faintest and brightest bins covering a large logarithmic range. The red data points are located at the bin centres in logarithmic flux space. For clarity, the axes on the left panel are linear below 1 mJy and logarithmic above this value. *Right panel:* Same as left-hand panel, but comparing *observed* high resolution JCMT SCUBA-2 $450\ \mu\text{m}$ images with super-resolved *observed* *Herschel* SPIRE counterparts. Note the Eddington bias in the faint fluxes, caused by pre-selection on faint features in the reconstruction (see text).

characteristics. It is therefore possible that a discriminator network would realise its training objective more effectively by focusing its attention on the fine details of the image noise, and disregarding the point source properties like astrometry and flux. Conversely, by using a hand engineered loss function, the network presented in this paper can be forced to focus on the image features that are most critical for the overall objective of super-resolving low resolution images.

Fig. 5 plots the offset distances and position angles between the locations of identifiable point sources in the super-resolved *Herschel* SPIRE data and the nearest sub-mm galaxy location in the corresponding JCMT SCUBA-2 imaging. Overall reconstruction accuracy is excellent, with a purity calculated at above 87% at all reconstructed source flux densities, and completeness above 95% at target source flux densities above 15 mJy. Nonetheless, some small offsets between the reconstructed and target source positions are apparent. These offsets are likely caused by the different pixel scales for the different *Herschel* SPIRE bands, and the JCMT data. These pixel scales are not exact multiples of each other and so pixels from the different image bands intercept flux from different parts of the sky, and may encode information about different subsets of the true source distribution. Even after interpolation, the sources which fall close to the edge of a pixel in the lower resolution bands have inherently uncertain positions, which is likely reflected in the CNN output. Further, the uncertain alignment of in particular the $350\ \mu\text{m}$ band might cause problems. The $12''$ and $6''$ of the $500\ \mu\text{m}$ and $250\ \mu\text{m}$

bands divides into each other, while the $8.33''$ of the $350\ \mu\text{m}$ band might cause some uncertainty in source location when the images are shifted during data augmentation. Finally, the redder sources might have higher astrometric uncertainty as they are less represented in the higher resolution *Herschel* bands. While further work might reduce this astrometric offset, Fig. 5 shows a tendency of astrometric precision better than the $12''$ pixel scale of the $500\ \mu\text{m}$ *Herschel* SPIRE band.

Following this successful proof of concept, there are several obvious next steps. These go beyond this initial analysis, and at least some of these will be presented in future papers.

Firstly, this deconvolution algorithm will be applied to all the *Herschel* SPIRE extragalactic survey data sets. For deeper fields with richer multi-wavelength complementary data, the deconvolution can be compared to other approaches that use this supplementary data as a prior (e.g. Hurley et al. 2017; Serjeant 2019).

Secondly, there are enhancements that can be made to the simulations, such as incorporating Galactic cirrus. Furthermore, Dunne et al. (2020) find that foreground large-scale structure can statistically magnify the background sub-mm source counts, so one improvement to the simulations would be to incorporate optical/near-infrared imaging and the effects of weak lensing. The deconvolutions would then be able to make use of the three SPIRE bands and the optical/near-infrared data. In the present analysis, the statistical clustering properties of sub-mm galaxies are implicitly (and non-trivially) used to

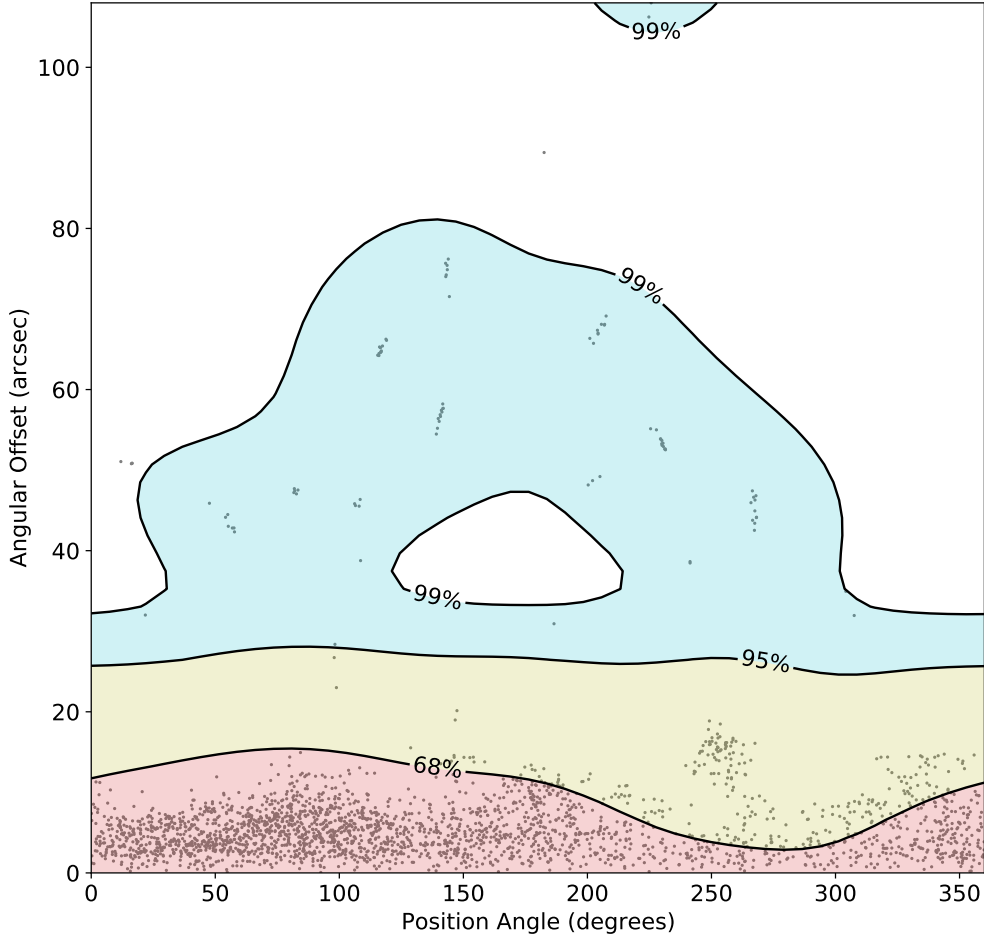


Figure 5. Angular distance versus position angle for offsets of extracted sources identified on the super-resolved images and the nearest counterpart point source in coincident SCUBA-2 imaging. Position angles are defined in degrees anti-clockwise from West. Grey markers correspond with individual objects that are identified in the super-resolved images. Contours showing the 68th, 95th and 99th percentiles are overlaid. Overall, the astrometric accuracy is excellent with 99% of the super-resolved objects having offsets ≤ 15 arcseconds. There is some evidence for clustering of offsets to the Northwest.

reconstruct the missing Fourier modes on scales smaller than the point spread function (section 1), so simulating a wider range of clustered multi-wavelength training data should improve the deconvolution. Strong gravitational lensing could also be included (e.g. Negrello et al. 2010), in which case the network could also encode multi-wavelength information, such as the presence of a foreground elliptical or cluster to signpost possible strong lensing. Extending the simulations and neural net training to a wider multi-wavelength domain has the potential in principle to implicitly incorporate more information than explicit multi-wavelength priors, albeit at a cost of less direct interpretability.

Thirdly, the loss function can be tailored to suit particular science goals. The present analysis represents a particular balance between source completeness, source reliability, flux reproducibility and astrometric accuracy, but other choices are possible. There is no reason to suppose that a single "best" deconvolution to suit all purposes

is possible even in principle. Indeed, the balances between angular resolution, point-source sensitivity and large-scale features are usually explicit and deliberate choices in astronomical image processing, driven in each case by the particular science goals (e.g. Briggs 1995; Serjeant et al. 2003; Smith et al. 2019; Danieli et al. 2020). One could imagine optimising the loss function not just for completeness or reliability or some balance thereof, but instead to reproduce the sub-mm galaxy source counts, or make the best estimate of the two-point correlation function of sub-mm galaxies, or reliably detect faint ultra-red sub-mm galaxies.

6 CONCLUSIONS

This paper has shown that it is possible to super-resolve *Herschel* SPIRE data using CNNs. In this paper an autoencoder was chosen. A

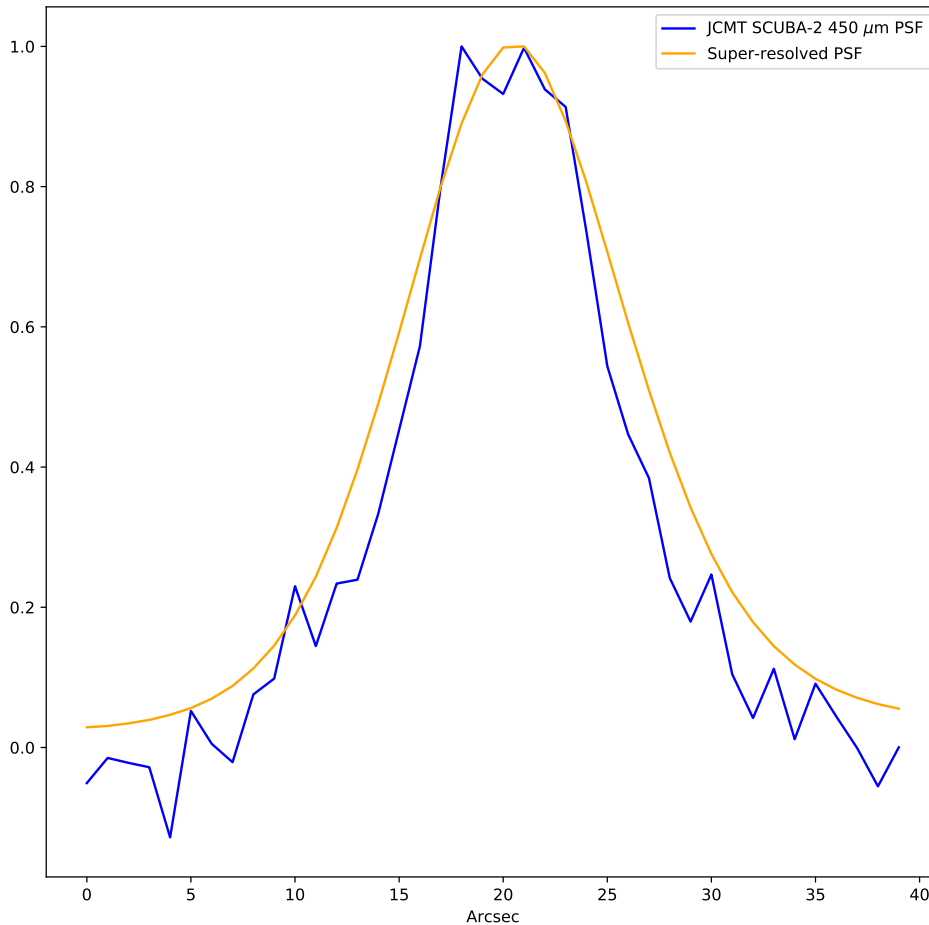


Figure 6. The stacked PSF of the real JCMT SCUBA-2 450 μm sources (blue), and the super-resolved sources (orange). The fluctuations in the real sources originate from the small sample of real images that can be used for training and validation. This causes a large overlap of sky area repeating the same source multiple times. The smoothness of the super-resolved PSF is achieved by the final activation layer suppressing the noise in the generated images.

new and innovative loss function was engineered to better replicate the image features of interest.

It is possible to reconstruct both astrometry and source flux using this method with some uncertainty. It is expected that the performance on particularly the source flux would improve with a larger, more varied training set of observed data, reducing the need for simulated data in the training phase. More realistic simulated data might also achieve this goal.

Ultimately this method will allow for further exploration of the fields observed by *Herschel* SPIRE as a complement to the observations carried out with JCMT and similar telescopes.

ACKNOWLEDGEMENTS

We thank the anonymous referee for many thoughtful and helpful comments that improved this paper. The sub-mm observations

used in this work include the STUDIES program (program code M16AL006), archival data from the S2CLS program (program code MJLSC01) and the PI program of Casey et al. (2013, program codes M11BH11A, M12AH11A, and M12BH21A). The James Clerk Maxwell Telescope is operated by the East Asian Observatory on behalf of The National Astronomical Observatory of Japan; Academia Sinica Institute of Astronomy and Astrophysics; the Korea Astronomy and Space Science Institute; Center for Astronomical Mega-Science (as well as the National Key R&D Program of China with No. 2017YFA0402700). Additional funding support is provided by the Science and Technology Facilities Council of the United Kingdom and participating universities and organizations in the United Kingdom and Canada. Additional funds for the construction of SCUBA-2 were provided by the Canada Foundation for Innovation. The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the indigenous Hawaiian community. We are

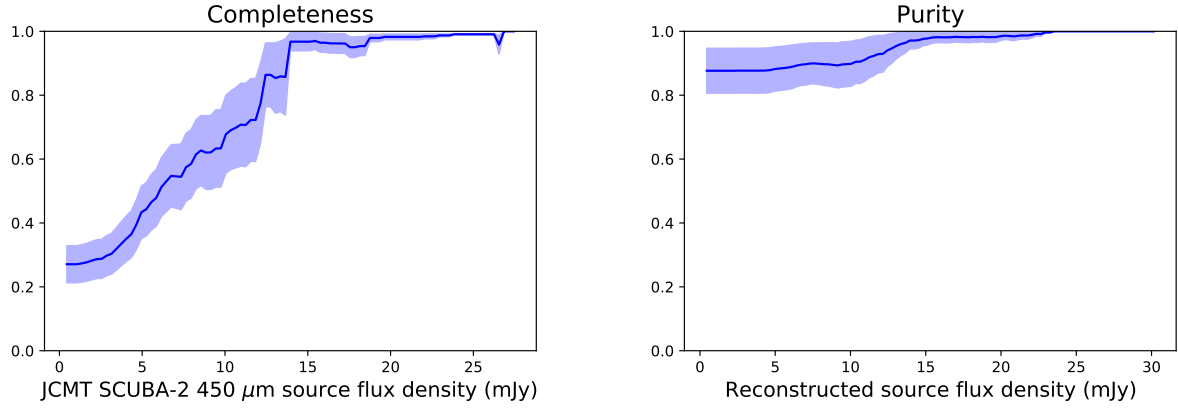


Figure 7. *Left panel:* Completeness of recreated sources inside $10''$ of detected $\text{SNR} \geq 5$ sources in the JCMT SCUBA-2 $450\ \mu\text{m}$ STUDIES survey maps. *Right panel:* Purity of sources detected in the reconstructed image as compared to sources with $\text{SNR} \geq 5$ in the JCMT SCUBA-2 $450\ \mu\text{m}$ STUDIES survey. In both cases, the horizontal axis shows minimum the flux density threshold under consideration. The shaded areas show the extent of the $\pm 1\sigma$ binomial uncertainties.

Table 3. The pull calculated for the reconstructed sources shown in the righthand panel of fig. 4.

Flux (mJy)	Pull
2.91	1.89
5.5	1.51
6.36	1.22
7.39	0.79
8.47	1.55
9.01	0.11
9.77	0.64
11.03	0.43
12.15	0.27
12.77	0.50
14.11	0.37
15.47	1.02
16.38	0.12
19.61	0.55
24.32	0.55
26.67	3.99

most fortunate to have the opportunity to conduct observations from this mountain.

This research has made use of data from HerMES project (<http://hermes.sussex.ac.uk/>). HerMES is a Herschel Key Programme utilising Guaranteed Time from the SPIRE instrument team, ESAC scientists and a mission scientist. The HerMES data was accessed through the Herschel Database in Marseille (HeDaM - <http://hedam.lam.fr>) operated by CeSAM and hosted by the Laboratoire d’Astrophysique de Marseille. The OBSIDs of the *Herschel* fields used were: 1342195856, 1342195857, 1342195858, 1342195859, 1342195860, 1342195861, 1342195862, 1342195863, 1342222819, 1342222820, 1342222821, 1342222822, 1342222823,

1342222824, 1342222825, 1342222826, 1342222846, 1342222847, 1342222848, 1342222849, 1342222850, 1342222851, 1342222852, 1342222853, 1342222854, 1342222879, 1342222880, 1342222897, 1342222898, 1342222899, 1342222900 and 1342222901.

This research made use of Astropy,⁶ a community-developed core Python package for Astronomy (Astropy Collaboration et al. 2013, 2018). This research made use of Photutils, an Astropy package for detection and photometry of astronomical sources (Bradley et al. 2020). Data analysis made use of the Python packages Numpy (Harris et al. 2020), Scipy (Virtanen et al. 2020) and Pandas (The pandas development team 2020; Wes McKinney 2010) as well as Tensorflow (Abadi et al. 2015) in Keras (Chollet 2015). Figures were made with the Python package Matplotlib (Hunter 2007).

SS and HD were supported in part by ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures, which in turn received funding from the European Union’s Horizon 2020 research and innovation programme under Grant Agreement no. 824064. SS and LL thank the Science and Technology Facilities Council for support under grants ST/P000584/1 and ST/T506321/1 respectively.

DATA AVAILABILITY

The observational data used for this paper are the *Herschel* SPIRE COSMOS data, the maps for which can be downloaded at <https://irsa.ipac.caltech.edu/data/COSMOS/images/herschel/spire/>, and the JCMT SCUBA-2 STUDIES data.

REFERENCES

- Abadi M., et al., 2015, TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, <http://tensorflow.org/>
 Astropy Collaboration et al., 2013, *A&A*, **558**, A33
 Astropy Collaboration et al., 2018, *AJ*, **156**, 123
 Bertin E., Fouqué P., 2010, SkyMaker: Astronomical Image Simulations Made Easy (ascl:1010.066)

⁶ <http://www.astropy.org>

- Bradley L., et al., 2020, *astropy/photutils*: 1.0.0, doi:10.5281/zenodo.4044744, <https://doi.org/10.5281/zenodo.4044744>
- Briggs D. S., 1995, in *American Astronomical Society Meeting Abstracts*, p. 112.02
- Casey C. M., et al., 2013, *MNRAS*, 436, 1919
- Chollet F., 2015, *Keras*, <https://github.com/fchollet/keras>
- Clark C. J. R., et al., 2018, *A&A*, 609, A37
- Danieli S., et al., 2020, *ApJ*, 894, 119
- Dempsey J. T., et al., 2013, *MNRAS*, 430, 2534
- Dudzevičiūtė U., et al., 2020, *MNRAS*, 494, 3828
- Dunne L., Bonavera L., Gonzalez-Nuevo J., Maddox S. J., Vlahakis C., 2020, *MNRAS*, 498, 4635
- Eales S., et al., 2010, *PASP*, 122, 499
- Eddington A. S., 1913, *MNRAS*, 73, 359
- Franco M., et al., 2018, *A&A*, 620, A152
- Geach J. E., et al., 2013, *Monthly Notices of the Royal Astronomical Society*, 432, 53
- Geach J. E., et al., 2017, *MNRAS*, 465, 1789
- Goodfellow I. J., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y., 2014, *Generative Adversarial Networks* ([arXiv:1406.2661](https://arxiv.org/abs/1406.2661))
- Goodfellow I., Bengio Y., Courville A., 2016, *Deep Learning*. MIT Press
- Greenslade J., et al., 2018, *Monthly Notices of the Royal Astronomical Society*, 476, 3336
- Griffin M. J., et al., 2010, *A&A*, 518, L3
- Harris C. R., et al., 2020, *Nature*, 585, 357
- Hayward C. C., Narayanan D., Kereš D., Jonsson P., Hopkins P. F., Cox T. J., Hernquist L., 2013, *MNRAS*, 428, 2529–2547
- Hodge J. A., et al., 2013, *ApJ*, 768, 91
- Holland W. S., et al., 2013, *MNRAS*, 430, 2513
- Hunter J. D., 2007, *Computing in Science & Engineering*, 9, 90
- Hurley P. D., et al., 2017, *MNRAS*, 464, 885
- Ioffe S., Szegedy C., 2015, *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift* ([arXiv:1502.03167](https://arxiv.org/abs/1502.03167))
- Iverson R. J., et al., 2007, *MNRAS*, 380, 199
- Jia P., Ning R., Sun R., Yang X., Cai D., 2021, *MNRAS*, 501, 291
- Jin S., et al., 2018, *ApJ*, 864, 56
- Levenson L., et al., 2010, *MNRAS*, 409, 83
- Lewis A. J. R., et al., 2018, *The Astrophysical Journal*, 862, 96
- Lutz D., 2014, *ARA&A*, 52, 373
- Ma C.-J., et al., 2015, *The Astrophysical Journal*, 806, 257
- Moriwaki K., Shirasaki M., Yoshida N., 2021, *The Astrophysical Journal*, 906, L1
- Negrello M., et al., 2010, *Science*, 330, 800
- Nguyen H. T., et al., 2010, *Astronomy and Astrophysics*, 518, L5
- Oliver S. J., et al., 2010, *A&A*, 518, L21
- Oliver S. J., et al., 2012, *MNRAS*, 424, 1614
- Pilbratt G. L., et al., 2010, *A&A*, 518, L1
- Ronneberger O., Fischer P., Brox T., 2015, in Navab N., Hornegger J., Wells W. M., Frangi A. F., eds, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Springer International Publishing, Cham, pp 234–241
- Schawinski K., Zhang C., Zhang H., Fowler L., Santhanam G. K., 2017, *Monthly Notices of the Royal Astronomical Society: Letters*, 467, L110
- Schreiber C., et al., 2017, *A&A*, 602, A96
- Serjeant S., 2019, *Research Notes of the American Astronomical Society*, 3, 133
- Serjeant S., et al., 2003, *MNRAS*, 344, 887
- Simpson J. M., et al., 2015, *ApJ*, 799, 81
- Smith M. W. L., et al., 2019, *MNRAS*, 486, 4166
- Srivastava N., Hinton G., Krizhevsky A., Sutskever I., Salakhutdinov R., 2014, *Journal of Machine Learning Research*, 15, 1929
- Starck J., Pantin E., Murtagh F., 2002, *Publications of the Astronomical Society of the Pacific*, 114, 1051
- The pandas development team 2020, *pandas-dev/pandas*: Pandas, doi:10.5281/zenodo.3509134, <https://doi.org/10.5281/zenodo.3509134>
- Valiante E., et al., 2016, *Monthly Notices of the Royal Astronomical Society*, 462, 3146
- Viero M. P., et al., 2013, *ApJ*, 772, 77
- Virtanen P., et al., 2020, *Nature Methods*, 17, 261
- Vojtekova A., Lieu M., Valtchanov I., Altieri B., Old L., Chen Q., Hroch F., 2020, *MNRAS*,
- Wang W.-H., et al., 2017, *ApJ*, 850, 37
- Wes McKinney 2010, in Stéfan van der Walt Jarrod Millman eds, *Proceedings of the 9th Python in Science Conference*. pp 56 – 61, doi:10.25080/Majora-92bf1922-00a

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.