RNAAS RESEARCH NOTES OF THE AAS



# Classifying Galaxy Merger Types via a CNN Approach on Simulation and Observational Data

To cite this article: Hrutvik Palutla *et al* 2026 *Res. Notes AAS* **10** 191

 

# Classifying Galaxy Merger Types via a CNN Approach on Simulation and Observational Data

Hrutvik Palutla, Ecaterina Sur, Karan Gupta

**Abstract**

Galaxy mergers dramatically reshape galaxies and are key drivers of cosmic evolution (Schweizer, 2000). We present a deep-learning approach to classify galaxy mergers by the morphological types of their progenitors. Using a large library of hydrodynamical merger simulations, the GalMer database, we train a model that takes 2D simulated merger snapshots, converts them into projected density maps, and then predicts whether the merging pair is spiral–spiral, spiral–elliptical, or elliptical–elliptical. This model was validated on real Sloan Digital Sky Survey merger images from the Galaxy Zoo merger catalog of Darg et al. (2010), whose visual classifications served as our ground-truth labels. Our approach automates progenitor-type classification for large survey samples, enabling new studies on galaxy evolution and the dependence of remnant merger morphology on progenitor morphologies. Our improved model achieved an overall accuracy of $88.37\%$.

## 1 Introduction

We investigated whether an ML model could accurately classify galaxy merger images according to the morphological types of the galaxies involved, from a single galaxy merger image, by developing a deep learning classification pipeline that takes raw merger images and outputs predicted morphological labels. Our approach involved preprocessing, augmenting the images to account for viewing angle, and experimenting with convolutional neural network (CNN) architectures. After training the model on simulated mergers, we evaluated its performance on real observational data.

Classifying merger remnants by morphologies is possible because theory and observational data suggest that different processes take place when galaxies of different type merge. For example, two merging ellipticals typically produce diffuse, fan-like structures on the outskirts (Schweizer, 1983), while merging disks or spirals tend to have more prominent bridges and tail-like structures (A. Toomre & J. Toomre, 1972).

 

## 2 Data Collection & Preprocessing

We used the GalMer Horizon Database (Chilingarian et al., 2010) to train our machine learning algorithms because it is one of the best and largest public galaxy merger simulation databases. It incorporates variables such as dark matter, mass ratios, disc stability, and orbital inclinations. GalMer file names encode the original galaxy types, so we mapped detailed Hubble labels into three broader classes: spiral–spiral, spiral–elliptical, and elliptical–elliptical.

The GalMer FITS files contain particle information rather than telescope images. To make CNN inputs, we used the particle positions and available stellar or gas mass fields, centered each snapshot, rotated it to one of the $17$ selected viewing angles (see below), projected it onto a two-dimensional image plane, and removed extreme projected outliers with a $98$th-percentile radius cut. The remaining particles were binned into a mass-weighted 2D histogram. We then smoothed the grid with a Gaussian kernel and applied $\log_{10}(H_{\mathrm{smooth}} + 1)$ contrast scaling, with display limits set by the 2nd and $99.8$th percentiles. This produced mass density maps where faint tidal features and bright cores both remained visible.

We generated $17$ fixed orientations for each GalMer snapshot, including the three principal projections and additional azimuth-inclination views sampled across a hemisphere. This expanded about $75{,}000$ snapshots to about $1.278$ million rendered views. Before training, we removed images with widely separated galaxies, an extremely faint component, or no clear interacting pair under automated blob detection. This filtering removed about $20$–$25\%$ of rendered maps.

To test our model, we used data from the Sloan Digital Sky Survey, specifically the Galaxy Zoo merger sample compiled by Darg et al. (2010). Galaxy Zoo volunteers first identified possible mergers, after which Darg and his research team visually checked candidate pairs and assigned morphology labels. Their catalog contained $3003$ merger images, of which three images were blank or unusable after preprocessing. The SDSS image cutouts were retrieved from the SDSS DR$16$ imaging database (Ahumada et al., 2020) using the sky coordinates and object IDs from their catalogue. This test set reflected the class imbalance typical of observed mergers: $2384$ spiral–spiral, $352$ spiral–elliptical, and $264$ elliptical–elliptical mergers.

For these observed SDSS images, we generated processed intensity maps rather than physical mass-density maps. The downloaded cutouts were cropped to $512 \times 512$ pixels at $0.2$ arcsec/pixel, corresponding to a $51.2$ arcsecond crop radius, converted to grayscale, and normalized. We estimated the sky with $3\sigma$-clipped statistics, subtracted the median plus twice the standard deviation, clipped negative pixels to zero, smoothed with a Gaussian kernel of $\sigma = 2$, applied $\log_{10}(I + 1)$ contrast scaling, and kept the dominant connected component above $0.02$ of the image maximum to remove unrelated stars and nearby sources while preserving the main interacting pair for classification without relying on color information from the original cutout.

 

## 3 Model Training

Our first training run used $60{,}000$ $128 \times 128$ maps from GalMer and a strongly imbalanced simulated training set, which produced high apparent overall accuracy ($80.2\%$) but poor recovery of the minority s–e and e–e classes. In practice, the model learned the dominant s–s class too easily and treated many non-s–s examples as variations of that class, especially when the elliptical component was faint or the pair was projected at an unfavorable angle. This showed that overall accuracy alone was not a reliable measure for this problem, since the observational test set was also dominated by s–s systems. The run therefore helped identify three concrete weaknesses in the pipeline: insufficient image resolution for faint tidal features, inadequate representation of minority merger types, and too many weak or poorly separated simulated interactions in the training set.

For the final model, we addressed these weaknesses by increasing image resolution to $512 \times 512$ pixels for both training and testing maps and using a more balanced training sample: $50\%$ s–s mergers, $25\%$ e–e mergers, and $25\%$ s–e mergers. We also filtered GalMer density maps by the approach described in the preprocessing section. The final CNN used five convolutional blocks with $64$, $128$, $256$, $512$, and $1024$ filters, max-pooling between blocks, ReLU activations, and two dense layers with $2048$ and $1024$ neurons before a three-class softmax output. We used dropout layers in the dense part to mitigate overfitting.

## 4 Final Results

In the second experiment, we trained the CNN on a smaller number of images ($N = 20{,}000$) but enforced a more balanced class distribution: $10{,}000$ s–s, $5000$ s–e, and $5000$ e–e training images. After training, we evaluated it on the same SDSS $3000$-image test set (see Figure 1).

The improved model reached an overall accuracy of $88.37\%$. Its precision, recall, and F1-score were $0.91$, $0.99$, and $0.95$ for s–s; $0.98$, $0.23$, and $0.37$ for s–e; and $0.64$, $0.75$, and $0.69$ for e–e. The number of correctly identified spiral–spiral mergers remained high. Precision for s–s improved slightly, while the recall of e–e improved significantly from $5\%$ in the initial model to $75\%$. This shows that balancing the training data, filtering the simulation images, and using higher-resolution density maps helped the CNN better learn features of elliptical–elliptical interactions.

The spiral–elliptical class remained the most difficult. In many cases where the model failed, it confused s–e and e–e mergers. This is likely because mixed mergers are a harder visual task than either same-type class. Spiral–elliptical systems have less symmetry and more variety: depending on merger stage and projection, the spiral disk may be faint, disrupted, or partially hidden, while the elliptical component may appear as a compact companion, a bulge-like feature, or diffuse structure. These cases give the CNN fewer consistent visual cues than s–s or e–e mergers, likely

 

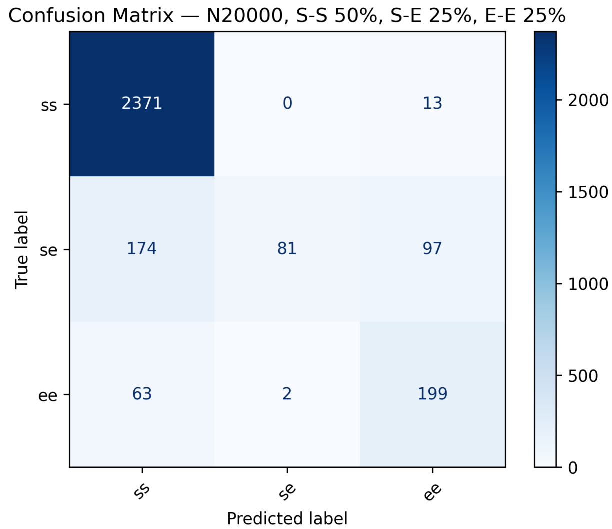


Figure 1: CNN confusion matrix on the $3000$ SDSS mergers. The rows are true labels, and the columns are predicted labels for the mergers.

contributing to the lower accuracy and confidence of its s–e predictions.

Once trained, the CNN can classify large merger samples much faster than visual inspection, making it useful as a scalable follow-up tool for survey data. Because simulations suggest that progenitor type affects later galaxy morphology (Dubois et al., 2016), automated progenitor-type labels can help connect observed merger remnants to broader questions in galaxy evolution.

We would like to thank Dr. Shyamal Mitra, Dr. Gebhardt, our peer mentors, and everyone in the Geometry of Space group at the University of Texas at Austin for providing essential guidance and feedback. Large language model tools were used for wording assistance and source organization; the approach, analysis, interpretation, and conclusions are the authors' own.

# References

Chilingarian, I. V., et al. 2010, A&A, 518, A61, The GalMer database: Galaxy Mergers in the Virtual Observatory

Darg, D. W., et al. 2010, MNRAS, 401, 1043, Galaxy Zoo: the fraction of merging galaxies in the SDSS and their morphologies

Dubois, Y., et al. 2016, MNRAS, 463, 3948, The Horizon-AGN simulation: morphological diversity of galaxies promoted by AGN feedback

Schweizer, F. 1983, in IAU Symposium No. 100, Internal Kinematics and Dynamics of Galaxies, ed. E. Athanassoula (Dordrecht: Reidel), 319, Observational Evidence for Mergers

 

Schweizer, F. 2000, in Phil. Trans. R. Soc. Lond. A, 358, 2063, arXiv:astro-ph/0002263v1, Interactions as a Driver of Galaxy Evolution

Toomre, A., & Toomre, J. 1972, ApJ, 178, 623, Galactic Bridges and Tails

Ahumada, R., et al. 2020, ApJS, 249, 3, The Sixteenth Data Release of the Sloan Digital Sky Surveys: First Release from the APOGEE-2 Southern Survey and Full Release of eBOSS Spectra