RNAAS RESEARCH NOTES OF THE AAS

A Systematic Comparison of Thresholding Algorithms for Rotational Standardization in Radio Galaxy Classification

To cite this article: Anna Iskra *et al* 2026 *Res. Notes AAS* **10** 138

DRAFT VERSION MAY 24, 2026
Typeset using LATEX **manuscript** style in AASTeX7.0.1

A Systematic Comparison of Thresholding Algorithms for Rotational Standardization in Radio Galaxy Classification

ANNA ISKRA,[1] FIONA LI,[1] AND CINDY FAN[1]

[1]*College of Natural Sciences, The University of Texas at Austin*

ABSTRACT

Morphological classification of radio galaxies with convolutional neural networks (CNNs) requires rotationally diverse training data, achieved either through augmentation or rotational standardization during preprocessing. K. Brand et al. (2023a) showed that rotational standardization can reduce computational overhead, and noted that improvements to thresholding would directly improve standardization performance. To our knowledge, we present the first systematic comparison of thresholding algorithms for this purpose, evaluating watershed segmentation, adaptive mean thresholding, and adaptive Gaussian thresholding against the original binary method of K. Brand et al. (2023a) on 960 radio galaxies across four morphological classes. Adaptive mean thresholding achieved the highest macro F1 among standardization methods (0.917), outperforming the original binary approach (0.837), though no method matched rotational augmentation (0.938). Class-specific failure modes varied across algorithms. We also document a reproducibility gap between our reimplementation and the standardization F1 scores reported by K. Brand et al. (2023a).

1. INTRODUCTION

Recent advancements in radio surveys have dramatically increased the pool of known radio galaxies, making efficient large-scale classification increasingly important. CNNs have become the leading tool for morphological classification, but require rotationally diverse training data. A common way to achieve this diversity is through rotational augmentation, where rotated copies of existing images are added to the training set. While effective, this approach increases dataset size and computational cost. K. Brand et al. (2023a) demonstrated that rotationally standardizing images during preprocessing—aligning all images to a common axis via principal component analysis (PCA) on thresholded galaxy pixels—can reduce this overhead, though standardization achieved slightly lower F1 scores than augmentation. Crucially, they noted that "any improvements in the thresholding algorithm will lead to an improvement in the results of [rotational standardization]" (K. Brand et al. 2023a, 296), as their approach relied on a binary mask dividing normalized images into zones of "galaxy" and "non-galaxy."

We investigate this directly. Using the same 960-galaxy FRGMRC dataset derived from the FIRST Survey (K. Brand et al. 2023a) and publicly released through Zenodo (K. Brand et al. 2023b), which spans four morphological classes (bent, compact, FRI, FRII), we compare four thresholding methods: the original binary algorithm of K. Brand et al. (2023a), watershed segmentation, adaptive mean thresholding, and adaptive Gaussian thresholding. Each thresholded mask is passed to PCA-based rotational standardization, and the resulting preprocessed datasets are used to train identical SCNNs taken directly from K. Brand et al. (2023a)'s codebase.

The dataset is noisy overall, with signal-to-noise ratio (SNR) distributions skewed left across all classes. FRI galaxies are particularly affected, with a median SNR of 11.9–this is consistent with their core-dominant morphology and diffuse, low-surface-brightness jets that place much of their flux near the noise floor. FRII galaxies, by contrast, are edge-brightened with concentrated lobes, yielding a median SNR of 71.2. This intrinsic disparity motivates the search for more sensitive thresholding: an ideal mask should capture faint FRI jets without misclassifying background in brighter FRII images.

2. METHODS

Regardless of the thresholding algorithm applied, we began with the same initial preprocessing step, scaling pixel values to $[0, 1]$ using $x_{ij} = (x_{ij} - \min(X))/(\max(X) - \min(X))$. Binary thresholding (K. Brand et al. 2023a) applies percentile-based intensity cutoffs depending on image noise level, with an intensity variability map fallback for significantly noisy images. Watershed segmentation treats an image as a topographical surface; we place markers at local maxima (via Otsu's method and distance transform) rather than local minima to reduce oversegmentation, then flood outward to form catchment basins. Adaptive mean thresholding calculates the threshold for each pixel as the arithmetic mean of its local neighborhood. Adaptive Gaussian thresholding uses a Gaussian-weighted mean, giving more influence to pixels closer to the center of the neighborhood. After masking, PCA is used to identify the axes of maximum pixel spread and each image is rotated to align with a fixed reference axis. We also train SCNNs on unprocessed data and on a rotationally augmented dataset (60° steps from 0° to 300°, producing six orientations per image) as baselines.

3. RESULTS

Table 1 presents per-class F1 scores and aggregate training statistics. Adaptive mean thresholding produced the strongest rotational standardization results, particularly for bent galaxies (F1 = 0.923) and compact galaxies (F1 = 0.930), achieving the highest macro F1 among standardization methods (0.917). Watershed segmentation performed similarly to mean thresholding on bent galaxies but showed weaker compact and FRI detection. Gaussian thresholding produced results comparable to the original binary method, with modest FRII gains but weaker performance on compact and FRI classes. Notably, none of the thresholding-based standardization methods matched augmentation, which remained the strongest overall strategy and achieved the highest F1 scores for three of four classes. Nevertheless, augmentation required the longest training time by far (4007 s), while standardization methods trained in 538–898 s, confirming the computational motivation for standardization despite the performance tradeoff.

Table 1. Performance of SCNNs using various preprocessing strategies.

Preprocessing	Bent F1	Compact F1	FRI F1	FRII F1	Macro F1	Loss	Time (s)
None	0.837	0.976	0.865	0.930	0.902	0.25	932
Original Thresholding	0.850	0.837	0.750	0.909	0.837	0.55	538
Watershed	0.923	0.933	0.875	0.921	0.913	0.31	898
Mean Thresholding	0.923	0.930	0.875	0.923	0.917	0.32	817
Gaussian Thresholding	0.850	0.844	0.750	0.907	0.838	0.51	734
Augmentation	0.950	0.933	0.909	0.946	0.938	0.18	4007

The classification performance of SCNNs across preprocessing methods provides insight into how morphological structure in radio galaxies interacts with the preprocessing steps. FRII galaxies were the easiest class for all models: their edge-brightened morphology produces concentrated lobes with high signal per pixel, leaving strong emissions well-preserved even when faint emission is removed. FRI galaxies were most vulnerable, frequently misclassified as compact; both classes contain a prominent central core, and FRI galaxies lose much of their extended emission during thresholding, indicating that their diffuse jets are not only faint in observation but also highly sensitive to image masking. All models misclassified some bent galaxies as FRII, likely due to morphological similarity—both classes often present two bright lobes, and bent galaxies whose curved jets are smoothed by thresholding can closely resemble symmetric FRIIs. Gaussian thresholding showed a consistent weakness for compact galaxies: because the Gaussian-weighted neighborhood places substantial emphasis on surrounding dark pixels, local thresholds are underestimated and compact sources are partially or fully removed from the mask.

Our rotationally standardized SCNNs—including the one using Brand et al.'s own thresholding method—performed below the standardization F1 scores reported by K. Brand et al. (2023a). These differences suggest that rotational standardization performance may be sensitive to implementation details in the thresholding and alignment procedure, as well as stochastic effects arising from random initialization and data loading. Such variability is not always explicitly reported in studies of automated radio galaxy classification; as radio surveys scale and automated classification becomes more widely used, careful documentation of preprocessing choices may become increasingly important alongside model architecture.

4. CONCLUSION

Among the thresholding methods we evaluated, adaptive mean thresholding most effectively enhances rotational standardization, narrowing the performance gap with augmentation while retaining its computational advantage. However, no single method consistently dominates across all four galaxy classes, suggesting that future pipelines may benefit from per-class or per-image algorithm selection. Overall, our results provide a benchmark comparison of thresholding algorithms for noisy radio galaxy images and highlight potential sensitivity to implementation details in a widely used preprocessing pipeline.

ACKNOWLEDGMENTS

We thank Dr. Shyamal Mitra and peer mentors Vivek Abraham, Joel Deville, and Sana Kohli for their guidance through the Geometry of Space Stream at the University of Texas at Austin.

DATA AND CODE AVAILABILITY

The FRGMRC dataset used in this work is publicly available on Zenodo at https://doi.org/10.5281/zenodo.7645530. The dataset is derived from FIRST Survey observations and was originally curated by Brand et al. (2023).

Code used for preprocessing, rotational standardization, and CNN training was implemented using the public repository provided by K. Brand et al. (2023a), with additional implementation of thresholding methods developed for this study. The analysis code is available at https://doi.org/10.5281/zenodo.20361774.

REFERENCES

Brand, K., Grobler, T. L., Kleynhans, W., et al. 2023a, Monthly Notices of the Royal Astronomical Society, 522, 292–311, doi: 10.1093/mnras/stad989

Brand, K., et al. 2023b, FRGMRC: FIRST Radio Galaxy Morphology Reference Catalogue, Zenodo, doi: 10.5281/zenodo.7645530