



Universiteit
Leiden
The Netherlands

Machine learning for radio galaxy morphology analysis

Mostert, R.I.J.

Citation

Mostert, R. I. J. (2024, January 25). *Machine learning for radio galaxy morphology analysis*. Retrieved from <https://hdl.handle.net/1887/3715061>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3715061>

Note: To cite this publication please use the final published version (if applicable).

Machine learning for radio galaxy morphology analysis

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof. dr. ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op donderdag 25 januari 2024
klokke 15.00 uur

door

Rafaël Inayat Jacobus Mostert

geboren te Delft
in 1993

Promotores: Prof. dr. H. J. A. Röttgering
Prof. dr. R. Morganti (ASTRON, Rijksuniversiteit Groningen)

Co-promotor: Dr. K. J. Duncan (University of Edinburgh)

Promotiecommissie: Prof. dr. T. H. W. Bäck
Dr. K. Polsterer (Heidelberg Institute for Theoretical Studies)
Prof. dr. A. Scaife (University of Manchester)
Prof. dr. I. A. G. Snellen
Prof. dr. S. Viti
Prof. dr. M. Wise (SRON Netherlands Institute for Space Research,
Universiteit van Amsterdam)

This dissertation was funded by ASTRON, Leiden Observatory,
and the Leiden Institute of Advanced Computer Science (LIACS).

ASTRON

ASTRON is the Netherlands Institute for Radio Astronomy. ASTRON is part of the institutes organisation of NWO — the Dutch research council.

LOFAR is the Low Frequency Array designed and constructed by ASTRON. It has observing, data processing, and data storage facilities in several countries, which are owned by various parties (each with their own funding sources), and which are collectively operated by the ILT foundation under a joint scientific policy. The ILT resources have benefited from the following recent major funding sources: CNRS-INSU, Observatoire de Paris and Université d'Orléans, France; BMBF, MIWF-NRW, MPG, Germany; Science Foundation Ireland (SFI), Department of Business, Enterprise and Innovation (DBEI), Ireland; NWO, The Netherlands; The Science and Technology Facilities Council, UK; Ministry of Science and Higher Education, Poland; The Istituto Nazionale di Astrofisica (INAF), Italy.

© Rafaël Inayat Jacobus Mostert, 2023

Machine learning for radio galaxy morphology analysis
PhD thesis, Universiteit Leiden, Leiden, Netherlands
ISBN: 978-94-6473-357-0

Cover art: Image created by a stable diffusion neural network hosted by stability.ai based on the prompt 'A field of LOFAR antennas underneath a sky of radio galaxies, painting by Vincent van Gogh'. The underlying stable diffusion model ([Rombach et al., 2021](https://arxiv.org/abs/2010.07421)) is available for use under the following license: <https://huggingface.co/spaces/CompVis/stable-diffusion-license>. In reality, LOFAR antennas are not steerable dishes as depicted. Future generative AI models might correctly paint LOFAR antennas as static.

Cover design by Rafaël Inayat Jacobus Mostert.

Printed by Ipskamp Printing.

An electronic copy of this thesis is available at <https://openaccess.leidenuniv.nl>.

Contents

1	Introduction	I
1.1	Radio interferometry with LOFAR	4
1.2	Radiation mechanisms	5
1.3	Radio-loud active galactic nuclei	6
1.3.1	Morphology	8
1.3.2	Open questions	10
1.4	Computer vision	11
1.5	Machine learning	12
1.5.1	Artificial neural networks	13
1.5.2	Neural networks for object detection	15
1.5.3	Self-Organising Maps	16
1.5.4	Decision trees for explainable predictions	17
1.6	This thesis	17
1.7	Future directions	19
2	Unveiling radio morphology using self-organising maps	21
2.1	Introduction	22
2.2	Data	24
2.3	Method: Rotation invariant self-organised maps	24
2.3.1	Rotation invariant SOM	25
2.3.2	Preprocessing: Creating a training dataset from LoTSS images	28
2.4	Results	31
2.4.1	Initial SOM training	32
2.4.2	Final 10×10 trained SOM	32
2.4.3	Morphology distribution of LoTSS extended radio sources	34
2.4.4	Discovering morphologically rare sources through outlier score	39
2.5	Discussion	46
2.5.1	Linear invariance and the challenge of radio source association	46
2.5.2	SOM for data-exploration. The LOFAR-PINK Visualisation Tool: an interactive webtool	47
2.5.3	Future work	47
2.6	Summary	48
2.A	First run self-organised map	49
2.B	Sample of 100 morphological outliers	49
2.C	Visualisation and exploration tool	55
3	Radio source-component association for the LOFAR Two-metre Sky Survey with region-based convolutional neural networks	57
3.1	Introduction	58
3.2	Existing automated radio component association approaches	59
3.3	Data	61

3.3.1	Selection of source components	62
3.3.2	Manual association process	63
3.4	Methods	65
3.4.1	The R-CNN architectures	65
3.4.2	Training and inference phase	67
3.4.3	Pre-processing the images and labels	70
3.4.4	Pre-computed regions	71
3.4.5	Removing unresolved sources	71
3.4.6	Data augmentation through rotation	73
3.5	Results	74
3.5.1	Baseline and upper-boundary performance	75
3.5.2	Classification backbone and learning rate experiments	76
3.5.3	Ablation study	77
3.5.4	Final results	78
3.6	Discussion	82
3.6.1	Prediction score versus catalogue accuracy	83
3.6.2	Scope of usability and limitations	85
3.6.3	Prediction for large and faint radio components	85
3.6.4	Comparison to the public LOFAR Galaxy Zoo	86
3.6.5	Future work	88
3.7	Conclusions	90
3.A	LoTSS-DR2 Zooniverse project interface	90
3.B	Manually corrected associations	91
3.C	Regions with low prediction scores	91
4	Finding AGN remnant candidates with machine learning	95
4.1	Introduction	97
4.2	Data	99
4.3	Methods	100
4.3.1	Data pre-processing	100
4.3.2	Deriving morphological features	105
4.3.3	Training the random forest classifier	109
4.4	Results	111
4.4.1	Resulting morphological features	111
4.4.2	Radio morphology in the trained SOM	111
4.4.3	Resulting hyperparameters for the trained random forest classifier	114
4.4.4	Feature importance	116
4.4.5	Random forest classifier performance	118
4.5	Discussion	123
4.5.1	Feature saliency	123
4.5.2	Estimating the number of ‘AGN remnant candidates’ after model prediction	123
4.5.3	Use and limitations	124
4.5.4	Future prospects	125
4.6	Conclusions	126

4.A	Examples of radio sources for varying concentration index, clumpiness index, and Gini coefficient values	127
4.B	Examples of sources from different Haralick clusters	127
4.C	Derivation of statistical estimates	127
5	Giant radio galaxy constraints from machine learning and Bayesian inference	137
5.1	Introduction	139
5.2	Theory	140
5.2.1	RG total and projected proper lengths	140
5.2.2	GRG projected proper length: general	140
5.2.3	GRG projected proper length: curved power law	142
5.2.4	GRG observed projected proper length	143
5.2.5	GRG number density	146
5.2.6	GRG lobe volume-filling fraction	147
5.2.7	GRG angular lengths	148
5.2.8	Inference	148
5.3	Data	149
5.4	Methods	149
5.4.1	Radio emission detection	151
5.4.2	Calculating radio to optical/infrared likelihood ratios	151
5.4.3	Sorting radio emission with a gradient boosting classifier	152
5.4.4	Associating radio emission into radio sources	152
5.4.5	Optical or infrared host galaxy identification	152
5.4.6	Re-assessing angular source lengths	154
5.4.7	Manual verification of obtained GRG sample	156
5.4.8	Merging our sample with the GRG sample in the literature	157
5.4.9	Bayesian parameter estimation	162
5.5	Results	165
5.5.1	GRG length distribution	165
5.5.2	GRG number density	168
5.5.3	GRG lobe volume-filling fraction	168
5.6	Discussion	170
5.6.1	Comparison with previous ML GRG search techniques	170
5.6.2	Comparison with previous inference strategies	171
5.6.3	Future work	171
5.7	Conclusions	172
5.A	Curved power law PDF for L	175
5.B	Likelihood trick	176
5.C	PyBDSF parameters	176
5.D	Adaptations of the radio ridgeline based host galaxy identification	177
5.E	Sky coverages	179
	English summary	193
	Nederlandstalige samenvatting	197

List of publications	203
Curriculum vitae	207
Acknowledgements	209

See the animal in the cage that you built

Right Where It Belongs, Trent Reznor

