Computer Science • 26(SI) 2025 https://doi.org/10.7494/csci.2025.26.SI.7072

Luis E. Suelves
William J. Pearson
Agnieszka Pollo

CHARACTERISTIC SKY-BACKGROUND FEATURES AROUND GALAXY MERGERS

Abstract *In the context of finding galaxy mergers in large-scale surveys, we applied machine-learning algorithms that made use of flux measurements instead of using images (as is the current standard). By training multiple NNs using the Sloan Digital Sky Survey class-balanced data set of mergers and non-mergers, we found that sky-background error parameters could provide a validation accuracy of 92.64 ± 0.15% and a training accuracy of 92.36 ± 0.21%. Moreover, analyzing the NN identifications led us to find that a simple decision diagram using the sky error for two flux filters was enough to gain a 91.59% accuracy. By understanding how the galaxies vary along the diagram and trying to parametrize the methodology in the deeper images of the Hyper Suprime-Cam, we are currently trying to define and generalize this sky error-based methodology.*

Keywords galaxies, evolution, interactions, photometry, data analysis, numerical methods

Citation Computer Science 26(SI) 2025: 109–125

1. Introduction

Galaxies are composed of stars, gas, dust, planets, and many other more exotic astronomical objects; however, they are also the main building blocks of the universe's large-scale structure. They show a wide variety of shapes and sizes that have arisen from the cosmological hierarchical growth of their structures and the combinations of the internal phenomena that have occurred to their components. The hierarchical growth of a structure explains how galaxies have been formed by the subsequent clumping of smaller galaxies into bigger ones [57, 58]. The galactic phenomena that shape their internal structures range from supernovae to the formations of new stars.

Galaxy mergers are a good trace of the hierarchical evolution of the universe; galaxies that approach each other with sufficiently slow relative velocity can merge into one final galaxy. As they approach each other, tidal forces distort their shapes [53]. When they cross each other, their internal structures (which resemble clouds of gas and stars) become even more distorted. The system energy decreases as they interact and finally coalesce into the resulting merged galaxy.

This process results in new galaxies that, in one of the most common cases, transform two spiral galaxies into a bigger elliptical galaxy, thus depleting the gas by forming many stars; these are known as wet mergers [9, 34]. A merger may generally trigger star formations along the process [22,38], even though this has not been found to always be the case [43]. This star formation seems to be triggered by the movement of gas toward the bulges [46] – the central and brighter parts of spiral galaxies. Besides, the time scales of galaxy mergers have been possible to estimate through simulations that have shown how they can last for $\sim$1 billion years [24, 29, 41].

In order to better understand the role that galaxy mergers play in galaxy evolution, it is crucial to be able to find them in large-area sky surveys such as the Sloan Digital Sky Survey (SDSS) [59], the Subaru/Hyper Suprime-Cam surveys, or the upcoming Large Survey of Space and Time (LSST) [21]. The number of galaxies (accounted for as the number of detections in astronomical images) is changing from the order of millions in the current SDSS to billions in the upcoming LSST, which is planning to start collecting data in 2025 and that will continue for 10 years. Many galaxy-merger events can be found among these galaxies, but they must be identified first. Galaxy-merger identification can be done through multiple techniques – the most simple being the visual inspection of galaxies' images [12,42,49]. However, the large number of galaxies makes visual inspections impractical, so it is necessary to have identification methods that can be automatized.

One traditional technique that allows for the faster identification is the cross-pairs method, where galaxies are determined to be merging if they are physically close to each other. This is done by calculating the galaxies' redshifts, which give information of their relative distance by the shift toward the redder colors of the galaxies' spectra [5, 13, 15, 16, 25–27, 39, 40, 45]. Another automatic technique is the calculation of morphological parameters, which estimate the shape deformations of galaxies by measuring the parameters in the images that quantify it. Through the resulting param-

eter space, mergers can be identified with more or less accuracy in characteristic regions [1,2,6,8,10,23,30,36,47,48,52]. A redshift can be measured by finding features in the electromagnetic spectra of galaxies that correspond to known emitted wavelengths; this can be done through the galaxies' spectra or, with less accuracy, through the photometric fluxes across multiple bands. Moreover, the advent of machine-learning (ML) techniques – especially convolutional neural networks (CNNs) – has increased the standards in automatic merger classification [3, 7, 14, 17, 33, 35, 42–44, 54, 56]. The use of each technique strongly depends on the type of data. If the resolution of an image is good enough, morphological parameters and CNNs are more suitable than cross-pair searches (which only require accurate distance estimations). Pre-mergers, or galaxies that are observed before merging into one, are easier to identify by cross-pairs and CNNs; however, post-mergers (the galaxies that have resulted in the mergers) cannot be found by cross-pair methods and are more difficult to identify in general.

In the work that is presented in these pages, we aimed to train a neural network (NN) by making exclusive use of photometric measurements. Photometry is the estimation of the galaxy's flux in the astronomical images that are obtained after careful calibration. The main reason for this has been to reduce the time and computational expenses of the methods that were introduced above, although finding new properties that characterized mergers was also a big motivation. Estimating morphological parameters or training ML models in images can take days when dealing with millions of sources; using direct data products from the surveys themselves could reduce the demand of computing resources. This would be because calculation morphological parameters would not be needed, and the ML models would not need the images but more simple data instead. In a talk that was presented during the 2nd International Workshop on Machine Learning and Quantum Computing Applications in Medicine and Physics (WMLQ 2024), we presented how the NN results helped us to find the potential strength of the SDSS sky-background error (`skyErr` from now on) for finding mergers, the interpretation of this finding, and the technical aspects that we we are taking into account in order to generalize the methodology outside the training data set.

This proceeding begins with introducing the data in Section 2, the methods that were used in Section 3, and the results in Section 4; we finalize with the interpretation and the current status of the methodology extension in Section 5.

2. Data

The SDSS galaxies of our training data set were observed in the SDSS Data Release 6 (DR6) [4]. The telescope's camera had five photometric filters: u, g, r, i, and z [20]; the photometric parameters of the SDSS galaxies were obtained for these five bands. The magnitudes were calculated by an asinh – inverse hyperbolic sine – magnitude function [31] from the measured fluxes. The model fluxes of the galaxies came from a linear combination of the exponential and De Vaucouleurs profiles that were described as in [50].

Besides the model magnitudes, we made use of aperture magnitudes known as fiber magnitudes. These were calculated as the flux within an aperture of 3 arcseconds of diameter around the centers of the galaxies, convolved with a 2-arcsecond seeing. They were defined to resemble the aperture that was observed by the SDSS fibers, which were used for the spectroscopic measurements of the survey.

In order to discern merging and non-merging galaxies, we needed to build a data set that included a representative sample of both. While the real fraction of galaxies that are undergoing merging processes is less than 10% in the nearby universe [42], we built a class-balanced training set; i.e., each class that the NN learned to recognize was equally represented. Our selection of galaxy mergers was taken from the catalog that was built in [11, 12]; this was composed of visually confirmed mergers from Galaxy Zoo Data Release 1 (GZ DR1) [28]. We complemented it with non-merging galaxies (also from GZ DR1). The GZ DR1 project provides its volunteers with images of galaxies from a set of SDSS DR6 galaxies, and volunteers answer questions regarding their morphologies through the GZ webpage.

The galaxies from GZ with a vote fraction $f_m > 0.4$ (which was the fraction of the volunteers that answered positively to the question that asked whether the galaxies were merging or not) were visually confirmed by the authors in [12]. Galaxies with $f_m > 0.4$ were considered to show clear merging features in [12] during their inspection of the GZ data. This resulted in 3 003 merging galaxies in the redshift range of [0.005, 0.1], including the merging companions and some multi-merger cases that were stored in a public catalog[1].

As a preliminary non-merger set, we considered all GZ DR1 galaxies with $f_m < 0.2$ [44], as they showed very little merging distortions according to [12]. We aimed to have a distribution of mergers and non-mergers that were similar in mass and distance in order to avoid extra differences between the two populations. We considered the r-band photometric magnitude as an adequate proxy for galactic mass [32] and estimated the distance through the spectroscopic redshift that was obtained from the fiber's spectrometry. Then, each merger got a 2D distance match with one non-merger, thus providing the final catalog.

The final data set consisted of 2 930 mergers, with cuts in the low spectrometric redshift at spec-z = 0.01 and at a high magnitude of r-mag = 18.05 [51]. This final set of 5 860 sources was then divided into 5 360 galaxies for a k-fold cross-validation, leaving the remaining 500 galaxies for testing.

Besides the NN training data set, we also worked on HSC galaxies from the North Ecliptic Pole (NEP). The images that were used were cutouts (squared sections of images that were centered on each galaxy) that came from the AKARI-NEP optical observations [19, 37]. From these, we created cutouts of a selection of mergers and non-mergers from the Galaxy Zoo: Cosmic Down! (GZ:CD) program. This GZ

[1] https://data.galaxyzoo.org/

edition was based on the NEP galaxies that were observed in the 2019 HSC deep-field observations of the Hawaii Two-0 (H20) survey [60].

We created a class-balanced set of mergers and non-mergers from the GZ:CD galaxies that were found in the HSC-NEP area. The selected galaxy mergers had merging votes that were above 0.7, following the criteria of the GZ release previous to GZ:CD [55]. Similar to the SDSS DR6 data set, we matched the resulting mergers in r-band magnitudes and photometric redshifts; this resulted in 256 mergers and their 256 corresponding non-mergers.

3. Methodology

The main technique that was applied for this work was a neural network (NN); this connected a set of input variables with a final output (which was the probability of a galaxy of being a merger or a non-merger in this work). Between the input and output, there were a series of subsequent layers of mathematical neurons; these neurons connected with all of the neurons in the layers before and after, or in other words. The input and output were the first and last layer. In each neuron, a linear combination was applied to its input values using the internal parameters that were updated during the training process. Through the training, these internal parameters were updated by the optimization of a loss function that quantified how the model outputs resembled the original labels.

The architecture of the NN consisted of 2 layers with 16 neurons each, with its input parameters being a heterogeneous combination of photometric measurements; the number of input variables ranged between 5 and 20. We applied a dropout rate of 0.2, and the chosen activation function for the layers was ReLU (even though the output probabilities were given by a softmax function). The initial training rate for the Adam optimization was $5{\times}10^{-5}$.

In our model, we applied k-fold cross-validation (specifically, with $k = 5$). The training galaxies were split equally and in a class-balanced manner in five sets so that the training was performed on four of them each time. The remaining fold was used for validation, which consisted of an independent estimation of the relevant NNs performance that differed from the training in that it was not used for updating the NNs parameters. By doing this in the five folds, the final performance parameters could be averaged over the five resulting trainings. Finally, the test set was only considered for obtaining the performance of the NN in a set of data that it had not seen at any point during the training. The NNs converged at around 900 epochs (when the validation accuracy and loss function stopped improving).

We quantified performance of the NN through four classification groups: true positives (TPs) (which were correctly classified as mergers); true negatives (TNs) (analogous for non-mergers); false negatives (FNs) (mergers that were wrongly identified as non-mergers); and false positives (FPs) (non-mergers that were mistaken as mergers).

The accuracy of the NN was used to assess the quality of the model as follows:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \ . \tag{1}$$

In order to further understand the results of the NN through `skyErr` (see Section 4 for its definition) and extend it to the HSC imaging, we estimated the medians and standard deviations of the pixel distributions within circular apertures on the cutouts of the HSC-NEP galaxies in our H20-GZ:CD catalog. We subsequently applied a cut on pixels with values that were greater than the median plus the standard deviation (`median + std`) on the cutouts. The goal of this was to analyze the low surface-brightness (LSB) pixels around the galaxies. The LSB pixels were then analyzed through multiple parameters that characterized their distribution. The parameters were as follows: `mean` and `median`, the interquartile range (`IQR`), three similar parameters (dubbed `Fraction 1`, `2`, and `3`), the `Skewness`, and `Kurtosis`. The three fractions were as follows:

$$\text{Fraction}\,1 = \frac{q_{75} - q_{50}}{q_{50} - q_{25}} \ , \ \text{Fraction}\,2 = \frac{q_{100} - q_{75}}{q_{25} - q_{0}} \ , \ \text{Fraction}\,3 = \frac{q_{100} - q_{50}}{q_{50} - q_{0}} \tag{2}$$

where the q_x variables are the x quantiles.

Figure 1 represents the process for getting the LSB pixels (shown in the bottom-left panels).

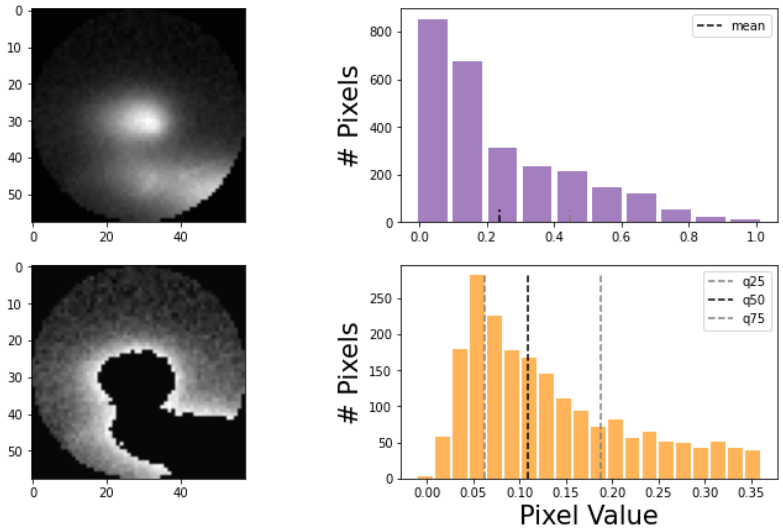

Figure 1. Top-left panel depicts the cutout of one merger; the top-right shows the pixel distribution of the given cutout and the small black and gray vertical lines represent the median and standard deviations, respectively; the bottom-left panel depicts the pixels within the clipped interval; the bottom-right histogram (orange) shows said pixel distribution, where the black and gray vertical lines now depict q50, q25, and q75 quartiles

The median and standard deviation of the distribution of the measurements per each pixel (top right) of the initial cutout (top left) were calculated so that the distribution (bottom-right) within the median plus the standard deviation of the resulting image (bottom-left) could then be analyzed by calculating the parameters that were defined above.

In order to simplify the parameter space that was obtained from the eight variables that were defined above, we applied the neighborhood components analysis (NCA) dimensionality-reduction technique [18]. The reason for using NCA was to define a region of the new space where the galaxy mergers were predominantly located instead of identifying them by the NN outputs. NCA is a non-parametric model that can generate a dimensionality-reduction embedding by taking known labels into account; it optimizes the distance in the embedding between the data points of the same class so that similarly labeled data points are located next to each other. The embedding transformation is a linear matrix between the input and the embedding space; thus, it can enhance the class separation that the combined eight dimensions may produce.

4. NN Results

Figure 2 indicates the results of training the NN; it shows the average and standard error of the validation accuracy that were calculated on the five folds that were used for the k-fold cross-correlation.

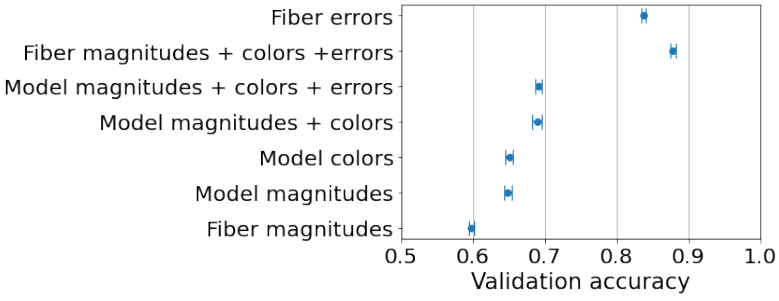

Figure 2. Five-fold cross-correlation validation accuracy for a selection of NN input combinations; indicated input combinations include parameters for each SDSS photometric band (i.e., fiber magnitudes include fiber magnitudes for five bands, or model colors include ten colors that could be obtained from five magnitudes, etc.): input parameters are min-max normalized between 0 and 1 for each galaxy

From bottom to top, the plot indicates how the accuracy increased through different combinations of inputs. For each galaxy, the given measurements were obtained for the five SDSS photometric bands, and they were min-max normalized for each galaxy before entering the NN at the input layer. The aperture magnitudes provided lower accuracy than the model magnitudes, which was expected given that

the model magnitudes included information on the galaxies' brightness profiles. The model colors had a similar value and error levels as the magnitudes did, which was to be expected given that the ten colors were the linear combinations of the five magnitudes. Surprisingly, the model magnitudes and colors worked better together than in the separated cases; this might have been because the extra information was positive for the NN performance. Adding the errors to this previous input, the accuracy was similar, but the error bar was smaller. Finally, the fiber counterpart reached an impressive accuracy of around 88%. Given the behavior of the fiber magnitudes and the model magnitudes, colors, and errors, it seemed difficult to understand what increased the accuracy. However, using exclusively fiber errors helped the NN reach around an 83% accuracy; this supported the fact that the key elements of the results were the fiber errors.

The fiber error was defined in a similar fashion to Equations A.4 and A.5 in [51]; thus, there were four different inputs for the fiber error: the aperture counts, the dark current variance of the CCD camera, the sky background, and `skyErr`. Figure 3 shows the NN accuracy results for some combinations of these parameters.

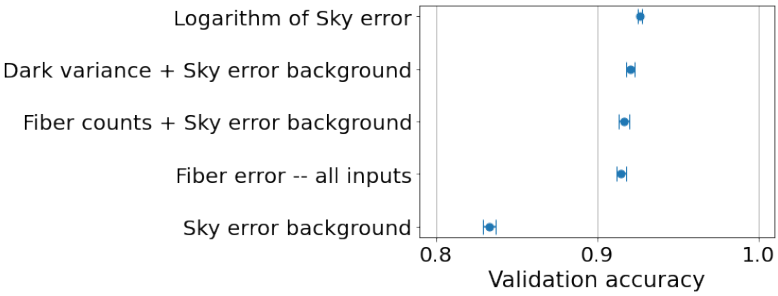

Figure 3. Five-fold cross-correlation validation accuracy for selection of NN input combinations; the inputs has been min-max normalized except for Logarithm Sky error on top; analogously to Fig. 2, each parameter was considered for each SDSS filter

The sky-background error alone provided a similar accuracy to that of the fiber error. The 20 variables that were the inputs of the fiber error itself reached a quite high accuracy of more than 91%, and the combination of the `skyErr` with the fiber counts or with the dark-current variance also reached more than 91%. The reason for this was the normalization: the `skyErr` was related to the merging state, but the dark variance depended only on the camera, so including it in the input modified how the sky-error parameters were normalized. In fact, the logarithm of `skyErr` was the best-performing NN: $92.64 \pm 0.15\%$ validation, and $92.36 \pm 0.21\%$ for the test set.

Further checks on the behavior of `skyErr` led to discovering that mergers could be found in the g and r-bands of the `skyErr` plane (as shown in Fig. 4). By drawing a boundary line in the plot and taking all of the sources above as mergers and below

as non-mergers, the obtained accuracy was 91.59%; this was a slightly lower accuracy result than the one obtained from the NN using the same input parameters. It definitely less-resource-consuming than the NN as the calculations are more simple: instead of all the NN internal equations, we just draw a line in a plane.

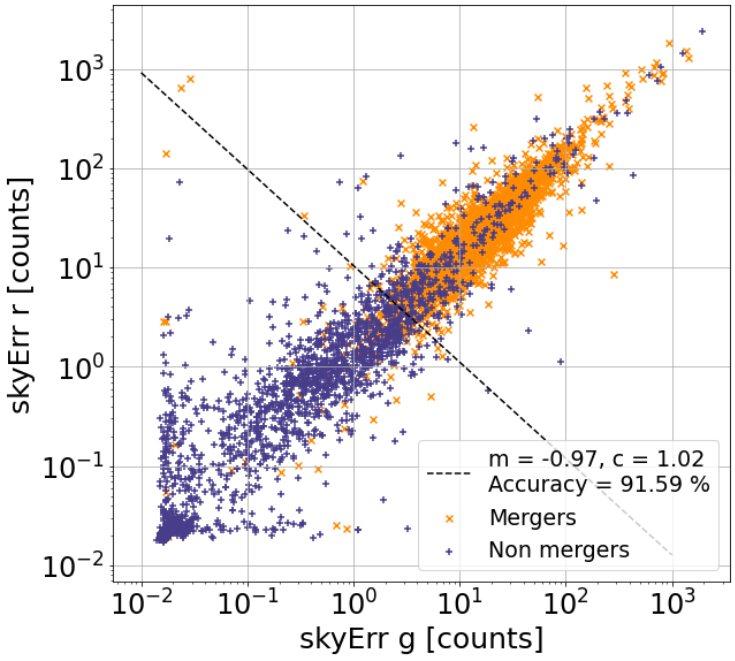

Figure 4. Training set galaxies in 2D plane of `skyErr` in r bands; mergers shown as orange crosses, and non-mergers shown as dark-blue plus symbols: boundary is dashed black line, with slope m, intercept c, and classification accuracy provided in label

5. Extension to HSC

The eight parameters that were described in Section 3 and calculated for the 256 mergers and 256 non-mergers in the HSC-NEP GZ:CD are shown in Figure 5. For each parameter, the bins were the same for each galaxy type (splitting the dynamical range equally). The number of galaxies per bin is shown on a logarithmic scale. Overall, only small separations between the mergers and non-mergers can be guessed from the distributions – mainly, in the lower ends of the skewness and kurtosis histograms. The NCA dimensionality reduction of these $236 \times 2 \times 8$ parameters is depicted in Figure 6, where the axes are the resulting axes from the embedding, and the red lines are narrowing where the abundance of mergers seems to be greater than non-mergers visually.

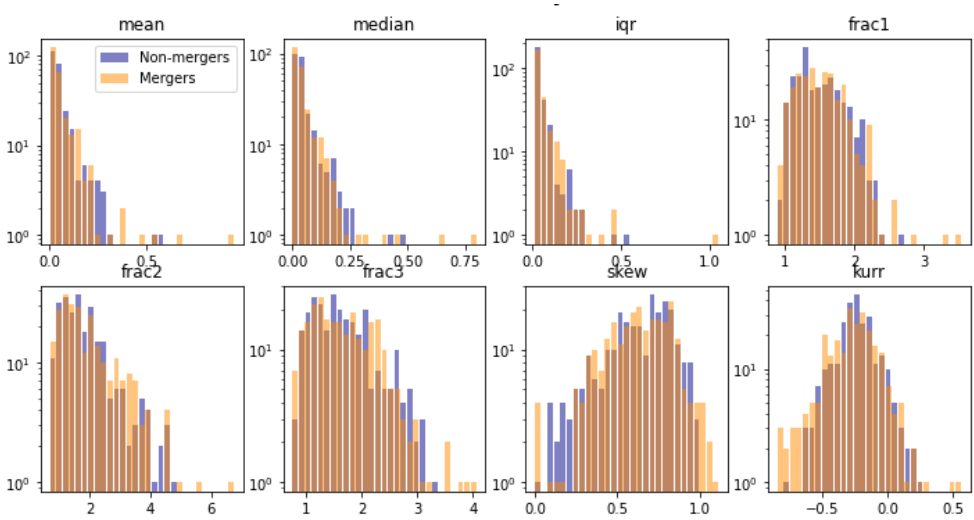

Figure 5. Histograms of eight parameters applied to HSC cutouts for mergers (in orange) and non-merger (in blue); y-axis shows number of galaxies in logarithmic scale, and x-axis indicates values of parameters that are indicated on top of each panel

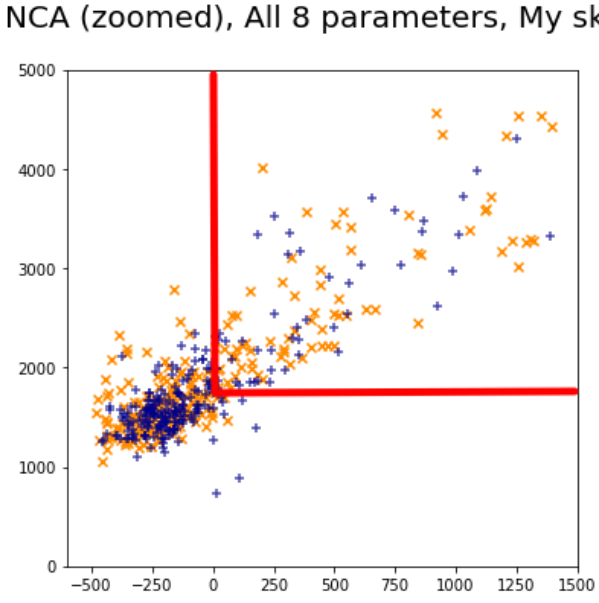

Figure 6. NCA dimensionality reduction of eight parameters shown in Figure 5: mergers are show as orange crosses, and non-mergers as blue plus signs; red lines that split top-right corner represent tentative decision boundary that could be applied to find mergers through such methodology

6. Conclusions

In this work, we trained multiple NNs on SDSS photometric data to identify galaxy mergers on a class-balanced data set of merging and non-merging galaxies. After testing multiple variables, we found not only a very high validation accuracy using the 5-band `skyErr` but also that the `skyErr` in the plane that was formed by the r and g band can be used to define a decision boundary to identify any mergers.

Then, we investigated a different way to characterize the sky background in HSC imaging by applying multiple parameters in the low S/N pixels of a smaller class-balanced data set. We checked how the mergers were located in the embedding that was obtained from applying the NCA dimensionality-reduction model to the parameters. This analysis was quite preliminary and very tentative; a further analysis of the background level and potential parameters that can encounter a signal that is similar to that of the SDSS `skyErr` is already ongoing and will be published in the near future.

Acknowledgments

L.E. Suelves was supported by the Estonian Ministry of Education and Research (grant TK202), Estonian Research Council Grant (PRG1006), and the European Union's Horizon Europe research and innovation program (EXCOSM, Grant No. 101159513). W.J. Pearson was supported by the Polish National Science Center projects (NCN) UMO-2020/37/B/ST9/00466 and UMO-2023/51/D/ST9/00147. A. Pollo was supported by NCN UMO 2023/50/A/ST9/00579.

References

[1] Abraham R.G., van den Bergh S., Glazebrook K., Ellis R.S., Santiago B.X., Surma P., Griffiths R.E.: The Morphologies of Distant Galaxies. II. Classifications from the Hubble Space Telescope Medium Deep Survey, *The Astrophysical Journals*, vol. 107, pp. 1–17, 1996. doi: 10.1086/192352.

[2] Abraham R.G., van den Bergh S., Nair P.: A New Approach to Galaxy Morphology. I. Analysis of the Sloan Digital Sky Survey Early Data Release, *The Astrophysical Journal*, vol. 588(1), pp. 218–229, 2003. doi: 10.1086/373919.

[3] Ackermann S., Schawinski K., Zhang C., Weigel A.K., Turp M.D.: Using transfer learning to detect galaxy mergers, *Monthly Notices of the Royal Astronomical Society*, vol. 479(1), pp. 415–425, 2018. doi: 10.1093/mnras/sty1398.

[4] Adelman-McCarthy J.K., Agüeros M.A., Allam S.S., Allende Prieto C., Anderson K.S., Anderson S.F., Annis J., *et al.*: The Sixth Data Release of the Sloan Digital Sky Survey, *The Astrophysical Journals*, vol. 175(2), pp. 297–313, 2008. doi: 10.1086/524984.

[5] Barton E.J., Geller M.J., Kenyon S.J.: Tidally Triggered Star Formation in Close Pairs of Galaxies, *The Astrophysical Journal*, vol. 530(2), pp. 660–679, 2000. doi: 10.1086/308392.

[6] Bershady M.A., Jangren A., Conselice C.J.: Structural and Photometric Classification of Galaxies. I. Calibration Based on a Nearby Galaxy Sample, *The Astronomical Journal*, vol. 119(6), pp. 2645–2663, 2000. doi: 10.1086/301386.

[7] Bottrell C., Hani M.H., Teimoorinia H., Ellison S.L., Moreno J., Torrey P., Hayward C.C., *et al.*: Deep learning predictions of galaxy merger stage and the importance of observational realism, *Monthly Notices of the Royal Astronomical Society*, vol. 490(4), pp. 5390–5413, 2019. doi: 10.1093/mnras/stz2934.

[8] Conselice C.J.: The Relationship between Stellar Light Distributions of Galaxies and Their Formation Histories, *The Astrophysical Journals*, vol. 147(1), pp. 1–28, 2003. doi: 10.1086/375001.

[9] Conselice C.J.: Early and Rapid Merging as a Formation Mechanism of Massive Galaxies: Empirical Constraints, *The Astrophysical Journal*, vol. 638(2), pp. 686–702, 2006. doi: 10.1086/499067.

[10] Conselice C.J., Bershady M.A., Jangren A.: The Asymmetry of Galaxies: Physical Morphology for Nearby and High-Redshift Galaxies, *The Astrophysical Journal*, vol. 529(2), pp. 886–910, 2000. doi: 10.1086/308300.

[11] Darg D.W., Kaviraj S., Lintott C.J., Schawinski K., Sarzi M., Bamford S., Silk J., *et al.*: Galaxy Zoo: the properties of merging galaxies in the nearby Universe – local environments, colours, masses, star formation rates and AGN activity, *Monthly Notices of the Royal Astronomical Society*, vol. 401(3), pp. 1552–1563, 2010. doi: 10.1111/j.1365-2966.2009.15786.x.

[12] Darg D.W., Kaviraj S., Lintott C.J., Schawinski K., Sarzi M., Bamford S., Silk J., *et al.*: Galaxy Zoo: the fraction of merging galaxies in the SDSS and their morphologies, *Monthly Notices of the Royal Astronomical Society*, vol. 401(2), pp. 1043–1056, 2010. doi: 10.1111/j.1365-2966.2009.15686.x.

[13] De Propris R., Liske J., Driver S.P., Allen P.D., Cross N.J.G.: The Millennium Galaxy Catalogue: Dynamically Close Pairs of Galaxies and the Global Merger Rate, *The Astronomical Journal*, vol. 130(4), pp. 1516–1523, 2005. doi: 10.1086/433169.

[14] Domínguez Sánchez H., Martin G., Damjanov I., Buitrago F., Huertas-Company M., Bottrell C., Bernardi M., *et al.*: Identification of tidal features in deep optical galaxy images with convolutional neural networks, *Monthly Notices of the Royal Astronomical Society*, vol. 521(3), pp. 3861–3872, 2023. doi: 10.1093/mnras/stad750.

[15] Duncan K., Conselice C.J., Mundy C., Bell E., Donley J., Galametz A., Guo Y., *et al.*: Observational Constraints on the Merger History of Galaxies since $z \approx 6$: Probabilistic Galaxy Pair Counts in the CANDELS Fields, *The Astrophysical Journal*, vol. 876(2), 110, 2019. doi: 10.3847/1538-4357/ab148a.

[16] Ellison S.L., Patton D.R., Simard L., McConnachie A.W.: Galaxy Pairs in the Sloan Digital Sky Survey. I. Star Formation, Active Galactic Nucleus Fraction, and the Luminosity/Mass-Metallicity Relation, *The Astronomical Journal*, vol. 135(5), pp. 1877–1899, 2008. doi: 10.1088/0004-6256/135/5/1877.

[17] Ferreira L., Conselice C.J., Duncan K., Cheng T.-Y., Griffiths A., Whitney A.: Galaxy Merger Rates up to z $\sim$ 3 Using a Bayesian Deep Learning Model: A Major-merger Classifier Using IllustrisTNG Simulation Data, *The Astrophysical Journal*, vol. 895(2), 115, 2020. doi: 10.3847/1538-4357/ab8f9b.

[18] Goldberger J., Roweis S., Hinton G.H., Salakhutdinov R.: Neighbourhood components analysis. In: *NIPS'04: Proceedings of the 18th International Conference on Neural Information Processing Systems*, vol. 17, pp. 513–520, 2004.

[19] Goto T., Toba Y., Utsumi Y., Oi N., Takagi T., Malkan M., Ohayma Y., *et al.*: Hyper Suprime-Camera Survey of the Akari NEP Wide Field, *Publication of Korean Astronomical Society*, vol. 32(1), pp. 225–230, 2017. doi: 10.5303/PKAS.2017.32.1.225.

[20] Gunn J.E., Carr M., Rockosi C., Sekiguchi M., Berry K., Elms B., de Haa E., *et al.*: The Sloan Digital Sky Survey Photometric Camera, *The Astronomical Journal*, vol. 116(6), pp. 3040–3081, 1998. doi: 10.1086/300645.

[21] Ivezić Ž., Kahn S.M., Tyson J.A., Abel B., Acosta E., Allsman R., Alonso D., *et al.*: LSST: From Science Drivers to Reference Design and Anticipated Data Products, *The Astrophysical Journal*, vol. 873(2), 111, 2019. doi: 10.3847/1538-4357/ab042c.

[22] Joseph R.D., Wright G.S.: Recent star formation in interacting galaxies – II. Super starbursts in merging galaxies., *Monthly Notices of the Royal Astronomical Society*, vol. 214, pp. 87–95, 1985. doi: 10.1093/mnras/214.2.87.

[23] Kent S.M.: CCD surface photometry of field galaxies. II. Bulge/disk decompositions., *The Astrophysical Journals*, vol. 59, pp. 115–159, 1985. doi: 10.1086/191066.

[24] Kitzbichler M.G., White S.D.M.: A calibration of the relation between the abundance of close galaxy pairs and the rate of galaxy mergers, *Monthly Notices of the Royal Astronomical Society*, vol. 391(4), pp. 1489–1498, 2008. doi: 10.1111/j.1365-2966.2008.13873.x.

[25] Lambas D.G., Tissera P.B., Sol Alonso M., Coldwell G.: Galaxy pairs in the 2dF survey – I. Effects of interactions on star formation in the field, *Monthly Notices of the Royal Astronomical Society*, vol. 346(4), pp. 1189–1196, 2003. doi: 10.1111/j.1365-2966.2003.07179.x.

[26] Le Fèvre O., Abraham R., Lilly S.J., Ellis R.S., Brinchmann J., Schade D., Tresse L., *et al.*: Hubble Space Telescope imaging of the CFRS and LDSS redshift surveys – IV. Influence of mergers in the evolution of faint field galaxies from z~1, *Monthly Notices of the Royal Astronomical Society*, vol. 311(3), pp. 565–575, 2000. doi: 10.1046/j.1365-8711.2000.03083.x.

[27] Lin L., Koo D.C., Willmer C.N.A., Patton D.R., Conselice C.J., Yan R., Coil A.L., *et al.*: The DEEP2 Galaxy Redshift Survey: Evolution of Close Galaxy Pairs and Major-Merger Rates up to z ~1.2, *The Astrophysical Journall*, vol. 617(1), pp. L9–L12, 2004. doi: 10.1086/427183.

[28] Lintott C., Schawinski K., Bamford S., Slosar A., Land K., Thomas D., Edmondson E., *et al.*: Galaxy Zoo 1: data release of morphological classifications for nearly 900 000 galaxies, *Monthly Notices of the Royal Astronomical Society*, vol. 410(1), pp. 166–178, 2011. doi: 10.1111/j.1365-2966.2010.17432.x.

[29] Lotz J.M., Jonsson P., Cox T.J., Primack J.R.: Galaxy merger morphologies and time-scales from simulations of equal-mass gas-rich disc mergers, *Monthly Notices of the Royal Astronomical Society*, vol. 391(3), pp. 1137–1162, 2008. doi: 10.1111/j.1365-2966.2008.14004.x.

[30] Lotz J.M., Primack J., Madau P.: A New Nonparametric Approach to Galaxy Morphological Classification, *The Astronomical Journal*, vol. 128(1), pp. 163–182, 2004. doi: 10.1086/421849.

[31] Lupton R.H., Gunn J.E., Szalay A.S.: A Modified Magnitude System that Produces Well-Behaved Magnitudes, Colors, and Errors Even for Low Signal-to-Noise Ratio Measurements, *The Astronomical Journal*, vol. 118(3), pp. 1406–1410, 1999. doi: 10.1086/301004.

[32] Mahajan S., Drinkwater M.J., Driver S., Hopkins A.M., Graham A.W., Brough S., Brown M.J.I., *et al.*: Galaxy and Mass Assembly (GAMA): blue spheroids within 87 Mpc, *Monthly Notices of the Royal Astronomical Society*, vol. 475(1), pp. 788–799, 2018. doi: 10.1093/mnras/stx3202.

[33] Margalef-Bentabol B., Wang L., La Marca A., Blanco-Prieto C., Chudy D., Domínguez-Sánchez H., *et al.*: Galaxy merger challenge: A comparison study between machine learning-based detection methods, *Astronomy & Astrophysics*, vol. 687, A24, 2024. doi: 10.1051/0004-6361/202348239.

[34] Mihos J.C., Hernquist L.: Gasdynamics and Starbursts in Major Mergers, *The Astrophysical Journal*, vol. 464, 641, 1996. doi: 10.1086/177353.

[35] Nevin R., Blecha L., Comerford J., Greene J.: Accurate Identification of Galaxy Mergers with Imaging, *The Astrophysical Journal*, vol. 872(1), 76, 2019. doi: 10.3847/1538-4357/aafd34.

[36] Nevin R., Blecha L., Comerford J., Simon J., Terrazas B.A., Barrows R.S., Vázquez-Mata J.A.: A declining major merger fraction with redshift in the local Universe from the largest-yet catalogue of major and minor mergers in SDSS, *Monthly Notices of the Royal Astronomical Society*, vol. 522(1), pp. 1–28, 2023. doi: 10.1093/mnras/stad911.

[37] Oi N., Goto T., Matsuhara H., Utsumi Y., Momose R., Toba Y., Malkan M., *et al.*: Subaru/HSC deep optical imaging of infrared sources in the AKARI North Ecliptic Pole-Wide field, *Monthly Notices of the Royal Astronomical Society*, vol. 500(4), pp. 5024–5042, 2021. doi: 10.1093/mnras/staa3080.

[38] Patton D.R., Grant J.K., Simard L., Pritchet C.J., Carlberg R.G., Borne K.D.: A Hubble Space Telescope Snapshot Survey of Dynamically Close Galaxy Pairs in the CNOC2 Redshift Survey, *The Astronomical Journal*, vol. 130(5), pp. 2043–2057, 2005. doi: 10.1086/491672.

[39] Patton D.R., Pritchet C.J., Carlberg R.G., Marzke R.O., Yee H.K.C., Hall P.B., Lin H., *et al.*: Dynamically Close Galaxy Pairs and Merger Rate Evolution in the CNOC2 Redshift Survey, *The Astrophysical Journal*, vol. 565(1), pp. 208–222, 2002. doi: 10.1086/324543.

[40] Patton D.R., Pritchet C.J., Yee H.K.C., Ellingson E., Carlberg R.G.: Close Pairs of Field Galaxies in the CNOC1 Redshift Survey, *The Astrophysical Journal*, vol. 475(1), pp. 29–42, 1997. doi: 10.1086/303535.

[41] Pearson W.J., Rodriguez-Gomez V., Kruk S., Margalef-Bentabol B.: Determining the time before or after a galaxy merger event, *Astronomy & Astrophysics*, vol. 687, A45, 2024. doi: 10.1051/0004-6361/202449532.

[42] Pearson W.J., Suelves L.E., Ho S.C.C., Oi N., Brough S., Holwerda B.W., Hopkins A.M., *et al.*: North Ecliptic Pole merging galaxy catalogue, *Astronomy & Astrophysics*, vol. 661, A52, 2022. doi: 10.1051/0004-6361/202141013.

[43] Pearson W.J., Wang L., Alpaslan M., Baldry I., Bilicki M., Brown M.J.I., Grootes M.W., *et al.*: Effect of galaxy mergers on star-formation rates, *Astronomy & Astrophysics*, vol. 631, A51, 2019. doi: 10.1051/0004-6361/201936337.

[44] Pearson W.J., Wang L., Trayford J.W., Petrillo C.E., van der Tak F.F.S.: Identifying galaxy mergers in observations and simulations with deep learning, *Astronomy & Astrophysics*, vol. 626, A49, 2019. doi: 10.1051/0004-6361/201935355.

[45] Rodrigues M., Puech M., Flores H., Hammer F., Pirzkal N.: Testing the hierarchical assembly of massive galaxies using accurate merger rates out to z ∼1.5, *Monthly Notices of the Royal Astronomical Society*, vol. 475(4), pp. 5133–5143, 2018. doi: 10.1093/mnras/sty098.

[46] Sanders D.B., Mirabel I.F.: Luminous Infrared Galaxies, *Annual Review of Astronomy and Astrophysics*, vol. 34, pp. 749–792, 1996. doi: 10.1146/annurev.astro.34.1.749.

[47] Sazonova E., Morgan C.R., Balogh M., Alatalo K., Benavides J.A., Bluck A., Brough S., *et al.*: RMS asymmetry: a robust metric of galaxy shapes in images with varied depth and resolution, *ArXiv e-prints, arXiv:240405792*, 2024. doi: 10.48550/arXiv.2404.05792.

[48] Snyder G.F., Lotz J., Moody C., Peth M., Freeman P., Ceverino D., Primack J., *et al.*: Diverse structural evolution at $z > 1$ in cosmologically simulated galaxies, *Monthly Notices of the Royal Astronomical Society*, vol. 451(4), pp. 4290–4310, 2015. doi: 10.1093/mnras/stv1231.

[49] Sola E., Duc P.A., Richards F., Paiement A., Urbano M., Klehammer J., Bílek M., *et al.*: Characterization of low surface brightness structures in annotated deep images, *Astronomy & Astrophysics*, vol. 662, A124, 2022. doi: 10.1051/0004-6361/202142675.

[50] Stoughton C., Lupton R.H., Bernardi M., Blanton M.R., Burles S., Castander F.J., Connolly A.J., *et al.*: Sloan Digital Sky Survey: Early Data Release, *The Astronomical Journal*, vol. 123(1), pp. 485–548, 2002. doi: 10.1086/324741.

[51] Suelves L.E., Pearson W.J., Pollo A.: Merger identification through photometric bands, colours, and their errors, *Astronomy & Astrophysics*, vol. 669, A141, 2023. doi: 10.1051/0004-6361/202244509.

[52] Takamiya M.: Galaxy Structural Parameters: Star Formation Rate and Evolution with Redshift, *The Astrophysical Journal Supplement Series*, vol. 122(1), pp. 109–150, 1999. doi: 10.1086/313216.

[53] Toomre A., Toomre J.: Galactic Bridges and Tails, *The Astrophysical Journal*, vol. 178, pp. 623–666, 1972. doi: 10.1086/151823.

[54] Walmsley M., Ferguson A.M.N., Mann R.G., Lintott C.J.: Identification of low surface brightness tidal features in galaxies using convolutional neural networks, *Monthly Notices of the Royal Astronomical Society*, vol. 483(3), pp. 2968–2982, 2019. doi: 10.1093/mnras/sty3232.

[55] Walmsley M., Lintott C., Géron T., Kruk S., Krawczyk C., Willett K.W., Bamford S., *et al.*: Galaxy Zoo DECaLS: Detailed visual morphology measurements from volunteers and deep learning for 314 000 galaxies, *Monthly Notices of the Royal Astronomical Society*, vol. 509(3), pp. 3966–3988, 2022. doi: 10.1093/mnras/stab2093.

[56] Wang L., Pearson W.J., Rodriguez-Gomez V.: Towards a consistent framework of comparing galaxy mergers in observations and simulations, *Astronomy & Astrophysics*, vol. 644, A87, 2020. doi: 10.1051/0004-6361/202038084.

[57] White S.D.M., Frenk C.S.: Galaxy Formation through Hierarchical Clustering, *The Astrophysical Journal*, vol. 379, 1991. doi: 10.1086/170483.

[58] White S.D.M., Rees M.J.: Core condensation in heavy halos: a two-stage theory for galaxy formation and clustering., *Monthly Notices of the Royal Astronomical Society*, vol. 183, pp. 341–358, 1978. doi: 10.1093/mnras/183.3.341.

[59] York D.G., Adelman J., Anderson J.E., Anderson S.F., Annis J., Bahcall N.A., Bakken J.A., *et al.*: The Sloan Digital Sky Survey: Technical Summary, *The Astronomical Journal*, vol. 120(3), pp. 1579–1587, 2000. doi: 10.1086/301513.

[60] Zalesky L.M.: The Hawaii Two-0 Twenty Square Degree Survey, *Bulletin of the American Astronomical Society*, vol. 53(1), 2021. https://baas.aas.org/pub/2021n1i215p05.

Affiliations

Luis E. Suelves
Tartu Observatory, University of Tartu, Observatooriumi 1, Toravere 61602, Estonia,
National Centre for Nuclear Research, Pasteura 7, 02-093 Warszawa, Poland, suelves@ut.ee

William J. Pearson
National Centre for Nuclear Research, Pasteura 7, 02-093 Warszawa, Poland

Agnieszka Pollo
National Centre for Nuclear Research, Pasteura 7, 02-093 Warszawa, Poland
Astronomical Observatory of Jagiellonian University, Faculty of Physics,
Astronomy, and Applied Computer Science, ul. Orla 171, 30-244 Krakow, Poland

Received: 10.03.2025
Revised: 12.03.2025
Accepted: 12.03.2025