RASTI **000**, 1–17 (2023) Preprint 29 February 2024 Compiled using rasti LATEX style file v3.0

ContinUNet: Fast Deep Radio Image Segmentation in the SKA Era with U-Net

Hattie Stewart,[1,2,3]★ Mark Birkinshaw,[1] Siu-Lun Yeung,[2] Natasha Maddox,[1] Ben Maughan,[1] Jeyan Thiyagalingam[2]
[1]*School of Physics, University of Bristol, HH Wills Physics Laboratory, Tyndall Avenue, Bristol BS8 1TL, UK*
[2]*SciML, Scientific Computing Department, Research Complex at Harwell, Rutherford Appleton Laboratory, Harwell Oxford, Didcot Didcot OX11 0FA*
[3]*UKRI Centre for Doctoral Training in Artificial Intelligence, Machine Learning & Advanced Computing*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

We present a new machine learning driven (ML) source-finding tool for next generation radio surveys that performs fast source extraction on a range of source morphologies at large dynamic ranges with minimal parameter tuning and post processing. The field of radio astronomy is on the brink of groundbreaking scientific advances with the construction of the Square Kilometre Array (*SKA*) radio telescope. Reaching science goals with current radio telescopes, which have modest data products compared to those promised by the *SKA*, is inhibited by a lack of accurate and automated source-finding techniques. We have developed a novel source-finding method, ContinUNet, powered by an ML segmentation algorithm, U-Net, that has proven highly effective and efficient when tested on *SKA* precursor data sets. Our model was trained and tested on simulated radio continuum data from *SKA* Science Data Challenge 1 and proved comparable to the state-of-the-art source-finding methods, PyBDSF and ProFound, in terms of recovery of the source population and their characteristics. ContinUNet was then tested on the MIGHTEE Early Science data without retraining and was able to extract point sources and extended sources with equal ease; processing a 1.6 $\deg^2$ field in <13 s on a supercomputer and ≈2 min on a personal laptop. We were able to associate components of extended sources without manual intervention due to the powerful inference capabilities learnt within the network. These advances make ContinUNet a promising tool for enabling science in the upcoming *SKA* era.

Key words: Machine Learning – Software – Data methods – radio continuum: galaxies – galaxies: evolution

1 INTRODUCTION

For many years radio telescopes have been used study Galactic and extragalactic processes through the thermal and non-thermal emission that they produce. With modern instruments like the Square Kilometre Array (*SKA*) pathfinder Low-Frequency Array (*LOFAR*; van Haarlem et al. 2013) and precursor *MeerKat* (Jonas et al. 2018), we have been able to study the Universe in greater depth than ever before. However, the data obtained from these surveys are becoming increasingly difficult to analyse due to their large size, making conventional source-finding techniques impractical. These conventional techniques may also make incorrect assumptions about the properties of the sources, potentially impacting the accuracy of statistical estimations of cosmological and astrophysical quantities.

Radio emission is a primary tracer of the star formation history in galaxies (e.g. Jarvis et al. 2014) via thermal emission in ionized hydrogen (HII) regions and non-thermal synchrotron radiation from supernova remnants. Deep radio surveys can detect radio emission from very distant star forming galaxies (SFGs) at different frequencies to track how they are evolving over time (e.g. Smith et al. 2021). Studying radio properties of these galaxies helps us understand the underlying mechanisms driving star formation throughout cosmic history. Active galactic nuclei (AGN) also produce radio emission, both from accretion onto the core and the generation of jets. Deep radio surveys can trace AGN activity to even higher redshifts, allowing us to understand how the source populations evolve with cosmic time (e.g. Smolcic et al. 2014; Whittam et al. 2023). Observations at multiple frequencies additionally provide information on the underlying emission mechanisms and how these sources age (Harwood et al. 2017).

In order to leverage the true science potential of modern radio surveys, sources must be extracted from the image data products and their parameters measured. The current state-of-the-art techniques used for source extraction in radio astronomy are PyBDSF (Mohan et al. 2015) and ProFound (Robotham et al. 2018). PyBDSF identifies sources in radio interferometric images by fitting a 2D Gaussian model to each component of emission. The algorithm involves a series of steps including background subtraction, island identification, Gaussian fitting, and deblending of overlapping sources. PyBDSF has been applied frequently across radio astronomical applications, for example for source-finding in the *LOFAR* Two-metre Sky Survey (Shimwell et al. 2017). ProFound detects sources in astronomical images by identifying objects as groups of contiguous pixels above a certain threshold. ProFound was initially developed for optical as-

★ E-mail: publishing@ras.org.uk (KTS)

tronomical data and used for measuring photometry in the Deep Extragalactic Visible Legacy Survey (DEVILS) data (Davies et al. 2018). ProFound was subsequently applied to radio data, Hale et al. (2019) performed source-finding and parameter extraction with ProFound on observations of the XMM-LSS field taken with the Very Large Array (*VLA*). Both methods provide tools for visual inspection and verification of the detected sources.

As the problem of analyzing large datasets will become even more significant in the *SKA* era, the *SKA* Observatory (SKAO) has initiated Science Data Challenges (SDC) to test new approaches that can handle the massive amount of *SKA* data whilst dealing with realistic image defects. Considering the large volumes of image data, it is an appropriate application for deep learning image segmentation methods. There have already been some applications of ML to the problem of source extraction in radio data in particular. ConvoSource (Lukic et al. 2019) uses a Convolutional Neural Network (CNN) to perform source extraction in the Science Data Challenge 1 (SDC1) data. Similarly, DeepSource (Sadr et al. 2018) uses a CNN to perform point source object detection. Both ConvoSource and DeepSource are able to extract point sources from radio data, but it is not evident that they are capable of extracting extended sources from radio data. Sortino et al. (2023) performed benchmarking on deep learning methods for object detection and segmentation in a recent review, which includes various deep learning segmentation methods such as U-Net. This review presents a brief analysis of a comprehensive list of different methods, but does not perform any extensive analysis or application of these techniques. More recently Riggi et al. (2023) published a source extraction tool, ceasar-mrcnn, a framework built around a Mask R-CNN trained using Evolutionary Map of the Universe (EMU; Norris et al. 2011) survey data from the Australian Square Kilometre Array Pathfinder (*ASKAP*; Hotan et al. 2021) radio telescope, which showed promising results.

It is widely accepted that thousands of annotated training samples are required for successful training of deep neural-networks (Ronneberger et al. 2015). However, U-Net is a deep method for image segmentation that can be trained on very few labelled images. U-Net is a segmentation algorithm taken from the field of biomedical imaging and was originally developed for segmenting electron microscopy data. U-Net is based on an encoder-decoder architecture, much like the Auto Encoder (AE) defined by Goodfellow et al. (2016) and Variational Autoencoder (VAE) devoloped by Kingma & Welling (2013), which have proven very effective for feature extraction in image data. The contracting path (encoder) captures context or features within the image and a symmetric expanding path (decoder) enables precise localisation. In Wang et al. (2023), a Vector Quantised VAE based method was applied to 3D magnetic resonance imaging (MRI) data images to detect anomalies in brain scans with results that outperform state-of-the-art methods. Yasutomi & Tanaka (2023) presented a Contrastive Conditioned VAE for efficiently extracting style features from text images in order to classify fonts. Meissen et al. (2022) used a feature-mapping function as a pre-proccesing step for unsupervised anomaly detection with an AE in brain MRI data. It is quickly evident that encoder-decoder architectures can be applied to a variety of image segmentation and classification cases with extremely promising results. U-Net has already been used for a variety of astrophysical research cases with proven success. For example, Zhou et al. (2022) applied ResUNet to extract foreground contamination in carbon monoxide intensity maps. A variation on U-Net, SegU-Net, presented in Bianco et al. (2023) was used to identify neutral and ionized regions in simulated 21-cm signal data. In Gupta & Reichardt (2020) the mResUNet was applied to simulated cosmic microwave background signals to extract the Sunyaev-Zel'dovich profiles and galaxy cluster masses. Makinen et al. (2020) perform foreground removal in 21-cm intensity mapping observations using U-Net with principle component analysis (PCA) as a pre-processing step.

Considering the urgency for faster and more efficient source extraction methods for the extremely large datasets produced by modern radio surveys, we present the ContinUNet framework for source-finding in 2D radio continuum data, based on the U-Net architecture (Ronneberger et al. 2015). We present ContinUNet in Section 2, an end-to-end pipeline for training a U-Net model on simulated radio continuum data and predicting source catalogues of new input data, describing the different modules developed for source-finding. We describe the SDC1 data set (Bonaldi & Braun 2018) used for training in Section 3, how we produce our segmentation maps for training from the provided truth catalogues and the training process itself. We perform an exhaustive comparison of our model with state-of-the-art source finders PyBDSF and ProFound in Section 4 and compare all methods to the ground truth catalogue provided with the SDC1 dataset. We extend our framework to real radio continuum data in Section 5 and apply ContinUNet to Early Science radio continuum data from the *MeerKat* International GHz Tiered Extragalactic Exploration (MIGHTEE) Survey (Jarvis et al. 2016). We are able to perform fast source-finding, extracting source counts comparable with those produced by the state-of-the-art source extraction methods at fast speeds. We also show examples of ContinUNet's ability to associate source components with no retraining, prior tuning or post source-finding clean up. We have developed a new ML source extraction method, benchmarked against state-of-the-art tools currently used by radio astronomers for source-finding, capable of detecting objects in large radio images with no parameter tuning and no tiling of the input data.

2 METHODS

Image segmentation is the process of partitioning an image into areas of interest, where a pixel-wise mask is created for each object in the image. This provides much more information than a simple bounding box and therefore is more suitable for this task. Semantic segmentation is a computer vision task that assigns a class label to each pixel. Binary segmentation methods only identify one type of object in an image. A segmentation map is an image that has every pixel labelled by a class, that represents the features within a corresponding data image. The segmentation maps for binary segmentation will typically have pixel values of 1 for the object identified (positive class), and 0 for the background. In order to perform image segmentation with ML, we require a labelled training set, consisting of data images and corresponding segmentation maps.

2.1 ContinUNet Model Framework

We propose a new framework ContinUNet for end-to-end source extraction from radio continuum images using U-Net architecture to learn the intrinsic source segmentation. The ContinUNet framework is split into modules; pre-processing, learning and inference. Raw data is parsed into the pre-processing module to generate a training data set. This training set is used to train the U-Net architecture, whose weights are saved into a model for the inference module. Pre-processed data cutouts can be parsed to the inference module to generate predicted source catalogues for the raw data input. This architecture is depicted in Fig. 1.

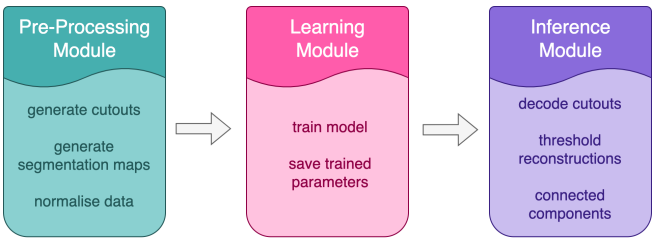

Figure 1. Graphic depicting the modular framework of ContinUNet. Pre-processing, learning and inference modules are described here.

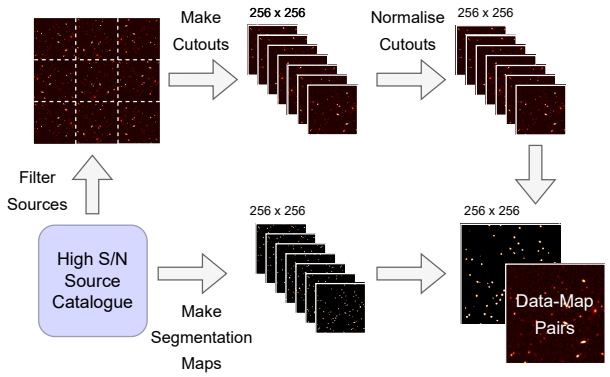

Figure 2. Graphic depicting the pre-processing module.

2.1.1 Pre-processing Module

The pre-processing module, Fig. 2 is used to convert our raw data images to a training set of data cutouts and corresponding segmentation maps. To create the segmentation maps for our radio data cutouts we can use source properties provided in the truth catalogues: Right Ascension (RA), Declination (Dec), position angle and major and minor axes. Using the location and properties of known sources we can create masked objects on an empty array to act as image labels. We do this by inserting elliptical masks for point sources and a more complex segmentation for extended sources. To create the segmentation for extended sources, we find their edges using a threshold on a cutout of the source from the data image and produce a segmentation 'stamp'. The class of each source as provided by the data set used are included in the segmentation map. However, only binary segmentation maps are currently produced by the model, where 0 represents background and 1 represents a source.

The data cutout values are normalised linearly to $0 <$ pixel value < 1, using:

$$I_{\text{normalised}} = \frac{I_{\text{original}} - I_{\text{min}}}{I_{\text{max}} - I_{\text{min}}} \tag{1}$$

where I is the pixel intensity value. Normalisation is standard practice in ML applications and is performed to ensure the training data has a consistent scale for the network to work with, leading to faster convergence during training. The generated cutouts and segmentation maps are hereafter referred to as data-map pairs.

2.1.2 Learning Module

The learning module uses a U-Net architecture, as shown in Fig. 3, to train a model for predicting segmentation maps from input data. The encoder has a similar architecture to a typical CNN; down-sampling is achieved through a series of convolutional layers, max pooling and dropout (Lecun et al. 1998). A final 1×1 convolutional layer is applied to map the component feature vectors to a desired number of classes. The decoder performs upsampling in the form of transpose convolutional layers, these fulfill the role of deconvolution and upsampling simultaneously. Feature extraction is performed intrinsically by the algorithm in the contracting path and features are captured within the latent space. The encoder-decoder structure of U-Net is similar to AEs and VAEs, except for the presence of skip connections and no Kullback–Leibler (KL) divergence term as in the case of VAEs. Skip connections carry context between corresponding encoding/decoding layers. They are implemented in the decoder path by concatenating the output of each layer with the output of the corresponding encoder layer. The result of this is that the outputs of the encoding layers are carried over to the decoding layers and image information at different dimensions is maintained throughout the network. The functions of the different layers represented in Fig. 3 are described below:

- '2DConv': 2D convolutional layer scans over the 2D input data with convolutional filters, summing the results into a 2D matrix of features. The convolutional layers perform feature extraction in the network.
- 'BatchNorm': Batch normalisation normalises the inputs to the layer and is used to accelerate and stabilise the training process. The values are normalised in batches by subtracting the mean and dividing by the standard deviation in order to re-center and re-scale the data.
- 'ReLu': activation layer with rectified linear unit (ReLu) function that introduces non-linearity in the network to alleviate the issue of vanishing gradients. It is a popular activation function for training deep convolutional models.
- 'MaxPool': 2D max pooling is a non-linear downsampling operation that reduces the spatial dimensions of the data parsed to the layer. The objective is to reduce computational complexity in the network by making the representation smaller and therefore reduce the amount of information in the image. Max pooling aids in making the feature representations invariant to changes in scale and orientation.
- 'Dropout': dropout is used to prevent overfitting in neural networks, this is called regularisation, and involves randomly setting some of the neurons in the layer to zero. The process improves the accuracy of predictions for new unseen data.
- '2DTransConv': 2D transpose convolutional layer is used for upsampling in the expanding path in order to reconstruct an image of the same size as the input data. It is the opposite of the 2D convolutional layer.
- 'Conc': concatenate is used to combine tensors along a given axis, in this case it is used to combine feature maps from encoding layers, carried across via skip connections, with those of the corresponding decoding layer. This process allows U-Net to learn multi-scale features, where features at one scale might build upon or relate to features at other scales and has provided significant improvements on previous segmentation models.

Normalised data cutouts are parsed to the encoder and are reduced to some low dimensionality space, the latent space. The latent space captures the features within the data before it is deconvolved back through the decoder to the original input size, this output is known as

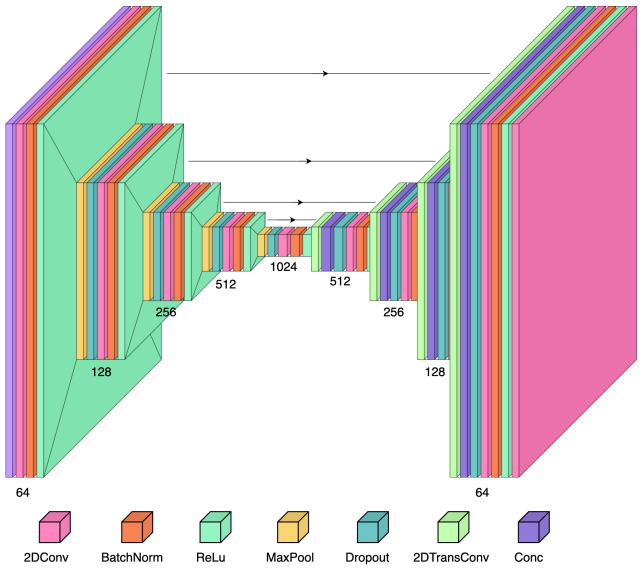

Figure 3. U-Net network architecture diagram made using the VISUALKERAS Python package (Gavrikov 2020). The arrows represent skip connections. The printed numbers are the number of tunable parameters at each of the layers, which are adjusted throughout the training process. The input image size is 256×256, being reduced to a representation of size 16×16 at the bottleneck. The model is trained by parsing each training image through the architecture and calculating the loss between the reconstructed image and the labelled segmentation map. The loss is minimised through training epochs until the model converges.

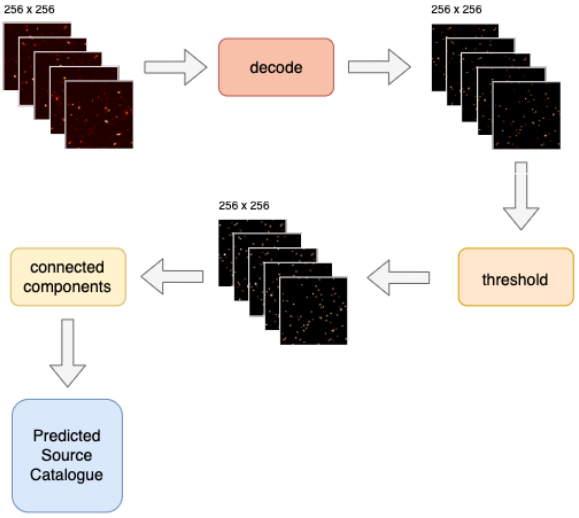

Figure 4. Graphic depicting the inference module. Data cutouts are parsed into the module to return a source catalogue via predictions made by the pre-trained model output by the learning module.

the reconstruction or decoded image. We use a Binary Cross Entropy (BCE) loss function, commonly used for binary classification problems. Here we are trying to classify each pixel as either 1 (source) or 0 (background). BCE loss measures the difference between the predicted probabilities for each pixel and their true binary labels. Deviation of the predicted probability distribution from the true labels results in an increase in BCE loss. BCE loss is defined as:

$$\mathcal{L}_{BCE} = -\frac{1}{N}\Big(\sum_{i=1}^{N} y_i \log(\hat{y}_i) + (1-y_i)\log(1-\hat{y}_i)\Big) \tag{2}$$

where y_i is the true binary label and $\hat{y}_i$ is the predicted probability of the positive class of a pixel within some data set consisting of N training samples. The $y_i \log(\hat{y}_i)$ term penalises the network when $y_i = 1$ but $\hat{y}_i$ is close to 0. Conversely, the $(1-y_i)\log(1-\hat{y}_i)$ term penalises the network when $y_i = 0$ but $\hat{y}_i$ is close to 1. The loss is therefore the negative log likelihood of $y_i|\hat{y}_i$. During training we minimise this loss in order to make the predicted probability for each pixel close to the true labels.

We train with training and validation data, consisting of 744 and 96 256×256 pixel size data-map pairs respectively, to ensure the model is not overfit. Once the model has converged, where training loss is minimised without increase in validation loss, we stop training and save the model. The output of the learning module is a pre-trained model that can be used to perform inference on new data.

2.1.3 *Inference Module*

The inference module is given in Fig. 4. Once our model is trained we can then generate predicted segmentation maps for our test data set. We parse the test set into the trained model and the reconstructed images are the predicted segmentation. The predicted segmentation is a probabilistic map of pixel values and must be converted to a binary segmentation map via post-processing. The decoded images are converted to binary by applying a thresholding method, selected based on metrics discussed in Section 4. Once thresholded, we can label each component and extract information about these regions using the SCIKIT-IMAGE [1] package. We extract the x and y pixel coordinates, the major and minor axis of the source, the orientation and the total intensity within that region. Some conversions are required to produce the required source parameters, including correcting the total intensity for the beam area which gives us the flux density of a source, and converting the orientation angle to a position angle by transforming the coordinate system of the region properties to be consistent with a polar coordinate system.

Pixel values in radio continuum data are given in units of Jy/beam, a conversion must be applied in order to obtain the total flux density of sources in Jy. The total intensity of a source can be measured by summing the pixel fluxes in that region, $\sum(S_{pix})$, which can then be converted into a total flux density value, S_{tot} according to:

$$S_{\rm tot}({\rm Jy}) = \sum(S_{\rm pix})\frac{8\log 2 \times \delta_{\rm RA} \times \delta_{\rm Dec}}{2\pi \times b_{\rm maj} \times b_{\rm min}} \tag{3}$$

where $\delta_{\rm RA}$ and $\delta_{\rm Dec}$ are the angular sizes of the pixels in the RA and Dec directions respectively and $b_{\rm maj}$ and $b_{\rm min}$ are the major and minor axes of the restoring beam.

3 SKA SCIENCE DATA CHALLENGE 1

This work has been carried out and tested on the *SKA* Science Data Challenge 1 images (Bonaldi & Braun 2018). The dataset consists of simulated radio continuum images taken at 3 different observing frequencies:

[1] https://scikit-image.org/

Table 1. Table of converted map sky sizes and pixel sizes for each frequency band. The map sky size is given in degrees on a side.

Frequency (Hz)	Map Sky Size (deg)	Pixel Size (arcsec)
5.6×10^8	5.50	0.600
1.4×10^9	2.20	0.240
9.2×10^9	0.33	0.037

- 560 MHz
- 1.4 GHz
- 9.2 GHz

and 3 integration times:

- 8 hours
- 100 hours
- 1000 hours

giving 9 images in total. The field of view (FoV) was selected for each frequency such that the image covers the angular size of the primary beam for a single telescope pointing. This results in a different map sky size for each frequency band. The size in pixels is consistent throughout all images, at 32768 x 32768 pixels, thus a pixel corresponds to a different angular size on the sky in each frequency band. The map sky sizes and the pixel sizes are given in Table. 1 respective to each frequency band.

Increasing the integration time increases the signal to noise of the images resulting in more of the faint source distribution being distinguishable from the noise, whereas sources can present quite differently at different frequencies. The statistics of the sources within each image therefore vary significantly with frequency and integration time. The longest integration will have a larger dynamic range in source flux density than the the shortest. An 8 hour image may contain fractionally more extended AGN than SFGs, whilst a 1000 hour image would have more SFGs. Thus training our model on images taken at different frequencies and integrations will help our trained model become invariant to changes in these observing factors and also understand how sources can appear differently with respect to the local population depending on such factors. Examples are shown in Fig. 5 where cutouts of the same field are taken from the 1.4 GHz images at the three different integrations to demonstrate the variation in the data.

A "Truth Catalogue" is provided for each frequency band, which includes the location and properties of all sources in each image. A "Training Catalogue" is also provided for each frequency band, which is a filtered truth catalogue containing the subset of sources within a region covering an area of 5% of the full image. These sources can be used as training labels for different machine learning applications. It is important to note that these catalogues contain a significant number of sources below the nominal flux limit which is of the order ∼1 nJy/beam (Bonaldi et al. 2020), and cannot be detected within the image itself. It is also worth noting that whilst each frequency band has its own truth catalogue, this is not the case for the different integration times. As such, the proportion of the catalogue detectable in the 8 hour image at 1.4 GHz for example, is not the same as that of 1000 hour 1.4 GHz image.

The sources injected into the sky images are a combination of SFGs and AGN. Sources whose major axes are larger than 3 pixels are considered extended, and those smaller than 3 pixels compact. As mentioned above, pixel size depends on frequency. Thus a source treated as extended in one continuum map may be compact in another.

There are 3 classes of object injected into the dataset:

- Steep-spectrum AGN, including Faranoff-Riley radio galaxies (Fanaroff et al. 1974)
- Flat-spectrum AGNs, galaxies with a compact core component and a single lobe viewed end-on
- Star forming galaxies

The compact sources have been modelled as 2D Gaussian objects. A library of real AGN images were used to inject the steep-spectrum AGNs into the data, scaled in total intensity and size, randomly flipped or rotated and then placed at random into the images. The flat-spectrum AGNs were added as a pair of components.

3.1 Pre-Processing

The U-Net architecture in ContinUNet's learning module requires enough images to adequately train on. Since the images are large (32768×32768 pixels), we can generate many cutouts for training. We parse all images from the 560 MHz and 1.4 GHz observing frequency bands to the pre-processing module described in Section 2.1.1. We currently have limited the training and testing of the model to two frequency bands. Whilst sources appear different at different frequencies, they possess characteristic statistical features that we want the network to learn. Training on different integration times and frequencies should make the model invariant to such changes in data. Sources at 9.2 GHz are much larger and with much lower source density. As a result, we could not find a cutout size suitable for training the 9.2 GHz band simultaneously with the other bands that would account for the very small and densely populated sources at one frequency and the sparse large sources at another. It is possible to take cutouts and resize them to train together, but we did not attempt this as resizing images may adversely affect the morphology of sources and introduce unwanted bias into the training set. In the future we plan to resolve this dynamic range issue and incorporate higher frequency bands into our training set. The images are trimmed to the area equivalent to the 5% training set provided for that frequency band. Each image is divided into 256×256 pixel cutouts.

We cannot include all sources from each cutout catalogue in the segmentation maps due to many of the sources having a very low signal-to-noise ratio (SNR), particularly very large diffuse sources that cannot be resolved from the foreground sources. We must make an SNR cut in the truth catalogues in order to produce segmentation maps suitable for training. It should be noted that these segmentation maps produced are imperfect as they do not include the full source distribution within the image, and we must assume that the model can not truly learn to understand what 'background' looks like.

3.2 Training

Once we have generated our data-map pairs, we split the data 80:10:10 for train, validation and test data sets respectively. Since we are using cutouts from all 560 MHz and 1.4 GHz images, the data in the different integration time images are duplicated but with a higher signal to noise for sources. We must ensure that cutouts of sources are not duplicated in the training and validation or training and test. If they were to be duplicated, it would result in over fitting to the training data, or our test set not being a true blind study. An example of duplicated cutouts is given in Fig. 5. To solve this, we split the data-map pairs for each image before concatenating the data sets together. Training the images together allows us to increase our training set size. This method gives us 744, 96 and 96 images for training, validation and testing respectively.

We train our data using the learning module described in Section

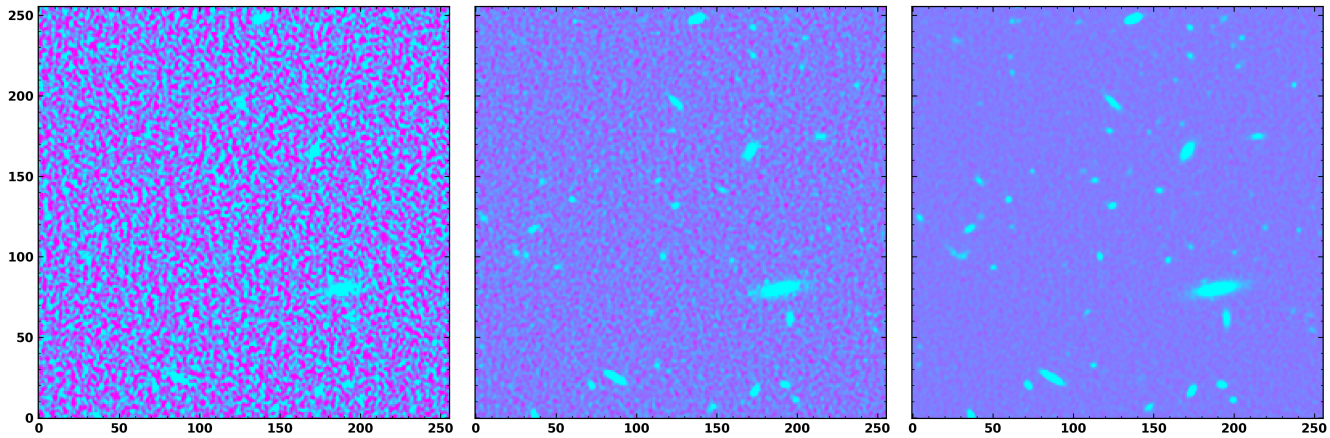

Figure 5. This figure shows cutouts of the SDC1 data, taken at the same coordinates for different integrations but the same frequency. The colour scale represents pixel intensity and is normalised between all images. These cutouts are all from the 1.4 GHz images taken (from left to right) from the 8, 100 and 1000 hour integrations respectively. We can see here how the longer exposure time affects the signal-to-noise ratio of the sources, with much more of the source distribution becoming apparent in the 1000 hour image and the morphology of the larger diffuse extended source becoming more clear.

2.1.2. The network is implemented using TensorFlow GPU. We train on training data and validation data simultaneously to avoid over fitting of the training data. This is done by balancing the training loss and the validation loss, the validation loss should always decrease with the training loss, if not the model is being over fit. We reduce the learning rate on plateau of the validation loss and utilise early stopping to avoid over fitting to the training data. The model is check pointed when the validation loss is improved and this model is saved. We reach equilibrium at 65 epochs, where validation and training loss have both reached a plateau.

3.3 Prediction

Using the trained model output by the learning module, we can generate predicted segmentation maps for our test data set using the inference module described in Section 2.1.3. We use a threshold method from scikit-image to convert the predicted segmentation into binary maps. We compared source-finding using a selection of thresholding methods available from scikit-image, but settle on the Otsu threshold (Otsu 1979) as it provides a suitable compromise between precision, recall and association of source components.

4 RESULTS

We perform source detection on the blind test data set using the inference module described in Section 2.1.3 by parsing normalised test data cutouts to the inference module to produce binary segmentation maps, source catalogues and contour plots.

4.1 Match Predictions to Ground Truth

In order to quantitatively assess our model's performance we must first determine which of our detected sources can be considered real detections. We must match predicted sources to ground truth sources for each cutout. Once we have matched our sources we can perform analysis on the matched and predicted sources to understand ContinUNet's source-finding capabilities. Due to the dense population of true sources in our cutouts, a simple nearest neighbour matching is not sufficient. We perform first a nearest neighbour matching on every ground truth source to find potential candidates in the predicted sources. The distance for this nearest neighbour matching is set as a function of source size, this allows for more lenience on larger sources and more strict matching on small point sources where the source population is most confused. The detection radius, r is given by:

$$r = \log \pi a_t b_t \times 10 \tag{4}$$

where a_t and b_t are the major and minor axes of the ground truth source respectively. For every candidate we calculate the intersection over union (IOU) score, this is also used as our segmentation analysis metric later, see Section 4.6 for details. The IOU score is the total area of intersection of two sources divided by the area of their union. This allows us to compare the predicted position angle, major and minor axes simultaneously and thus choose the predicted source that most closely matches the ground truth source in location, shape and size. We set a threshold of 0.3 for IOU score and if any ground truth source has no candidate matches that have an IOU score above this threshold the source is considered missed and marked as a false negative. Matched sources are marked as true positives. Figure 6 is a diagram depicting the algorithm used to match detections.

4.2 Method Comparison

We use PyBDSF and ProFound as benchmark methods to assess the quality of our model's performance as a source detector. We run both PyBDSF and ProFound on our test cutouts. For PyBDSF we use the `process_image` method with default parameters. We use the `MakeSegIm` method from ProFound to generate the predictions. Both methods produce source catalogues for each cutout that we perform source matching on according to the algorithm described in Section 4.1.

We perform quantitative analysis on detected sources from all methods. Both PyBDSF and ProFound were used with default parameters on their source-finding methods. This choice was made as we want to test the performance of all methods as 'out-of-the-box' source-finding methods. Our aim is to develop a method that can be

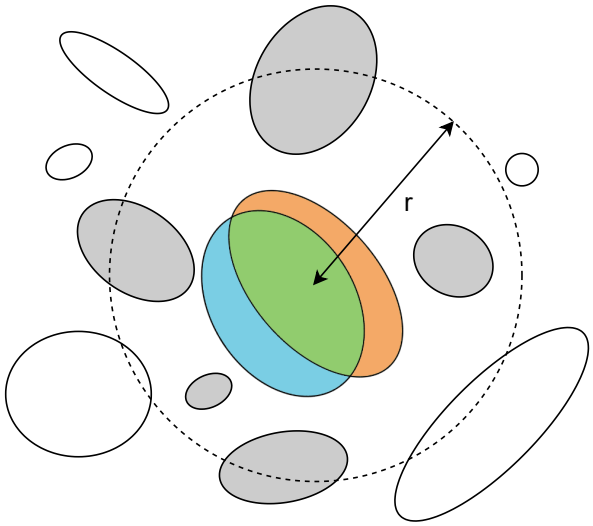

Figure 6. Graphic depicting the source matching algorithm. The orange ellipse in the center is the ground truth source to be matched. The dashed circle marks the region at detection radius, r as given by Equation 4. All grey ellipses whose centers lie within this region are predicted sources marked as candidates. The white ellipses whose centers lie outside this radius are predicted sources not marked as candidates for matching. The blue ellipse is the predicted source with the highest IOU score of all the candidates and is selected as the matched source. The intersecting area of the ground truth source and the matched predicted source is shown in green.

run with no parameter tuning and little to no expertise required in order to produce a source finder with optimum end user experience and consistent and fast results.

For the comparisons we include only the results for the 1.4 GHz 1000 hour SDC1 image. We include only the 1000 hour results as each frequency band has its own truth table, but this is not true for each integration time. As such, true sources are included in this table that are not detectable in the image. Thus we use the 1000 hour image as all of the sources that are detectable in the 8 hour and 100 hour images exist in the 1000 hour image. We do see a slight improved performance of ContinUNet over the other methods for the shorter integration times, suggesting that ContinUNet can perform source detection at lower SNR, but we do not include the figures for the sake of brevity. Radio continuum at lower frequencies typically show more extended sources, as emission at lower frequencies is typically dominated by older electron populations. Thus comparing performance on the 560 MHz band images would preferentially bias each model's ability to handle diffuse extended emission. However, we only include results for the 1.4 GHz band as there are sufficient extended sources within these data to demonstrate the comparison between each model fairly whilst maintaining brevity.

We include an example of one of the test data set cutouts taken from the 1.4 GHz and 1000 hour images with the ellipses of sources detected by ContinUNet, ProFound and PyBDSF overplotted in Fig. 7.

4.3 Source Population Recovery

We assess the recovery of the source population by each method using precision, recall and F_1 given by Equation 5, 6 and 7 respectively, where TP is true positive, TN is true negative, FP is false positive and FN is false negative. Precision is defined as:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{5}$$

and can be considered the same as purity, it is the proportion of your predicted sources that are real detections. This is an important metric for source-finding as detecting too many false positives will reduce the purity of a sample and thus impact source counts measured from it. A high precision also means that less clean up of predicted catalogues is required. Recall is defined as:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{6}$$

and can be considered the same as completeness, that is the proportion of the true sources that were positively identified by the model. Recall is also an important metric for source-finding as we want to recover as much of the source distribution in one pass as possible. This improves the completeness of our sample and means less work is required to recover the entire population. F_1 score is defined as:

$$\text{F}_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{7}$$

and is a combination of precision and recall with both metrics weighted the same. This metric is useful as we can consider both the precision and recall together to give a single metric for source population recovery. We chose F_1 over F_β, where the relative importance of precision and recall can be weighted accordingly, as we consider both precision and recall to be equally important metrics for source-finding. We give precision, recall and F_1 score as a function of flux density for all methods in Fig. 8, 9 and 10 respectively.

Information about the radio source populations can be inferred by measuring population statistics. It can be assumed that the number of sources detected at a given flux density is an indicator of both the properties of the emitting source population and the cosmology of that region, since at any given flux density we see both luminous distant sources and faint local ones. Source counts are measured in radio observations to understand the distribution of these properties in the entire source population. The source count distribution describes the number of sources per unit flux density. Source counts can be measured in different frequency channels to construct a picture of how source populations in different parts of the sky are changing over cosmic time (Mandal et al. 2020).

We calculate the differential source counts for predicted sources and matched sources for all models. Source counts are used to quantify the number of sources N within a flux density S bin per unit steradian observed on the sky, defined as:

$$\log\left(\frac{dN(S)}{dS}S^{2.5}\right) = \sum_{i=0}^{n} a_i (\log S)^i \tag{8}$$

The source counts are Euclidean normalised, denoted as $n(S)S^{2.5}$ (Hale et al. 2019). The Euclidean source counts are given for all source predicted by each method in the 1.4 GHz 1000 hour SDC1 image in the left hand figure of Fig. 11 and for all matched sources in the right hand figure.

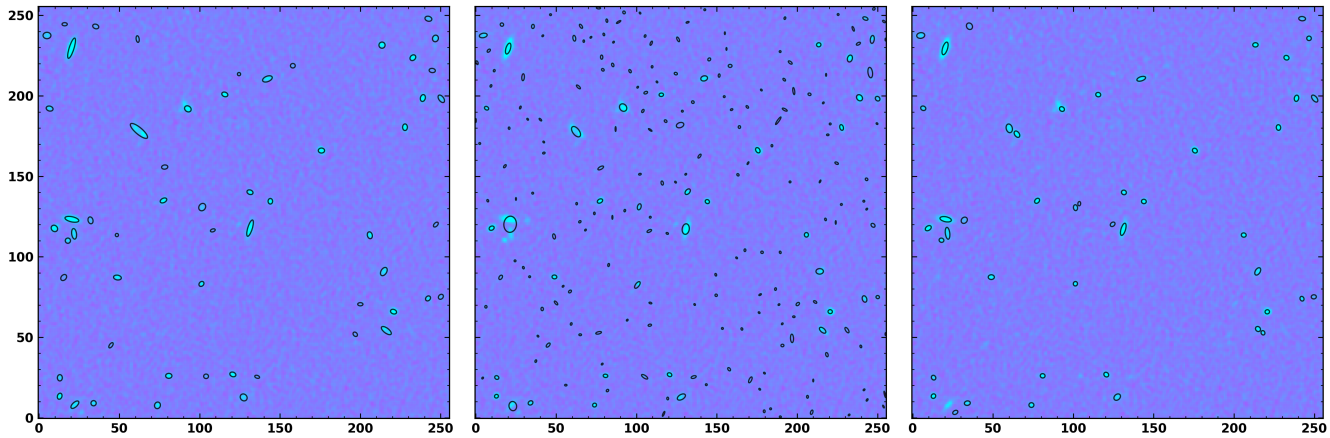

Figure 7. Example cutout taken from the 1.4 GHz 1000 hour test cutouts, with the ellipses of the sources predicted by each method overplotted in black. Left: ContinUNet, middle: ProFound, right: PyBDSF.

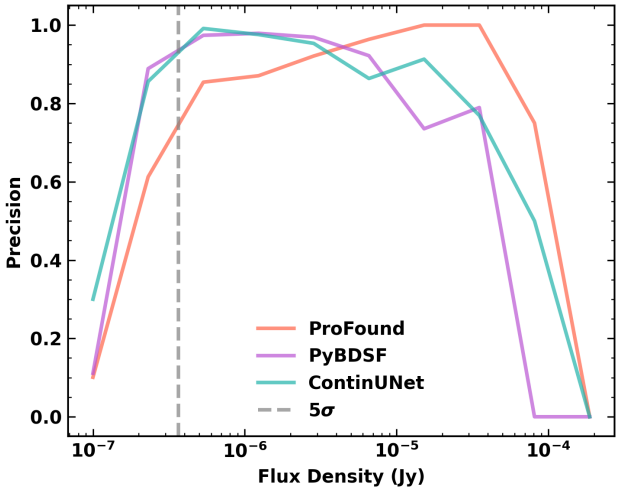

Figure 8. Precision at different flux density bins for sources detected by all methods in the 1.4 GHz 1000 hour SDC1 image.

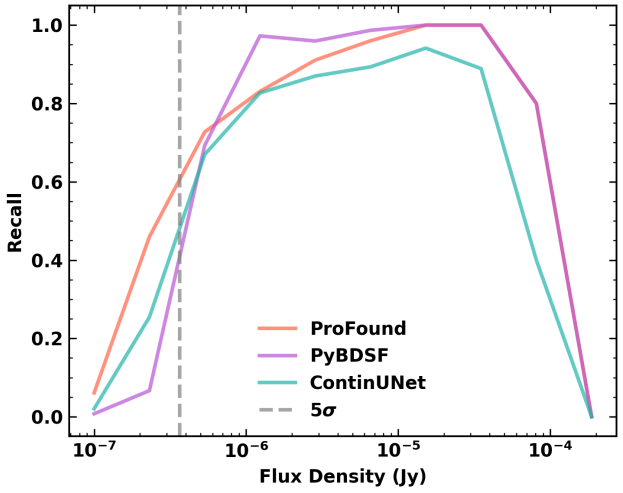

Figure 9. Recall at different flux density bins for all sources detected by all methods in the 1.4 GHz 1000 hour SDC1 image.

4.4 Centroid Location Accuracy

Accurately detecting the central coordinates of a source is key for source-finding as any cross matching with source catalogues in other wavelengths will be strongly impacted by inaccuracies in these values. This issue will become more apparent as we move into the era of next generation surveys such as *SKA* as we expect a high source population density in the resulting catalogues. We compare the accuracy of detection of centroid location for matched sources for all methods. The distribution of spatial distance between detections and their matched ground truth source is shown in Fig. 12, where spatial distance d is given by:

$$d = \sqrt{({x_t}^2 - {x_p}^2) + ({y_t}^2 - {y_p}^2)} \tag{9}$$

and x_t, y_t, x_p, y_p are the image coordinates for the true and predicted source.

4.5 Flux Density

We compare the flux density distributions for predicted and matched sources for all methods. Accuracy in measurements of flux density will impact source counts extracted from surveys which will in turn impact our understanding of the distribution of AGNs and SFGs.

We compare the flux density distribution for all sources predicted by each method to the total flux density distribution of all sources and also the flux density distribution of matched sources predicted by each method in Fig. 13. For matched sources we compare the distribution of the ratio of predicted to true flux density in Fig. 14 and compare predicted versus true flux density as a 2D histogram in Fig. 15.

4.6 IOU Score

We can calculate the intersection over union (IOU) score for all matched sources. This is a way to determine the accuracy of the segmentation. We perform it here only on the matched sources to

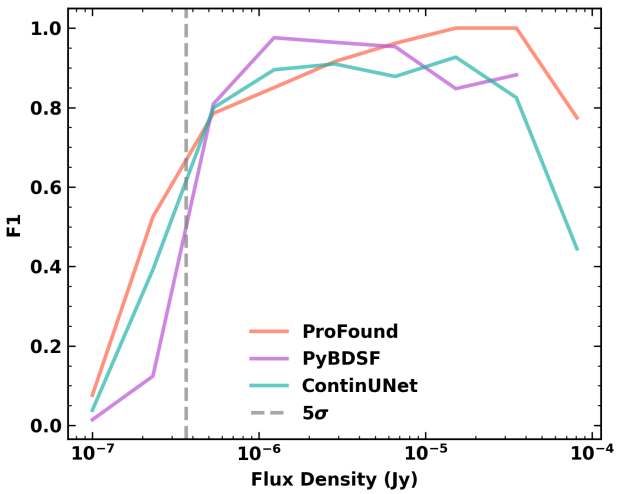

Figure 10. F1 score at different flux density bins for all sources detected by all methods in the 1.4 GHz 1000 hour SDC1 image.

determine how accurately each method has reconstructed the shape of the ground truth sources. It is measured by calculating the intersecting area of the predicted and true source segmentation, divided by the area of the union. The union is the total area of both segmentation minus the intersection. The distribution of IOU scores for matched sources is shown in Fig. 16. The equation for IOU is given by:

$$\text{IOU} = \frac{\text{Area of Intersection}}{\text{Area of Union}} \tag{10}$$

5 MIGHTEE EARLY SCIENCE DATA

Having performed extensive analysis on the test data set from SDC1, we applied the ContinUNet framework to real radio continuum data. We use the science ready continuum image of the COSMOS deep field reduced by (Heywood et al. 2021) from the MIGHTEE Survey (Jarvis et al. 2016). The image contains a wide variety of sources from small to giant radio galaxies (Delhaize et al. 2021) and is extremely densely populated with compact sources. There are both low resolution/high sensitivity and high resolution/low sensitivity images of the COSMOS field produced from the MIGHTEE survey data. We use the low resolution image as it more sensitive, contains more sources and has more diffuse connecting emission between large extended sources, although the central region of the image suffers from great confusion than the high resolution image. The low resolution image covers an area 1.62 deg^2, with an angular resolution of 8".6, thermal noise level of 1.7μJybeam^{-1} and has a classical confusion limit of approximately 4.5 μJybeam^{-1}. Whereas, the high resolution image has an angular resolution of 5" and a 1σ noise level of 5.5μJybeam^{-1}.

We process the MIGHTEE low resolution image with ContinUNet and extract a source catalogue using the inference module described in Section 2.1.3. We do not perform any tiling on the image but process the full field in one pass. The model was not retrained for inference on MIGHTEE data, we used the pre-trained model trained on the simulated SDC1 data, the only modification is the threshold method used. In the SDC1 data we use the Otsu threshold (Otsu 1979), for MIGHTEE we use the Triangle threshold method (Zack et al. 1977). We use this threshold due to the increased dynamic range of this image in comparison to the test images, due to the smaller (256×256 pixel) size of the latter compared with the full MIGHTEE image. The lower value of the triangle threshold allows for the connection of components of extended sources in the binary segmentation maps.

We create a model map of sources in the image by multiplying our binary segmentation map by the input image. The model map is then used to create a residual image by subtracting it from the input image. The MIGHTEE image and the residual image after removing the contribution from the sources extracted by ContinUNet are shown in Fig. 17. We find that ContinUNet is capable of segmenting real radio sources without the need for retraining. A closer view of the performance of ContinUNet on the MIGHTEE continuum image in Fig. 18, which shows a sub section of the image at the center of the FoV, corresponding predicted segmentation map produced by ContinUNet and the resulting model map produced of that region. This image sub section showcases the variation in dynamic range in both size, brightness and complexity of sources within MIGHTEE. The model map produced by ContinUNet clearly demonstrates the performance of our model at managing complex data with minimal tuning.

We calculate the Euclidean normalised source counts for sources detected by ContinUNet above 5σ using Equation 8. We compare these counts to those of the component catalogue predicted using PyBDSF in (Hale et al. 2022) which contains 9896 components, and the source catalogue cross correlated with optical and near-infrared data produced in (Whittam et al. 2023) which contains 5223 sources, in Fig. 21. The 5σ limit is determined using the confusion limit of 4.5μJy rather than the theoretical noise limit of 1.7μJy.

We also include cutouts of two complex extended sources from the MIGHTEE image detected by ContinUNet with their segmentation boundaries over plotted in Fig. 19 and Fig. 20. These sources are present in the cross matched catalogue (Whittam et al. 2023) and consist of 17 and 47 associated PyBDSF components respectively. These components were associated by hand through manual intervention performed by the MIGHTEE collaboration.

5.1 Computation Time

We perform some light benchmarking for ContinUNet using the MIGHTEE image (≈1 square degree, ≈120 MB, ≈ 5500 square pixels) on a local development environment on an Apple MacBook Pro (2.4 GHz Quad-Core Intel Core i5 Processor, 16GB 2133 MHz LPDDR3 onboard memory) and with CPU on a supercomputer (AMD EPYC 7702P CPU, 64 cores, 128 threads, 1TB RAM).

The processing time refers to the wall time, i.e. the time experienced by the end user when performing source-finding. This is not the same as the total computational processing time which is the actual processing time on every core used for each process. The load processing time is the time taken to load the image into memory and perform pre-processing steps such as reshaping the array and normalising the image, this is the processing time for the pre-processing module. The inference time is the time taken to produce the predicted segmentation map from the loaded image. This is the inference module processing time and intuitively the bottleneck of the source-finding. The following times are both part of the post-processing module, but separated into labelling time and cleaning time to demonstrate the limiting step. Labelling time is the time to cluster pixels in the thresholded segmentation map and label

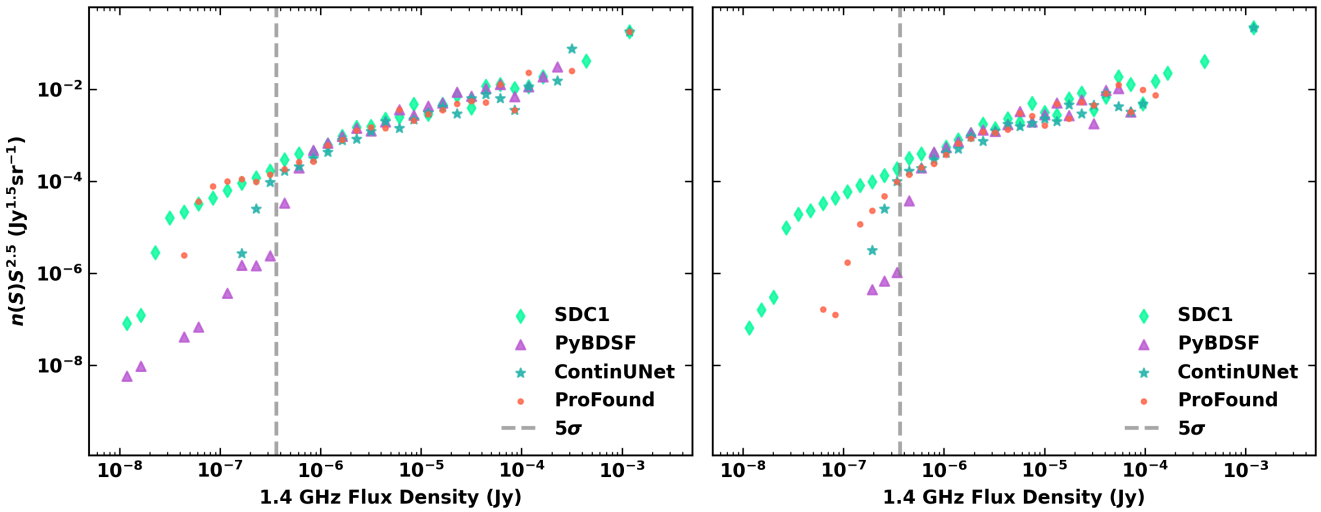

Figure 11. Left: Normalised 1.4 GHz Euclidean source counts for sources predicted by ContinUNet, ProFound and PyBDSF, with the true source counts as given by the SDC1 truth catalogues. Right: Normalised 1.4 GHz Euclidean source counts for source predicted by ContinUNet, ProFound and PyBDSF that have been matched to a ground truth source, with the true source counts as given by the SDC1 truth catalogues. Only sources predicted in the 1000 hour image from the 1.4 GHz image are included.

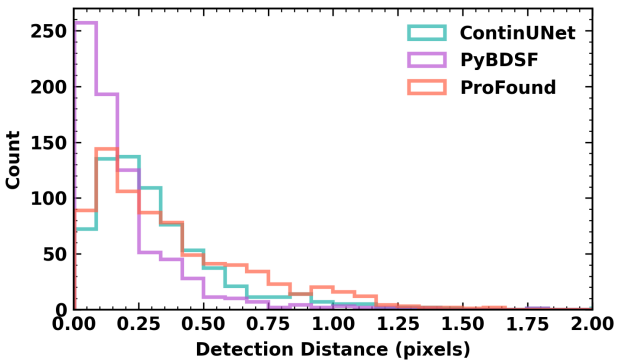

Figure 12. Distribution of distance in Cartesian space of centers of predicted sources and the true source they have been matched to. Sources are those that have been detected by ContinUNet, ProFound and PyBDSF in the 1.4 GHz 1000 hour SDC1 image and have been matched to a true source, the sources included in this figure are those whose measured flux density is greater than the 5σ background noise.

each source accordingly, it is performed by SCIKIT-LEARN [2] and thus performance cannot be optimised here without an additional computational approach. Cleaning time is the time to perform cleaning steps on the sources such as area and flux corrections and noise cuts, this step has been developed using numpy methods which are already optimised for high performance computing. The total time is the sum of all of these steps and demonstrates the end to end time from parsing the MIGHTEE image to ContinUNet to generating a cleaned post-processed source catalogue and corresponding segmentation and model maps and residual image. Computation times for all three hardware environments are given in Table 5.1.

[2] https://scikit-learn.org/stable/

	MacBook Pro	HPC CPU
Load Time (s)	1.296	0.198
Inference Time (s)	**68.769**	**8.969**
Labelling Time (s)	51.426	3.691
Cleaning Time (s)	2.637	0.0142
Total Time (s)	**124.158**	**12.873**

Table 2. Computational times for different processes within the ContinUNet source-finding framework when run on the MIGHTEE low resolution image on two different computational environments.

The model used for inference was trained using a GPU (NVIDIA TU104GL [Tesla T4] 2560 cores, 16GB RAM), on a training set of simulated data of size 744×256×256 for training and 96×256×256 for validation. The model converged after 65 epochs and training took 6 minutes 14 sec.

6 DISCUSSION

In the following section we will discuss the performance of the detection methods on the simulated SDC1 data and discuss the results of applying ContinUNet to real MIGHTEE data. We will then discuss what the outcomes of both investigations mean for source-finding in upcoming *SKA* data releases.

6.1 Source Count Recovery

In terms of precision, recall and F1, none of the tools assessed clearly outperforms the others at all flux densities, see Fig. 8, 9 and 10 respectively. For precision, PyBDSF and ContinUNet show similar trends with flux density, but ProFound behaves quite differently, with precision increasing from lower to higher flux densities. F1 score is an appropriate metric for source-finding as it takes into account precision and recall with equal weighting, giving a good representation of how the models perform in terms of recovering the entirety of the

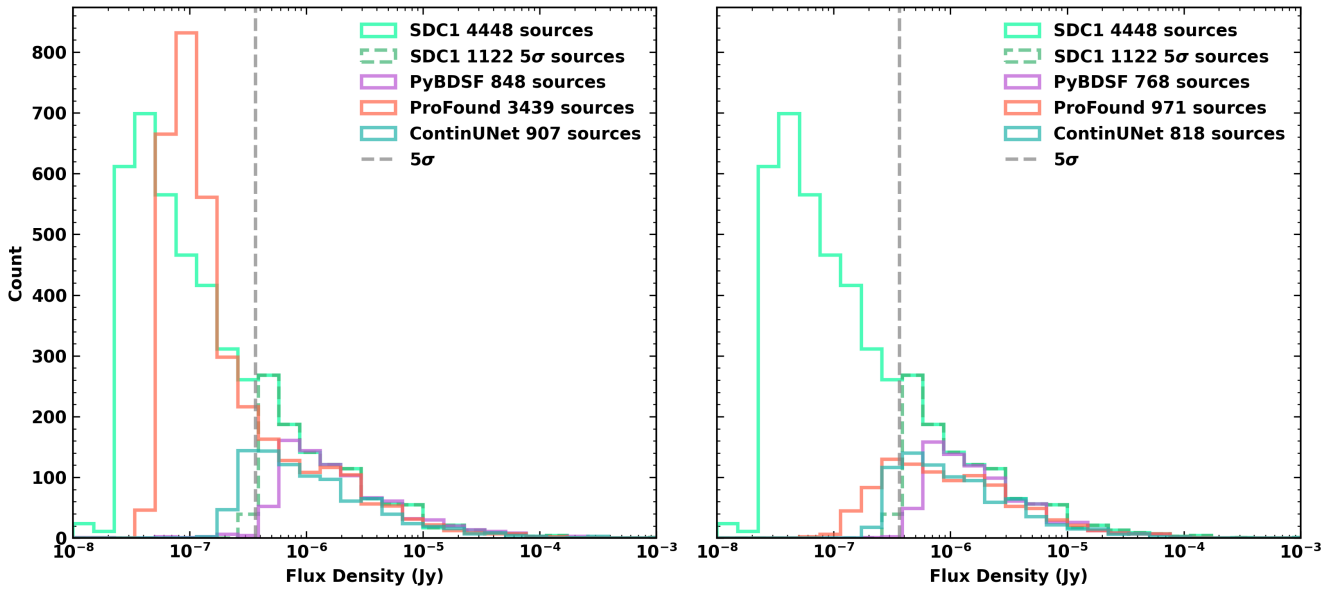

Figure 13. Left: distribution of measured flux densities of sources predicted by ContinUNet, ProFound and PyBDSF in the 1.4 GHz 1000 hour SDC1 image. Right: distribution of measured flux densities of matched sources predicted by ContinUNet, ProFound and PyBDSF in the 1.4 GHz 1000 hour SDC1 image. The flux densities of the full source population injected into the data set are included for reference, as are the flux densities of true sources whose flux density is greater than 5σ. The numbers of sources are the total sources respresented by each distribution.

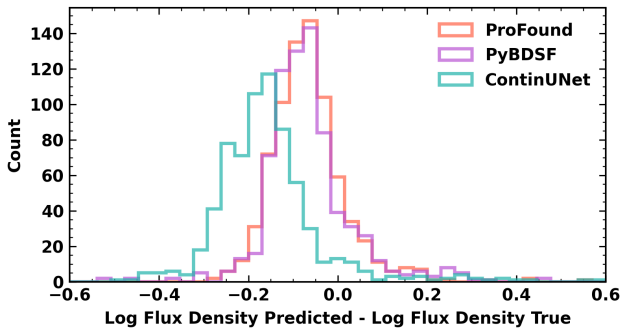

Figure 14. Distribution of difference between predicted log flux density and true log flux density of matched sources detected by ContinUNet, ProFound and PyBDSF in the 1.4 GHz 1000 hour SDC1 image. The sources included in this figure are those whose measured flux density is greater than the 5σ background noise. A ratio of 1 is a perfect reconstruction of the flux density of a source.

source population whilst minimising the number of false detections. For all methods there is a drop off in precision, recall and thus F1 score at the highest flux densities, where we would expect all metrics to approach 1. This is due to a bias introduced during matching. The high density of the source population means that we cannot match using a nearest neighbours algorithm, thus we match using the IOU score of the predicted sources. However, this introduces bias as IOU score will decrease with increasing source size as there are more pixels to accurately match. The larger, more complex and extended sources are typically found at higher flux densities. This results in a drop in both precision and recall as sources labelled as false positives are correctly detected but their segmentation is not good enough to count as a match.

We see a high number of false positives detected by ProFound at lower flux densities in Fig. 13, and there is a strong discrepancy between the predicted and matched ProFound source counts. This high false positive rate explains the lower precision at lower flux densities for ProFound in Fig. 8. However, since the majority of these false positives have a measured flux density below the 5σ noise limit, the drop in precision is also below the noise limit and therefore is not a significant drawback as these sources can be easily removed with flux cuts.

In Fig. 11 we again see comparable performance between all methods when recovering the source population above the 5σ noise limit. There is a distinctive drop off at 5σ for the PyBDSF source counts and also flux distributions in Fig. 13, due to the 5σ flux cut which is set as default by PyBDSF. PyBDSF in general both predicts and matches fewer sources than ContinUNet and ProFound and many fewer at lower flux densities.

6.2 Source Parameter Recovery

The parameters recovered by ContinUNet, ProFound and PyBDSF are all quite tightly clustered with the ground truth. For matched sources all methods show a high degree of accuracy for detection of the source center location, with the majority of predictions being accurate to within 0.25 pixels as shown in Fig. 12. In Fig. 15 we see that the matched PyBDSF and ProFound sources have a measured flux density more tightly aligned with the true flux density compared with ContinUNet. Figure 14 shows the distribution of predicted log flux density minus true log flux density, to represent the accuracy of flux recovery, a perfect flux density measurement would sit at zero on a distribution. Whilst PyBDSF and ProFound are skewed to an under prediction of flux density of $\approx$20%, ContinUNet under predicts by $\approx$30%. We will be looking into improving the flux recovery of

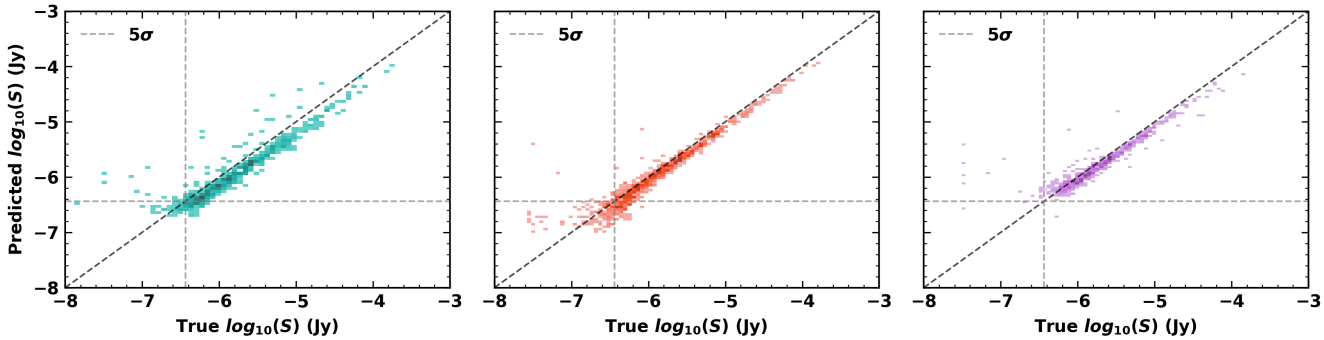

Figure 15. 2D histogram of true flux density vs predicted flux density of matched sources detected in the 1.4 GHz 1000 hour SDC1 image. Left to right are matched sources predicted by ContinUNet, ProFound and PyBDSF. The grey dashed lines represent the 5σ noise level for the image and the black dashed line is a one-to-one line. A perfectly recovered flux would sit on this line.

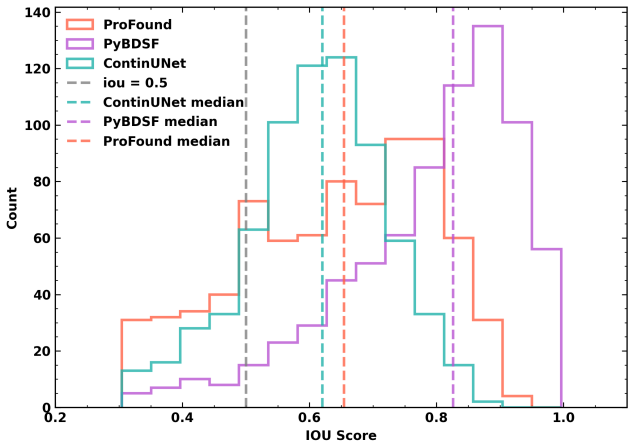

Figure 16. IOU scores for sources predicted by ContinUNet, ProFound and PyBDSF in the 1.4 GHz 1000 hour SDC1 image that have been matched to a true source. An IOU score of 0.5 or above is considered a good segmentation, a value of 1 is a perfect segmentation. The sources included in this figure are those whose measured flux density is greater than the 5σ background noise.

ContinUNet, but in this paper we are presenting the method for source-finding as oppose to photometry.

6.3 Source Segmentation

It is difficult to find a suitable metric for measuring segmentation accuracy as many of the sources cannot be modelled as elliptical. The extended sources have much more complex morphology, and although we have made estimates of the source boundaries when generating our segmentation maps for training, we do not have true segmentation for these sources and so cannot directly compare the segmentation of each method. We calculate IOU scores of ellipses created from the major and minor axes and position angle of the true sources and the predicted sources detected by each model. The distribution of IOU scores for matched sources predicted by each method is shown in 16, which shows superior performance in segmentation for PyBDSF in comparison to the other methods, with a high median IOU score. However, the median IOU score for all methods sit comfortably above 0.5, an accepted threshold for good segmentation. There may be systematic bias that favours the performance of PyBDSF due to the fact that the data set is simulated by injecting Gaussian blobs into the image and PyBDSF performs source-finding by fitting Gaussian models to sources. It is clear from Fig. 16 that PyBDSF performs well at characterising such sources. Whilst majority of the sources in the SDC1 data set are compact (or unresolved) sources that are well characterised by a Gaussian model, this approach is not appropriate for modelling extended or more complex sources, which we see evidence for in the real data experiments in Section 6.4.

6.4 Real Data Application

The SDC1 data set may not be representative of real data and thus it is important to see how our model performs in a real data case. The MIGHTEE data is a suitable representation of *SKA* data as *MeerKat* is a precursor telescope to *SKA* and the data are rich, densely populated with sources, and have a large dynamic range in terms of flux density and source complexity. We applied ContinUNet to the MIGHTEE Early Science data without retraining. We performed source-finding on MIGHTEE on a laptop in ≈2 minutes, the only modification required was changing the threshold method. Figure 17 shows the MIGHTEE image on the left and the residuals after removing the contribution from the detected sources on the right. The source contribution has been well characterised and the main residuals are from faint extended emission of large radio galaxies. ContinUNet transferred quickly and easily to real data and showed impressive performance at segmenting large extended sources. A closer view of this segmentation is given in Fig. 18, which shows the segmentation map and the model map for a subsection of the MIGHTEE image. The extended sources shown in this figure exceed the size limit of extended sources present in the training set, whose major or minor axes are smaller than 100 pixels. The largest source detected by ContinUNet in the MIGHTEE image has a major axis of ≈170 pixels. The fact that ContinUNet was able to recover these sources is extremely promising as it suggests that the model has learned the statistical and morphological features of radio sources, thus we can perform segmentation that is invariant to scaling changes. In the future we intend to supplement the training data with real data to improve robustness to variations in scale and SNR.

Source-finding was performed on the MIGHTEE Early Science data COSMOS field using PyBDSF (Hale et al. 2022), which produced a catalogue of Gaussian components where one or more was

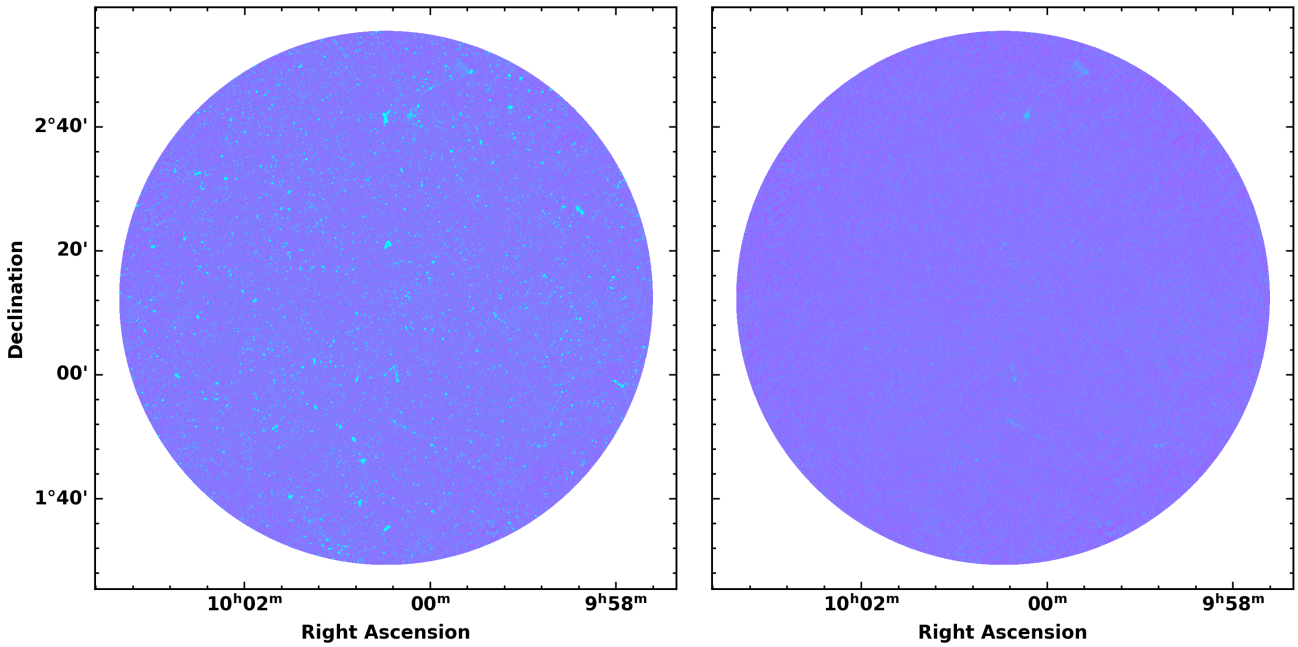

Figure 17. Left: MIGHTEE continuum image. Right: Residuals after subtraction of the ContinUNet model map. The residuals are quite clean of source contributions, with the only contributions left behind being from the GRGs and extended emission from large, bright complex AGN.

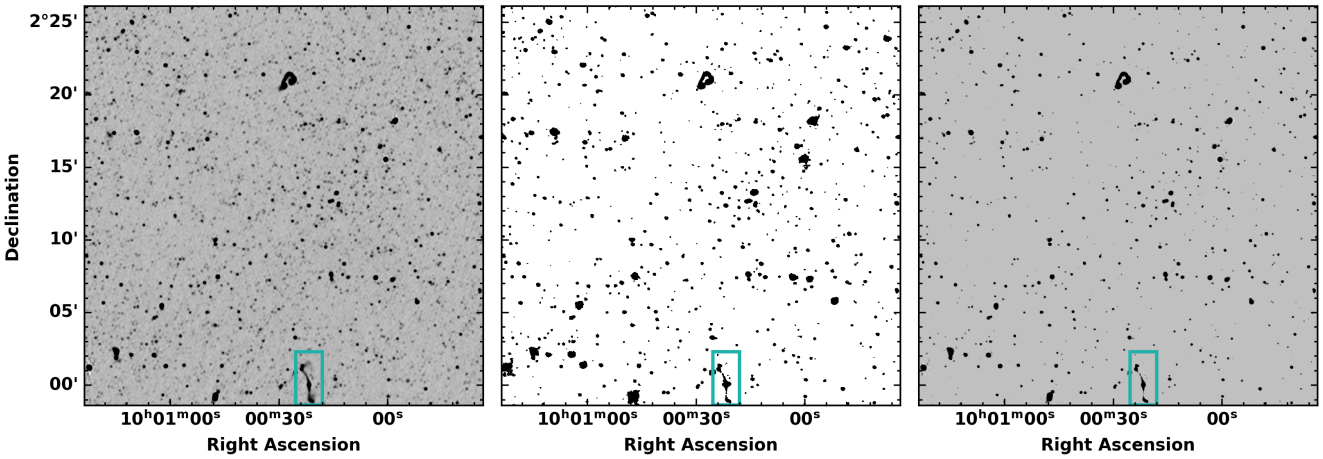

Figure 18. Left: Sub section of the MIGTHEE continuum image taken from the center of size 1500×1500 pixels. Middle: Corresponding segmentation map of sources predicted by ContinUNet, including sources below 5σ flux cutoff. Right: Model map of sub section produced by multiplying the segmentation map with the input data image. This central region of the FoV suffers the greatest confusion, but ContinUNet is still able to associate components of complex extended sources as can be seen in the highlighted teal box.

associated with a source in the image. Further processing on this catalogue was performed and the components were associated by hand through manual intervention and cross-matching with optical and near-infrared data (Whittam et al. 2023). Both of the sources shown in Fig. 19 and Fig. 20 can be found in the published catalogue. Figure 19 shows a near perfect segmentation of the bent tail radio galaxy by ContinUNet. However, this source was characterised by 17 PyBDSF components as stated in the manually cross-matched catalogue. Examples of complex extended sources are extremely common in the MIGHTEE image, and can be seen in Fig. 17 (left); such sources are not well characterised by Gaussian components. Considering that the MIGHTEE field is a fraction of the FoV of *SKA*, the approach of manual intervention cannot scale practically to next generation surveys, particularly considering expected frequency of such cases expected for *SKA*. Figure 20 shows another example of an extended radio galaxy. This source is actually two radio galaxies, but cannot be separated with the MIGHTEE data alone; cross-matching with higher frequency radio data allowed the two sources to be separated (Whittam et al. 2023). The top right hand tail is the second galaxy that is not associated with the main double radio galaxy in the centre

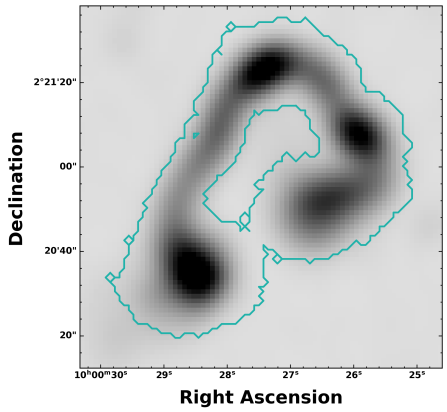

Figure 19. Cutout of extended source detected using ContinUNet with segmentation boundaries taken directly from the segmentation map over-plotted in teal. In the cross matched catalogue this source contained multiple PyBDSF components, $N_{\rm comp}$ = 17 (Whittam et al. 2023). ContinUNet has segmented the shape of this source whilst correctly associating all of the clear components as one source.

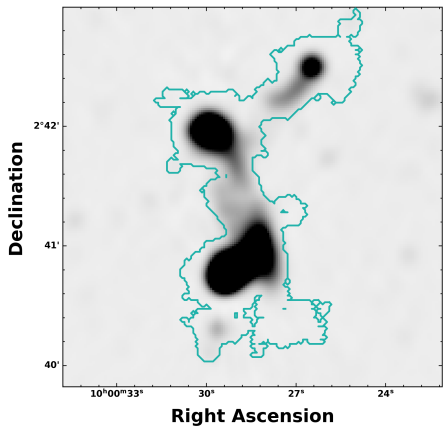

Figure 20. Cutout of extended source detected using ContinUNet with segmentation boundaries taken directly from the segmentation map over-plotted in teal. In the cross matched catalogue this source contained multiple PyBDSF components, $N_{\rm comp}$ = 47 (Whittam et al. 2023). ContinUNet has produced a segmented source that includes the multiple components visible with little connecting emission.

of Fig. 20. This main galaxy is comprised of 47 PyBDSF components, which had to be manually associated. There is an obvious reduction in overhead in the case where one galaxy must be split into two over the need to associate 47 components. These two examples demonstrate clearly the improved automation of the source-finding process in *SKA*-like data with ContinUNet.

Figure 21 shows the Euclidean normalised source counts for sources detected in the MIGHTEE image by ContinUNet, the PyBDSF component catalogue (Hale et al. 2022) and the manually cross-matched source catalogue (Whittam et al. 2023). The ContinUNet source counts sit below those of the component catalogue, which we expect as we have already seen ContinUNet correctly associate source components which are unassociated in the PyBDSF catalogue. This is because ContinUNet is able to associate components with no matching required by eye, as the model has learnt

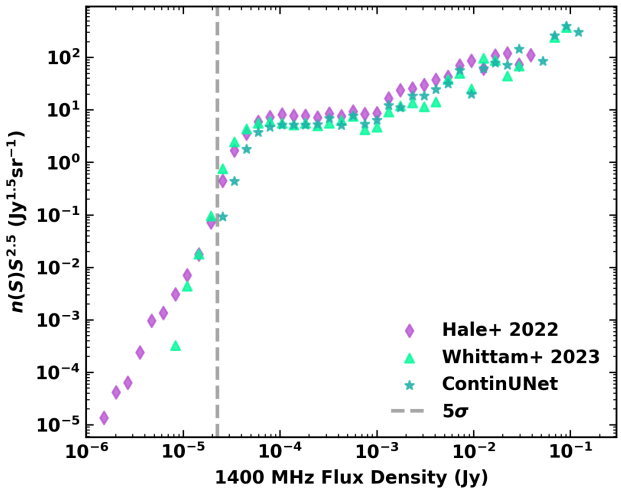

Figure 21. Euclidean normalised source counts for MIGHTEE continuum low resolution image at 1.4 GHz, for ContinUNet detected sources, PyBDSF components (Hale et al. 2022) and cross matched sources (Whittam et al. 2023). For the ContinUNet detected sources, we include only predicted sources detected at a flux density above 5σ where σ is the confusion limited noise, since meaningful science cannot be inferred from sources detected below the confusion limit.

to segment radio sources and associate components simultaneously. This is demonstrated further in Fig. 21, where the brightest end of the source distribution is not accounted for by the PyBDSF component catalogue, but is by the cross-matched source catalogue. These brightest sources typically contain more than one component, which will be represented lower down on the flux density scale than the flux contribution from the entire source. The source counts from ContinUNet more closely match those of the cross-matched catalogue, particularly at high flux densities where components were handled automatically.

In addition to the demonstrated improved automation for source-finding in MIGHTEE, we have also shown the method to be extremely fast. Table 5.1 shows that ContinUNet performs source-finding on the MIGHTEE low resolution image, with end-to-end processing taking <13 s and can be run on a personal machine. This is because inference is fast in comparison to the training, which only has to be performed once and carries the majority of the computational complexity of this method. We have not performed benchmarking at this stage on the MIGHTEE data, but we know that PyBDSF fits Gaussian components to each source, which is a computationally complex and expensive method. We also know that ProFound has memory limitations, and it cannot be used for source-finding in the MIGHTEE image of an area of 1.62 deg^2 on a machine with 16GB RAM, posing a significant draw back for *SKA* data whose images will be ≈30,000 times this size.

Performing source-finding in MIGHTEE with ContinUNet showed that our method is able to associate source components based on the intrinsic source segmentation learned in the latent space of the network, and is able to do so without the need for retraining or substantial parameter tuning. The end-to-end source-finding is fast and scalable and requires significantly less manual intervention than one of the current state-of-the-art methods.

7 CONCLUSIONS

We investigated a new ML driven source-finding method, ContinUNet, in preparation for the data produced by the next generation of radio interferometers such as the *SKA* and compared its performance to the state-of-the-art source-finding methods PyBDSF and ProFound. ContinUNet was trained and tested on the simulated SDC1 radio continuum data. We performed rigorous testing of all methods on the SDC1 data set, and presented our results for the 1.4 GHz 1000 hour image. ContinUNet performed comparably to the state-of-the-art on all metrics in the simulated test data. ContinUNet transferred to real data with no retraining required, and the only tuning performed was to change the thresholding method. When performing source-finding on real data, ContinUNet handled the increased dynamic range of the MIGHTEE data robustly, and demonstrated an ability to associate components of extended sources with no post-processing required. Such sources have been split into multiple components by PyBDSF and are non trivial to associate through manual intervention. Perhaps the most promising outcome is the computational speed and scaling potential; ContinUNet can process an image of one square degree in less than 13 s when run on a supercomputer.

Improvements to the model will be made by improving the segmentation maps used for training and increasing the training set size. However, this work serves as a proof of concept that U-Net can be used for source-finding in *SKA* like data, and that ContinUNet is a promising method that can and should be used for source-finding in precursor data sets such as the MIGHTEE survey.

Currently the model only performs source-finding and no classification of the predicted sources. In the future we would like to add interpretability to the latent space of the model in order to extract source classification as part of the source-finding method. These classifications could include distinctions between FR I and FR II galaxies, compact and extended sources or galaxy morphology such as lopsided or bent tail, depending on the science context of the application of ContinUNet.

Having the ability to perform extremely fast source-finding to a comparative if not improved standard to the state-of-the-art methods, whilst considering computational limitations, that can also reduce the expert labour required in post-processing is particularly important for the upcoming *SKA* science goals. ContinUNet has been designed to perform source-finding out-of-the-box with no parameter tuning. The positive outcomes of improved speed and automation of source-finding in large and complex data sets will prove invaluable as we move into the *SKA* era.

ACKNOWLEDGEMENTS

This paper is dedicated to Professor Mark Birkinshaw, who sadly passed away before it was completed. The author's would like to thank him for his support and guidance and recognise his invaluable contribution to this work. The author's would also like to thank Rhys Shaw, Johannes Allotey and Matt Selwood for their helpful contributions and feedback. This research was supported by the Scientific Machine Learning group at the Rutherford Appleton Laboratory, funded through the Science and Technology Facilities Council. This work was supported by the UKRI Centre for Doctoral Training in Artificial Intelligence, Machine Learning & Advanced Computing, funded by grant EP/S023992/1. This research made use of the *SKA* Science Data Challenge 1 dataset, produced by the SKAO. The MIGHTEE Early Science data, produced by the *MeerKAT* radio telescope was used along with radio source and component catalogues made available by the MIGHTEE collaboration.

DATA AVAILABILITY

The *SKA* Science Data Challenge data is available at: https://www.skao.int/en/science-users/160/skao-data-challenges. The MIGHTEE Early Science data can be found at: https://archive-gw-1.kat.ac.za/public/repository/10.48479/emmd-kf31/index.html. The ContinUNet source-finding package is available at: https://github.com/hstewart93/continunet.

REFERENCES

Bianco, M., Giri, S. K., Prelogović, D., Chen, T., Mertens, F. G., Tolley, E., Mesinger, A., & Kneib, J.-P., 2023. Deep learning approach for identification of hhh regions during reionization in 21-cm observations-ii. foreground contamination, *MNRAS*, **000**, 1–16.

Bonaldi, A. & Braun, R., 2018. Square kilometre array science data challenge 1.

Bonaldi, A., An, T., Bruggen, M., Burkutean, S., Coelho, B., Goodarzi, H., Hartley, P., Sandhu, P. K., Wu, C., Yu, L., Haghighi, M. H. Z., Anton, S., Bagheri, Z., Barbosa, D., Barraca, J. P., Bartashevich, D., Bergano, M., Bonato, M., Brand, J., de Gasperin, F., Giannetti, A., Dodson, R., Jain, P., Jaiswal, S., Lao, B., Liu, B., Liuzzo, E., Lu, Y., Lukic, V., Maia, D., Marchili, N., Massardi, M., Mohan, P., Morgado, J. B., Panwar, M., Prabhakar, Ribeiro, V. A. R. M., Rygl, K. L. J., Ali, V. S., Saremi, E., Schisano, E., Sheikhnezami, S., Sadr, A. V., Wong, A., & Wong, O. I., 2020. Square kilometre array science data challenge 1: analysis and results.

Davies, L. J., Robotham, A. S., Driver, S. P., Lagos, C. P., Cortese, L., Mannering, E., Foster, C., Lidman, C., Hashemizadeh, A., Koushan, S., O'Toole, S., Baldry, I. K., Bilicki, M., Bland-Hawthorn, J., Bremer, M. N., Brown, M. J., Bryant, J. J., Catinella, B., Croom, S. M., Grootes, M. W., Holwerda, B. W., Jarvis, M. J., Maddox, N., Meyer, M., Moffett, A. J., Phillipps, S., Taylor, E. N., Windhorst, R. A., & Wolf, C., 2018. Deep extragalactic visible legacy survey (devils): Motivation, design, and target catalogue, *Monthly Notices of the Royal Astronomical Society*, **480**, 768–799.

Delhaize, J., Heywood, I., Prescott, M., Jarvis, M. J., Delvecchio, I., Whittam, I. H., White, S. V., Hardcastle, M. J., Hale, C. L., Afonso, J., Ao, Y., Brienza, M., Brüggen, M., Collier, J. D., Daddi, E., Glowacki, M., Maddox, N., Morabito, L. K., Prandoni, I., Randriamanakoto, Z., Sekhar, S., An, F., Adams, N. J., Blyth, S., Bowler, R. A., Leeuw, L., Marchetti, L., Randriamampandry, S. M., Thorat, K., Seymour, N., Smirnov, O., Taylor, A. R., Tasse, C., & Vaccari, M., 2021. Mightee: Are giant radio galaxies more common than we thought?, *Monthly Notices of the Royal Astronomical Society*, **501**, 3833–3845.

Fanaroff, B. L., Riley, J. M., Fanaroff, B. L., & Riley, J. M., 1974. The morphology of extragalactic radio sources of high and low luminosity, *MNRAS*, **167**, 31P–36P.

Gavrikov, P., 2020. visualkeras, https://github.com/paulgavrikov/visualkeras.

Goodfellow, I., Bengio, Y., & Courville, A., 2016. *Deep Learning*, MIT Press, http://www.deeplearningbook.org.

Gupta, N. & Reichardt, C. L., 2020. Mass estimation of galaxy clusters with deep learning i: Sunyaev-zel'dovich effect.

Hale, C. L., Robotham, A. S. G., Davies, L. J. M., Jarvis, M. J., Driver, S. P., & Heywood, I., 2019. Radio source extraction with profound.

Hale, C. L., Whittam, I. H., Jarvis, M. J., Best, P. N., Thomas, N. L., Heywood, I., Prescott, M., Adams, N., Afonso, J., An, F., Bowler, R. A. A., Collier, J. D., Cook, R. H. W., Davé, R., Frank, B. S., Glowacki, M., Hatfield, P. W., Lovell, S. K. C. C., Maddox, N., Marchetti, L., Morabito, L. K., Murphy, E., Prandoni, I., Randriamanakoto, Z., & Taylor, A. R., 2022. Mightee: Deep 1.4 ghz source counts and the sky temperature contribution of star

forming galaxies and active galactic nuclei, *Monthly Notices of the Royal Astronomical Society*, **520**, 2668–2691.

Harwood, J. J., Hardcastle, M. J., Morganti, R., Morganti, R., Croston, J. H., Bruggen, M., Brunetti, G., Rottgering, H. J., Shulevski, A., White, G. J., & White, G. J., 2017. Fr ii radio galaxies at low frequencies ii. spectral ageing and source dynamics, *Monthly Notices of the Royal Astronomical Society*, **469**, 639–655.

Heywood, I., Jarvis, M. J., Hale, C. L., Whittam, I. H., Bester, H. L., Hugo, B., Kenyon, J. S., Prescott, M., Smirnov, O. M., Tasse, C., Afonso, J. M., Best, P. N., Collier, J. D., Deane, R. P., Frank, B. S., Hardcastle, M. J., Knowles, K., Maddox, N., & Murphy, E. J., 2021. Mightee: total intensity radio continuum imaging and the cosmos / xmm-lss early science fields, *MNRAS in press*, **20**, 1–21.

Hotan, A. W., Bunton, J. D., Chippendale, A. P., Whiting, M., Tuthill, J., Moss, V. A., Mcconnell, D., Amy, S. W., Huynh, M. T., Allison, J. R., Anderso, C. S., & Bannister, K. W., 2021. Australian square kilometre array pathfinder: I. system description, **38**.

Jarvis, M. J., Seymour, N., Afonso, J., Best, P., Beswick, R., Heywood, I., Huynh, M., Murphy, E., Prandoni, I., Schinnerer, E., Simpson, C., Vaccari, M., & White, S., 2014. The star-formation history of the universe with the ska, *Proceedings of Science*, **9-13-June-2014**, 68.

Jarvis, M. J., Taylor, A. R., Agudo, I., Allison, J. R., Deane, R. P., Frank, B., Gupta, N., Heywood, I., Maddox, N., McAlpine, K., Santos, M. G., Scaife, A. M., Vaccari, M., Zwart, J. T., Adams, E., Bacon, D. J., Baker, A. J., Bassett, B. A., Best, P. N., Beswick, R., Blyth, S., Brown, M. L., Brüggen, M., Cluver, M., Colafranceso, S., Cotter, G., Cress, C., Davé, R., Ferrari, C., Hardcastle, M. J., Hale, C., Harrison, I., Hatfield, P. W., Klöckner, H. R., Kolwa, S., Malefahlo, E., Marubini, T., Mauch, T., Moodley, K., Morganti, R., Norris, R., Peters, J. A., Prandoni, I., Prescott, M., Oliver, S., Oozeer, N., Röttgering, H. J., Seymour, N., Simpson, C., Smirnov, O., Smith, D. J., Spekkens, K., Stil, J., Tasse, C., van der Heyden, K., Whittam, I. H., & WIlliams, W. L., 2016. The meerkat international ghz tiered extragalactic exploration (mightee) survey, *Proceedings of Science*, p. 6.

Jonas, J. L., the MeerKAT Team, Africa, S., & Jonas, M. L. J., 2018. The meerkat radio telescope, *Proceedings of Science*, **277**, 001.

Kingma, D. P. & Welling, M., 2013. Auto-encoding variational bayes.

Lecun, Y., Bottou, L. E., Bengio, Y., & Abstract|, P. H., 1998. Gradient-based learning applied to document recognition.

Lukic, V., Gasperin, F. D., & Brüggen, M., 2019. Convosource: Radio-astronomical source-finding with convolutional neural networks.

Makinen, T. L., Lancaster, L., Villaescusa-Navarro, F., Melchior, P., Ho, S., Perreault-Levasseur, L., & Spergel, D. N., 2020. deep21: a deep learning method for 21cm foreground removal.

Mandal, S., Prandoni, I., Hardcastle, M. J., Shimwell, T. W., Intema, H. T., Tasse, C., van Weeren, R. J., Algera, H., Emig, K. L., Röttgering, H. J. A., Schwarz, D. J., Siewert, T. M., Best, P. N., Bonato, M., Bondi, M., Jarvis, M. J., Kondapally, R., Leslie, S. K., Mahatma, V. H., Sabater, J., Retana-Montenegro, E., & Williams, W. L., 2020. Extremely deep 150 mhz source counts from the lotss deep fields.

Meissen, F., Paetzold, J., Kaissis, G., & Rueckert, D., 2022. Unsupervised anomaly localization with structural feature-autoencoders.

Mohan, N., Rafferty, D., Mohan, N., & Rafferty, D., 2015. Pybdsf: Python blob detection and source finder, *ascl*, p. ascl:1502.007.

Norris, R. P., Hopkins, A. M., Afonso, J. C., Brown, S. A., Condon, J. J., Dunne, L. E., Feain, I. A., Hollow, R. A., Jarvis, M. F., Johnston-Hollitt, M. G., Lenc, E. A., Middelberg, E. H., Padovani, P. I., Prandoni, I. J., Rudnick, L. K., Seymour, N. L., Umana, G. M., Andernach, H. N., Alexander, D. M., Appleton, P. N., Bacon, D. P., Banfield, J. A., Becker, W. Q., Brown, M. J. I., Ciliegi, P. S., Jackson, C. A., Eales, S. T., Edge, A. C., Gaensler, B. M., Giovannini, G. J., Hales, C. A., Hancock, P. V., Huynh, M. T., Ibar, E. X., Ivison, R. J., Kennicutt, R. Z., Kimball, A. E., Koekemoer, A. M., Koribalski, B. S., opez S anchez, A. R. L., Mao, M. Y., Murphy, T. V., Messias, H. C., Pimbblet, K. A., Raccanelli, A. P., Randall, K. E., Reiprich, T. H., Roseboom, I. G., Saikia, D. J., Sharp, R. G., Slee, O. B., Smail, I. U., Thompson, M. A., Urquhart, J. S., Wall, J. V., & b P Zhao, G., 2011. Emu: Evolutionary map of the universe, *Publications of the Astronomical Society of Australia*, **28**, 215–248.

Otsu, N., 1979. Threshold selection method from gray-level histograms., *IEEE Trans Syst Man Cybern*, **SMC-9**, 62–66.

Riggi, S., Magro, D., Sortino, R., Marco, A. D., Bordiu, C., Cecconello, T., Hopkins, A. M., Marvil, J., Umana, G., Sciacca, E., Vitello, F., Bufano, F., Ingallinera, A., Fiameni, G., Spampinato, C., & Adami, K. Z., 2023. Astronomical source detection in radio continuum maps with deep neural networks, *Astronomy and Computing*, **42**, 100682.

Robotham, A. S. G., Davies, L. J. M., Driver, S. P., Koushan, S., Taranu, D. S., Casura, S., & Liske, J., 2018. Profound: Source extraction and application to modern survey data.

Ronneberger, O., Fischer, P., & Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation.

Sadr, A. V., Vos, E. E., Bassett, B. A., Hosenie, Z., Oozeer, N., & Lochner, M., 2018. Deepsource: Point source detection using deep learning, *MNRAS*, **000**, 1–15.

Shimwell, T. W., Röttgering, H. J., Best, P. N., Williams, W. L., Dijkema, T. J., Gasperin, F. D., Hardcastle, M. J., Heald, G. H., Hoang, D. N., Horneffer, A., Intema, H., Mahony, E. K., Mandal, S., Mechev, A. P., Morabito, L., Oonk, J. B., Rafferty, D., Retana-Montenegro, E., Sabater, J., Tasse, C., Weeren, R. J. V., Brüggen, M., Brunetti, G., Chy, K. T., Conway, J. E., Haverkorn, M., Jackson, N., Jarvis, M. J., McKean, J. P., Miley, G. K., Morganti, R., White, G. J., Wise, M. W., Bemmel, I. M. V., Beck, R., Brienza, M., Bonafede, A., Rivera, G. C., Cassano, R., Clarke, A. O., Cseh, D., Deller, A., Drabent, A., Driel, W. V., Engels, D., Falcke, H., Ferrari, C., Fröhlich, S., Garrett, M. A., Harwood, J. J., Heesen, V., Hoeft, M., Horellou, C., Israel, F. P., Kapińska, A. D., Kunert-Bajraszewska, M., McKay, D. J., Mohan, N. R., Orrú, E., Pizzo, R. F., Prandoni, I., Schwarz, D. J., Shulevski, A., Sipior, M., Smith, D. J., Sridhar, S. S., Steinmetz, M., Stroe, A., Varenius, E., Werf, P. P. V. D., Zensus, J. A., & Zwart, J. T., 2017. The lofar two-metre sky survey: I. survey description and preliminary data release, *Astronomy and Astrophysics*, **598**, A104.

Smith, D. J., Haskell, P., Gürkan, G., Best, P. N., Hardcastle, M. J., Kondapally, R., Williams, W., Duncan, K. J., Cochrane, R. K., McCheyne, I., Röttgering, H. J., Sabater, J., Shimwell, T. W., Tasse, C., Bonato, M., Bondi, M., Jarvis, M. J., Leslie, S. K., Prandoni, I., & Wang, L., 2021. The lofar two-metre sky survey deep fields: The star-formation rate-radio luminosity relation at low frequencies, *Astronomy and Astrophysics*, **648**, A6.

Smolcic, V., Padovani, P., Delhaize, J., Prandoni, I., Seymour, N., Jarvis, M., Afonso, J., Magliocchett, M., Huynh, I. M., Vaccari, M., & Karim, A., 2014. Exploring agn activity over cosmic time with the ska, *Proceedings of Science*, **9-13-June-2014**, 69.

Sortino, R., Magro, D., Fiameni, G., Sciacca, E., Riggi, S., Demarco, A., Spampinato, C., Hopkins, A. M., Bufano, F., Schillirò, F., Bordiu, C., & Pino, C., 2023. Radio astronomical images object detection and segmentation: A benchmark on deep learning methods.

van Haarlem, M. P., Wise, M. W., Gunst, A. W., Heald, G., McKean, J. P., Hessels, J. W. T., de Bruyn, A. G., Nijboer, R., Swinbank, J., Fallows, R., Brentjens, M., Nelles, A., Beck, R., Falcke, H., Fender, R., Hörandel, J., Koopmans, L. V. E., Mann, G., Miley, G., Röttgering, H., Stappers, B. W., Wijers, R. A. M. J., Zaroubi, S., van den Akker, M., Alexov, A., Anderson, J., Anderson, K., van Ardenne, A., Arts, M., Asgekar, A., Avruch, I. M., Batejat, F., Bähren, L., Bell, M. E., Bell, M. R., van Bemmel, I., Bennema, P., Bentum, M. J., Bernardi, G., Best, P., Bîrzan, L., Bonafede, A., Boonstra, A.-J., Braun, R., Bregman, J., Breitling, F., van de Brink, R. H., Broderick, J., Broekema, P. C., Brouw, W. N., Brüggen, M., Butcher, H. R., van Cappellen, W., Ciardi, B., Coenen, T., Conway, J., Coolen, A., Corstanje, A., Damstra, S., Davies, O., Deller, A. T., Dettmar, R.-J., van Diepen, G., Dijkstra, K., Donker, P., Doorduin, A., Dromer, J., Drost, M., van Duin, A., Eislöffel, J., van Enst, J., Ferrari, C., Frieswijk, W., Gankema, H., Garrett, M. A., de Gasperin, F., Gerbers, M., de Geus, E., Grießmeier, J.-M., Grit, T., Gruppen, P., Hamaker, J. P., Hassall, T., Hoeft, M., Holties, H. A., Horneffer, A., van der Horst, A., van Houwelingen, A., Huijgen, A., Iacobelli, M., Intema, H., Jackson, N., Jelic, V., de Jong, A., Juette, E., Kant, D., Karastergiou, A., Koers, A., Kollen, H., Kondratiev, V. I., Kooistra, E., Koopman, Y., Koster, A., Kuniyoshi, M., Kramer, M., Kuper, G., Lambropoulos, P., Law, C., van Leeuwen, J., Lemaitre, J., Loose, M., Maat, P., Macario, G., Markoff,

S., Masters, J., McFadden, R. A., McKay-Bukowski, D., Meijering, H., Meulman, H., Mevius, M., Middelberg, E., Millenaar, R., Miller-Jones, J. C. A., Mohan, R. N., Mol, J. D., Morawietz, J., Morganti, R., Mulcahy, D. D., Mulder, E., Munk, H., Nieuwenhuis, L., van Nieuwpoort, R., Noordam, J. E., Norden, M., Noutsos, A., Offringa, A. R., Olofsson, H., Omar, A., Orrú, E., Overeem, R., Paas, H., Pandey-Pommier, M., Pandey, V. N., Pizzo, R., Polatidis, A., Rafferty, D., Rawlings, S., Reich, W., de Reijer, J.-P., Reitsma, J., Renting, G. A., Riemers, P., Rol, E., Romein, J. W., Roosjen, J., Ruiter, M., Scaife, A., van der Schaaf, K., Scheers, B., Schellart, P., Schoenmakers, A., Schoonderbeek, G., Serylak, M., Shulevski, A., Sluman, J., Smirnov, O., Sobey, C., Spreeuw, H., Steinmetz, M., Sterks, C. G. M., Stiepel, H.-J., Stuurwold, K., Tagger, M., Tang, Y., Tasse, C., Thomas, I., Thoudam, S., Toribio, M. C., van der Tol, B., Usov, O., van Veelen, M., van der Veen, A.-J., ter Veen, S., Verbiest, J. P. W., Vermeulen, R., Vermaas, N., Vocks, C., Vogt, C., de Vos, M., van der Wal, E., van Weeren, R., Weggemans, H., Weltevrede, P., White, S., Wijnholds, S. J., Wilhelmsson, T., Wucknitz, O., Yatawatta, S., Zarka, P., Zensus, A., & van Zwieten, J., 2013. Lofar: The low-frequency array, *Astronomy Astrophysics*, **556**, A2.

Wang, M., Li, J., Li, Z., Luo, C., Chen, B., Xia, S.-T., & Wang, Z., 2023. Unsupervised anomaly detection with local-sensitive vqvae and global-sensitive transformers.

Whittam, I. H., Prescott, M., Hale, C. L., Jarvis, M. . J., Heywood, I., An, F., Glowacki, M., Maddox, N., Marchetti, L., Morabito, L. K., Adams, N. J., Bowler, R. A. A., Hatfield, P. W., Varadaraj, R. G., Collier, J., Frank, B., Taylor, A. R., Santos, M. G., Vaccari, M., Afonso, J., Ao, Y., Delhaize, J., Knowles, K., Kolwa, S., Randriamampandry, S. M., Randriamanakoto, Z., Smirnov, O., Smith, D. J. B., & White, S. V., 2023. Mightee: multi-wavelength counterparts in the cosmos field.

Yasutomi, S. & Tanaka, T., 2023. Style feature extraction using contrastive conditioned variational autoencoders with mutual information constraints.

Zack, G. W., Rogers, W. E., & Latt, S. A., 1977. Automatic measurement of sister chromatid exchange frequency, *Journal of Histochemistry and Cytochemistry*, **25**, 741–753.

Zhou, X., Gong, Y., Deng, F., Zhang, M., Yue, B., & Chen, X., 2022. Foreground removal of co intensity mapping using deep learning.

This paper has been typeset from a TEX/LATEX file prepared by the author.