Identifying extra-terrestrial intelligence using machine learning

Małgorzata Gutowska[1], Michael Scriney[2], and Andrew McCarren[3]

[1] Dublin City University, Dublin 9, Ireland
`malgorzata.gutowska2@mail.dcu.ie`
[2] Insight Centre for Data Analytics, Dublin City University, Dublin 9, Ireland
`michael.scriney@insight-centre.org`
[3] VistaMilk Research Centre, School of Computing, Dublin City University, Ireland
`andrew.mccarren@dcu.ie`

Abstract. Since the date of establishment of the SETI Institute, its scientists have used various approaches in their search for extra-terrestrial intelligence (SETI). A novel idea involved image categorisation techniques in classifying radio signals represented by 2D spectrograms. The dataset of simulated radio signals, created for classification purposes have been used in this work to train models based on neural network architectures. It is shown in this paper that combining three different models, trained on features obtained by various techniques, has a positive impact on model accuracy and performance. Features learned by a convolutional neural network (CNN), bottleneck features from existing models and manually extracted features from the spectrograms comprised the three feature sets used as training data for the combined model. It was also shown that combining different methods of spectrogram generation resulted in improving the accuracy of the final model.

Keywords: Convolutional Neural Network · Image Processing · Spectrograms.

1 Introduction

Scientists from the SETI Institute (Search for Extra-terrestrial Intelligence) monitor radio signals coming from multiple directions from space, searching for signs of extra-terrestrial intelligence [20]. The artificial signals are thought likely to be narrow-band, in contrast to natural radiation featuring wide spectra. Therefore, the efforts to identify artificially generated electromagnetic waves are focused on searching for narrow-band signals. The Allen Telescope Array (ATA) is an array of 42 radio telescopes in northern California [24], one of whose main activities is the search for ETI. This telescope allows for simultaneous observations of 3 very small windows of the sky, which helps in highlighting potential signs of intelligent origin [5]. Narrow band signals of frequency drifting linearly over time highlight candidates of potential artificial origin [5]. The sensitive algorithm within the specialised software at ATA detects various types of narrow-band signals. While

the rate of not detecting potentially interesting signals is acceptably low, many other similar signal types produce false positives, resulting in additional valueless sets of observations [5]. Transformed into a form of 2D spectrograms, narrow-band radio signals with a drifting frequency can be easily distinguished from other signal types by a human due to their shape. This characteristic can be utilised in an automated fashion with an adoption of rapidly developing image recognition techniques such as convolutional neural networks (CNN). SETI scientists simulated 140,000 signals and used these as training data for image classification algorithms. A competition, which was held in July 2017, produced a range of classification models, of which the best achieved results of 95% accuracy [6]. A subset of the simulated and labelled data (35,000 signals) was made available after the competition was over [7]. The competition winners used high powered computing provided by IBM [5]. The purpose of this research is to find out whether a satisfactory classification of SETI signals can be achieved using convolutional neural networks and other machine learning techniques on limited resources accessible to citizen scientists. The aim of this work was to reproduce or improve the results of top scoring models using Colaboratory platform with GPU processing and 18 GB RAM.

2   Background

The ATA radio telescopes record analogue voltage signals representing electromagnetic field oscillations reaching individual antennas. Before the signals are stored, they are demodulated down from GHz range to lower frequencies and digitised, which results in a stream of complex-valued time-series data. The digitisation process produces 104 million complex samples per second, corresponding to about 70MHz bandwidth window around central frequency from the range 1-10 GHz, which is the observation range for ATA [5].

2.1   Spectrograms

Usage of frequency domain for signals oscillating in time, in many use cases gives a more meaningful illustration of signal's characteristics than a time domain representation. A transition from time to frequency domain can be achieved by Fourier Transform (FT), based on the assumption that each periodic signal can be decomposed to a sum of sines and cosines:

$$S(t) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos n\omega t + \sum_{n=1}^{\infty} b_n \sin n\omega t \tag{1}$$

or, if represented in a complex space, it can be expressed as

$$S(t) = \sum_{-\infty}^{\infty} c_n e^{jn\omega t} \tag{2}$$

where $S(t)$ is the approximation of signal $s(t)$ in the time domain and ω is signal frequency. The coefficients a_n and b_n or c_n describe a signal in the frequency domain and they are expressed as:

$$a_n = \frac{2}{T} \int s(t) \cos n\omega t\, dt, b_n = \frac{2}{T} \int s(t) \sin n\omega t\, dt \tag{3}$$

and

$$c_n = \frac{1}{T} \int s(t) e^{-jn\omega t}\, dt \tag{4}$$

where T is an oscillation period. The link between sine, cosine and exponential functions is provided by the Euler formulas [15]. The result from the FT is the relationship between a signal's power and its frequency. By including additional time dependency and representing the power by pixels brightness, it is possible to produce 2D spectrograms, which are the types of images classified in this work.

2.2 Observed signals

Radio signals carrying a large amount of energy at a narrow frequency range are often called "narrowband", as they feature narrow bandwidth. In the observed signals, the carrier (central) frequency usually changes with time. This frequency drift (Doppler's drift) is caused by the Earth's rotation and/or the movement or acceleration of the source. These signals are of the greatest interest to SETI scientists. Narrow-band signals with a constant frequency (without drift) are of less interest, because they are more likely to come either from terrestrial broadcasting stations or geostationary satellites [19]. SonATA, the software managing the signals recording and pre-processing at ATA, is also responsible for determining the candidates for intelligent signals. A signal is flagged as a candidate, if it demonstrates characteristics of a narrow-band signal; in such cases follow-up observations are conducted in order to assess the possibility that the signal is of extra-terrestrial origin. The relatively high rate of false positives, which is a weak point of the currently used detection algorithm of SonATA, can easily lead to wasting large amounts of observation time [5].

2.3 Signals simulation

A need for a multi-categorical classifier brought an idea of training a neural-network-based model on spectrogram images. Simulations of radio signals provide a controlled dataset of labelled images. SETI scientists, based on their domain expertise, have simulated 6 categories, which reflect the most commonly observed signal groups, plus an additional category of background noise. Two examples of spectrograms of the simulated signals and their labels are presented in Figure 1. Pictures of all categories are shown in [9]. The full list contained seven categories: `narrowband`, `narrowbanddrd`, `squiggle`, `squarepulsednarrowband`, `squigglesquarepulsednarrowband`, `brightpixel`, `noise`. Simulation process
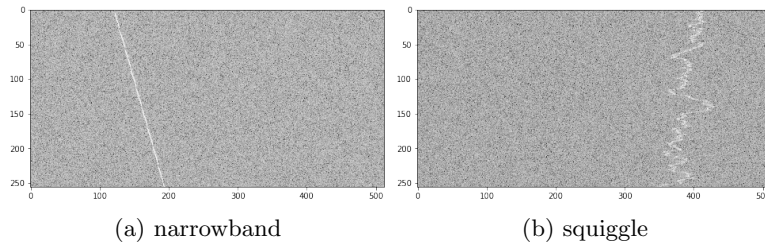
(a) narrowband   (b) squiggle

Fig. 1: Spectrograms representing examples of simulated signals of two categories. The scale used on simulated images is arbitrary and does not reflect any physical units.

and parameters of individual categories have been presented in detail by [5]. One of the goals of the simulation was to generate signals of a wide range of difficulties in distinguishing between individual categories.

2.4   CNN architectures

In this paper we propose a technique that can be used on low performing home computers and give compariable results to the winning competitors. Across recent years, innovative CNN architectures such as VGG, ResNet, Inception, DenseNet or NASNet [18] were each setting a new state-of-the-art classification on datasets such as ImageNet [8] or CIFAR [14]. ResNet's innovation was an introduction of identity connections and optimisation of the residual mapping instead of the original mapping [11]. By developing the residual learning it was possible to train extremely deep networks without gradient vanishing and accuracy degradation. Another innovation – DenseNet – introduced connections between all convolutional layers [12].

3   Methods

The following section describes the methods utilised in this work and the data pre-processing steps, which are as follows:

- obtaining spectrograms from time-series data
- spectrograms enhancement
- exploration of existing Convolutional Neural Networks (CNN)
- creation of manual features
- building the combined architecture

3.1   From time-series data to spectrograms

The simulated data consisted of a series of complex numbers in the time domain. Complex numbers are often chosen for mathematical convenience, when it is required to perform operations on amplitude and frequency.

For example, a signal from a real space written as $x(t) = A_1 \cos \omega t$ in a complex space can be expressed as $x(t) = A_2 e^{j\omega t}$ and then by using the Euler formula (§2), it can be written as a sum of Real and Imaginary components: $x(t) = A_2 \cos \omega t + A_2 j \sin \omega t = a(t) + jb(t)$ The raw data used in this work consisted of numbers of the above pattern.

The following steps were performed to obtain spectrograms from the simulated data:

1. Time-series data has been split into time intervals. The length of the interval would later reflect a time unit on the spectrogram's vertical axis.

2. The time segments have been stacked horizontally to create a 2D array.

3. The array has been processed with Hann window [10]. This technique is used to prevent against spectral leakage and generation of artificial frequencies. This step was performed only in some cases, as described later.

4. Fourier Transform (FT) has been performed on each time segment and a square of its modulus has been taken.

5. A logarithm has been applied on the FT results. This operation brings higher granularity into the lower ranges of a colour histogram. This step was not done in all of the cases, which will be discussed later.

6. The resulting values have been normalised to the range 0-255, for the array to be plotted as an image, where brighter colours correspond to higher signal power.

3.2   Enhancing the spectrograms

The maximum image size attainable from the simulated data was 196,608 pixels, reflecting a resolution of 384x512. However, for the majority of use cases lower resolution images (128x256) have been created.

Monochromatic pictures have only one colour channel, whereas colour pictures can have 2-4 channels; e.g. the RGB colour space uses 3 channels where each colour – red, green and blue are described separately on a scale from 0 to 255. Convolutional networks are therefore well suited to operate on colour images by convolving 3-dimensional filters simultaneously through all colour layers and forming independent weight tensors. Even though spectrograms are monochromatic, in this research separate channels have been used to carry the results of different techniques of image generation. The following methods were explored for this purpose:

- a logarithm has been applied over the FT (Figure 2a),
- the logarithm has been omitted (Figure 2b),
- Hann window has been applied before the FT and the logarithm was taken from the FT result (Figure 2c),
- Radon transform [16] has been applied over the spectrogram obtained as in the first point.

(a) Spectrogram obtained using logarithm over the Fourier Transform output; no Hann window used.

(b) Spectrogram obtained without applying the logarithm; no Hann function used.

(c) Spectrogram obtained using the Hann window before Fourier Transform and the logarithm over the FT output.
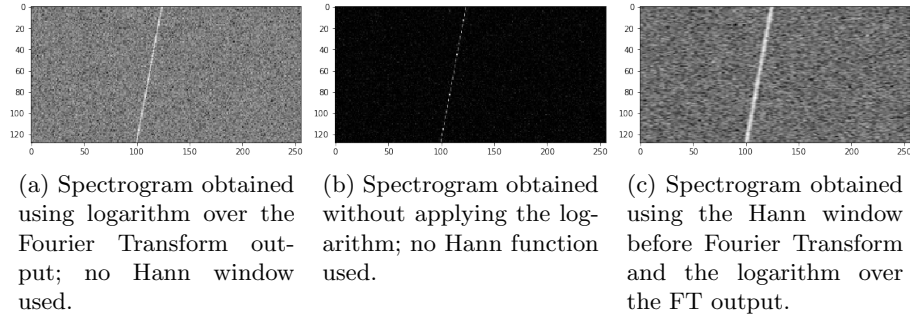
Fig. 2: Spectrogram enhancing. The spectrograms obtained by presented methods have been combined into multichannel images.

While the Radon transform can be useful in detecting straight lines within a noisy background [17], it was eventually omitted in the CNN training, due to lack of sufficient distinction. The spectrograms presented in Figure 2 have been merged into 3-channel pictures and used in CNN's. Another image pre-processing technique which has been explored was a noise reduction technique. Two algorithms were utilised for this purpose; contrast enhancing custom function and the median-based denoising method from OpenCV library [2]. While this process seemed to work well with many spectrograms, it failed in cases where signal to noise ratio was very low. Overall it did not appear to improve the native CNN's pattern detection methods and thus was omitted in further CNN trials. However, it was adopted in the manual feature extraction, which is described later.

3.3   Exploring CNN architectures

Prototyping with CNN's has been performed on 4 categories, with total size of the data set being 4000. Further method exploration was performed on a set of 7000 signals, including all 7 categories, whereas the full data set size used here was 35,000 signals. The CNN architectures, which have been tested in this research include:

- 3 and 4-convolutional-block networks of *Convolutional + ReLU*[4] *+ Max Pooling*[5] layers with variable filter sizes: $3 - 7$,
- VGG16 and VGG19 [21]
- Xception [4]
- ResNeXt101 [25]
- InceptionResNetV2 [22]
- DenseNet201 (201 layers) [12]

---

[4] ReLU is a non-linear activation function
[5] A max value of 2x2 pixels area

Most of the above networks were used through transfer learning, rather than training from random weights. The CNNs of 3 and 4 convolutional blocks, as well as VGG16 have been fully trained from randomly distributed weights, yet the accuracy from the VGG did not outperform the former one.
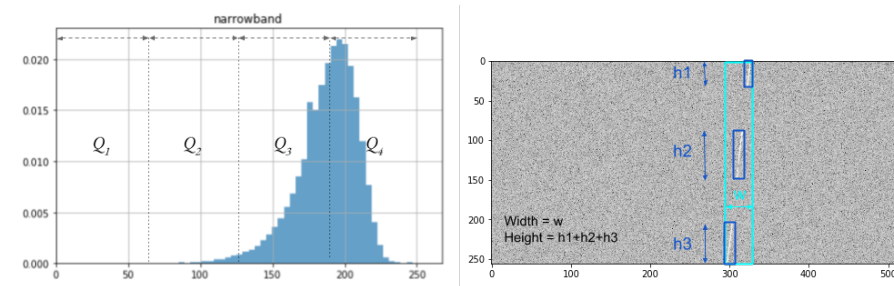
In addition to the above architectures a few variants of combined models were tested, such as DenseNet or ResNeXt accompanied with a neural network of fully connected layers trained on manually crafted features.

3.4   Manually crafted features

Apart from variants of CNN's, a more traditional machine learning method was also part of this research. A regular neural network was trained on a set of features, which were manually crafted to reflect the main characteristics of all classes. The following features were generated:

- the number of pixels in each of the 4 brightness ranges, such as: $0 - 62.5$, $62.5 - 125$, etc., marked as Q_1, $- Q_4$ in Figure 3a,
- characteristics of pixel brightness distribution, such as mean, variance and skewness,
- estimated width of the bounding box around the signal, as shown in Figure 3b; this was obtained using a custom function,
- estimated total height of individual bounding boxes, as in Figure 3b; this was generated using a custom method,
- estimated length of the signal's line using Canny edge detection algorithm [3].

The width and height of a signal line were calculated by counting total number of columns or rows containing pixels of a brightness over a predefined threshold.

(a) Colour (brightness) histogram of a sample spectrogram with annotated four quarters.

(b) Bounding boxes around signal labelled with width and height.

Fig. 3: Illustration of extraction of some of the features.

3.5   Combined architecture

Residual Networks (ResNet) [11] and Densely Connected Convolutional Networks (DenseNet) [12] were shown to perform exceptionally well with the spectrogram images from SETI [6]. In this work, an architecture of the DenseNet (or ResNeXt) model combined with other models was created. The motivation behind combining two or more models was a belief that different types of networks would be able to learn different characteristics of original images, therefore combining them would lead to the "best of each" net result. Both versions of combined architectures – with DenseNet and ResNeXt performed similarly, however, to achieve the same result, DenseNet required much less memory due to a compact architecture featuring less parameters.

Densely Connected Convolutional Network. DenseNet had the highest performance out of all the networks tested [12]. The innovative feature of this architecture is passing the output from all preceding convolution layers to all subsequent layers within a dense block. This way, low level features combined with more complex features are simultaneously passed to the classification layer. This is not without significance for this particular use case, where images contain mostly simple shapes.

Final model implementation. The model developed in this work is shown in Figure 4. Pretrained on the ImageNet dataset, the DenseNet201 model has been used for extracting features out of the training dataset (bottleneck features), by trimming the top part of its architecture and using only the convolutional part. A separate 4-convolutional-block CNN was trained on the spectrogram data from random weights, which produced similarly structured bottleneck features. Two fully connected layers were then added on top of each model and the output was passed onto the common layer. A third element of the whole architecture was a regular network with one hidden and one output layer trained on manually extracted features. The output produced was passed again to the common layer, which was comprised of the outputs of all three individual models. The common layer combined the outputs together, giving each of them the same number of features. Two fully connected layers were added on top of the common layer, with the softmax activation, producing the final output of probabilities for seven categories. In order to achieve the optimal batch size for training, the split of 35,000 pictures into train, validation and test datasets was performed with the ratio of 76:12:12 respectively. The data split have been performed once and then the same subsets were fed into three models. The test data was used at the evaluation stage only, after the training process had completed. During training, images were loaded in batches in order to accommodate available memory resources. In order to overcome overfitting, dropout was applied onto three branches of the combined model.

The gradient descent process was accomplished with the Adam optimiser [13], with the initial learning rate $1e^{-3}$ decaying with subsequent epochs. Training ran through 9 epochs, with each of them taking approximately 43s.
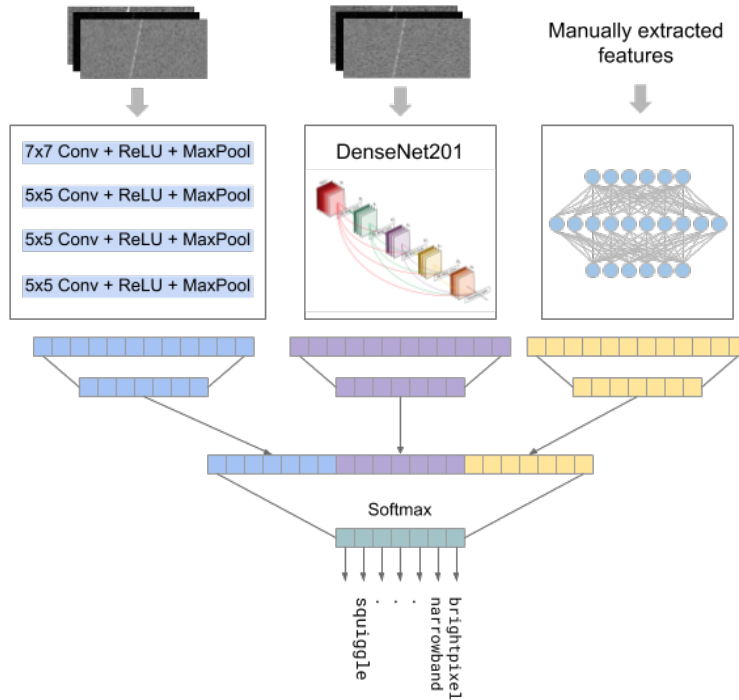
Fig. 4: Combined model architecture. CNN trained on the spectrograms, pre-trained DenseNet [12] and regular network trained on extracted features have been combined together to produce the final output.

4 Results and Discussion

Table 1 compares accuracy values achieved by various methods of this research and Table 2 shows the winning SETI competition results with the best performing method from this work. The performance was measured for both types of images: monochromatic (including only one type of spectrogram) and multi-channel (containing three versions of spectrograms). The triple model achieved the following scores: accuracy – 87.9%, logarithmic loss – 0.3815. Detailed performance scores are shown in Table 3.

From among the single models, the best accuracy has been achieved by the CNN with 4 convolutional blocks, though the combined model outperformed all single model architectures. While other combined architectures were explored, the one combining DenseNet with manual features (MF) and the CNN-trained features yielded the best results. The Table 3 shows the accuracy for all categories. The category with the highest F_1 score was squiggle, which by visual inspection is also the most obvious to recognise. The next best recognised signal is brightpixel. While the precision of brightpixel detection is quite high, the recall is not as good, as many actual brightpixel signals were not distinguish-

Table 1: Performance of models developed in this work (individual and merged).

Model	Images	Dataset	Accuracy
DenseNet	mono	35K Signals	73.96%
DenseNet-MF	mono	35K Signals	82.49%
DenseNet	3 chnl	35K Signals	76.80%
DenseNet-MF	3 chnl	35K Signals	84.48%
4-bl CNN	3 chnl	35K Signals	78.52%
4-bl CNN-DenseNet-MF	3 chnl	35K Signals	87.86%

Table 2: Comparison between the best performing method in this study to competition winners.

Team	Rank	Dataset	Accuracy	LogLoss
Effsubsee [23]	1	140K signals	94.6%	0.1881
Signet [1]	2	140K signals	94.7%	0.2263
Snb1	3	140K signals	87.5%	0.3847
4-bl CNN-DenseNet-MF	3	35K Signals	**87.9%**	**0.3815**

Table 3: Final model performance scores

Category	Precision	Recall	F_1
brightpixel	0.94	0.88	0.91
narrowband	0.84	0.89	0.87
narrowbanddrd	0.87	0.80	0.83
noise	0.82	0.98	0.89
squarepulsednarrowband	0.86	0.82	0.84
squiggle	0.92	0.95	0.93
squigglesquarepulsednarrowband	0.92	0.83	0.87

able from `noise` signals. The `noise` classification has relatively low precision and in addition to `brightpixel`, many other signals were misclassified as `noise`. The lowest confusion with `noise` appears for `squiggle` and `narrowband`. As the correct classification of the `narrowband` category is of the greatest interest to SETI researchers [5], this can be perceived as an advantage of this model. The poorest classification occurs in `narrowbanddrd` signals, which are often confused with `narrowband`. Indeed, in many cases from the `narrowbanddrd` class, the frequency drift was so small that it would be impossible to see a difference when visually inspecting both types of signal. A fact worth noting is that the model presented in this work achieved better overall results in noise detection than both competition winning models, which reported the F_1 score values as: 88.54% and 88.04% [6], significantly lower than the average accuracy of both models: 94.61% and 94.74% respectively. The model presented in this paper achieved 89%, which is above the average accuracy of the model itself. The reason behind this is the inclusion of manually extracted features, that added an ability to separate the noise category relatively well from other categories. Overall, the model presented

in this work did not outperform already existing models, as it would have placed slightly higher than 3rd place. However, it should be noted that the winning team utilised the computational and storage resources provided by IBM for this purpose [5]. The dataset size used in this research was also significantly smaller, as the full dataset generated for this purpose had 140,000 images, while the one used in this work had 35,000. By using the combined model it was possible to achieve the reported accuracy during training taking 6 minutes and 27s. The training ran through 9 epochs and stopped afterwards, when the validation loss stopped decreasing. This is not the total time however, as training of the custom CNN took 40 minutes, similar to feature extraction from DenseNet. The longest time taken was to calculate the manual features for all images, especially, the width and height of the bounding box, which altogether took around 20 hours. The main reason for this process taking so much time was usage of non-optimised custom functions. In summary, given the resources used, rather short training time, and the fact that images of reduced resolution were used in both convolutional networks, a relatively good accuracy was achieved thanks to additional innovations introduced in this work.

5   Conclusions and Future Work

In this work a means of improving classification accuracy (in relation to other examined methods) by both enhancing the source data and consolidating NN architectures was presented. Three methods were integrated by creating spectrograms from raw time-series data into 3 image channels and the inclusion of manually extracted features from the spectrograms. In addition, the consolidation of three neural network architectures into a single model was implemented. By using these methods an accuracy of 87.86% with F_1 scores ranging from 0.83 to 0.93 was obtained. This research shows that including manually extracted features helps in classifying specific categories, such as the SETI category `noise`. Finally, this work showed favourable results in comparison to competitor methods implemented on more advanced architecture environments. Further work in this area could be a promising direction for additional model enhancement. Future work seeks to explore the addition of a wider range of manual features, as well as other denoising techniques.

References

1. Bastian, B.: Ml4seti (2017), https://github.com/sagelywizard/ml4seti?source=post_page
2. Bradski, G.: The OpenCV Library. Dr. Dobb's Journal of Software Tools (2000)
3. Canny, J.: A computational approach to edge detection. In: Fischler, M.A., Firschein, O. (eds.) Readings in Computer Vision, pp. 184 – 203. Morgan Kaufmann, San Francisco (CA) (1987). https://doi.org/10.1016/B978-0-08-051581-6.50024-6
4. Chollet, F.: Xception: Deep learning with depthwise separable convolutions. CoRR **abs/1610.02357** (2016), http://arxiv.org/abs/1610.02357

5. Cox, G.A., Egly, S., Harp, G.R., Richards, J., Vinodababu, S., Voien, J.: Classification of simulated radio signals using wide residual networks for use in the search for extra-terrestrial intelligence. CoRR **abs/1803.08624** (2018), http://arxiv.org/abs/1803.08624

6. Cox, G.A.: Using artificial intelligence to search for extraterrestrial intelligence (2017), https://medium.com/codait/using-artificial-intelligence-to-search-for-extraterrestrial-intelligence-ec19169e01af

7. Cox, G.A.: Welcome to the seti institute code challenge! (2018), https://github.com/setiQuest/ML4SETI/blob/master/tutorials/Step_1_Get_Data.ipynb

8. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: ImageNet: A Large-Scale Hierarchical Image Database. In: CVPR09 (2009)

9. Gutowska, M., McCarren, A.: The search for extra-terrestrial intelligence: classification of radio signals with machine learning. Insight Technical Report **1** (September 2019), https://doras.dcu.ie/23780/1/2019-mcm-gutowsm2-Final_Practicum_Paper.pdf

10. Harris, F.J.: On the use of windows for harmonic analysis with the discrete fourier transform. In: Proc. IEEE. pp. 51–83 (1978)

11. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. CoRR **abs/1512.03385** (2015), http://arxiv.org/abs/1512.03385

12. Huang, G., Liu, Z., Weinberger, K.Q.: Densely connected convolutional networks. CoRR **abs/1608.06993** (2016), http://arxiv.org/abs/1608.06993

13. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)

14. Krizhevsky, A.: Learning multiple layers of features from tiny images. University of Toronto (05 2012)

15. Moskowitz, M.A.: A course in complex analysis in one variable. World Scientific (2002)

16. Murphy, L.M.: Linear feature detection and enhancement in noisy images via the radon transform. Pattern Recognition Letters **4**(4), 279 – 284 (1986). https://doi.org/10.1016/0167-8655(86)90009-7

17. Nikitin, I.: Statistical analysis of narrow-band signals at setilive.org. arXiv e-prints arXiv:1502.04887 (Feb 2015)

18. Pan, S.J., Yang, Q.: A survey on transfer learning. IEEE Transactions on Knowledge and Data Engineering **22**(10), 1345–1359 (Oct 2010). https://doi.org/10.1109/TKDE.2009.191

19. Richards, J.: Seti institute (2016), https://old.seti.org/users/jrichards

20. SETI: Seti institute (2019), https://seti.org

21. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. CoRR **abs/1409.1556** (2014)

22. Szegedy, C., Ioffe, S., Vanhoucke, V.: Inception-v4, inception-resnet and the impact of residual connections on learning. CoRR **abs/1602.07261** (2016), http://arxiv.org/abs/1602.07261

23. Vinodababu, S.: Wide-residual-nets-for-seti (2017), https://github.com/sgrvinod/Wide-Residual-Nets-for-SETI?source=post_page

24. Welch, J., et al.: The Allen Telescope Array: The First Widefield, Panchromatic, Snapshot Radio Camera for Radio Astronomy and SETI. IEEE Proceedings **97**, 1438–1447 (Aug 2009). https://doi.org/10.1109/JPROC.2009.2017103

25. Xie, S., Girshick, R.B., Dollár, P., Tu, Z., He, K.: Aggregated residual transformations for deep neural networks. CoRR **abs/1611.05431** (2016), http://arxiv.org/abs/1611.05431