

A machine learning approach to galactic emission-line region classification

Carter L. Rhea,^{1,2★†} Laurie Rousseau-Nepton,³ Ismael Moumen,^{2,3,4} Simon Prunet,⁵ Julie Hlavacek-Larrondo,¹ Kathryn Grasha^{1b},^{6,7} Carmelle Robert,⁴ Christophe Morisset,⁸ Grazyna Stasinska,⁹ Natalia Vale-Asari^{1b},^{10,11,12} Justine Giroux,⁴ Anna McLeod^{1b},^{13,14} Marie-Lou Gendron-Marsolais,¹⁵ Junfeng Wang,¹⁶ Joe Lyman^{1b}¹⁷ and Laurent Chemin¹⁸

¹Département de Physique, Université de Montréal, Succ. Centre-Ville, Montréal, Québec H3C 3J7, Canada

²Centre de Recherche en Astrophysique du Québec (CRAQ), Québec QC G1V 0A6, Canada

³Canada-France-Hawaii Telescope, Kamuela, HI 96743, USA

⁴Département de physique, de génie physique et d'optique, Université Laval, Québec QC H3A 1B9, Canada

⁵Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, F-06300, France

⁶Research School of Astronomy and Astrophysics, Australian National University, Weston Creek, ACT 2611, Australia

⁷ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Australia

⁸Instituto de Astronomía, Universidad Nacional Autónoma de México, AP 106, 22800 Ensenada, B. C., Mexico

⁹Laboratoire Univers et Théorie, Observatoire de Paris, Université PSL, Université Paris Cité, CNRS, F-92190 Meudon, France

¹⁰Departamento de Física-CFM, Universidade Federal de Santa Catarina, C.P. 476, 88040-900, Florianópolis, SC, Brazil

¹¹School of Physics and Astronomy, University of St Andrews, North Haugh, St Andrews KY16 9SS, UK

¹²Royal Society–Newton Advanced Fellowship, UK

¹³Centre for Extragalactic Astronomy, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK

¹⁴Institute for Computational Cosmology, Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK

¹⁵European Southern Observatory, Alonso de Coórdova 3107, Vitacura, Casilla 19001, Santiago de Chile

¹⁶Department of Astronomy, Xiamen University, Xiamen, Fujian 361005, China

¹⁷Department of Physics, University of Warwick, Coventry CV4 7AL, UK

¹⁸Instituto de Astrofísica, Universidad Andres Bello, Fernandez Concha 700, Las Condes, Santiago RM, Chile

Accepted 2023 June 21. Received 2023 June 20; in original form 2023 April 24

ABSTRACT

Diagnostic diagrams of emission-line ratios have been used extensively to categorize extragalactic emission regions; however, these diagnostics are occasionally at odds with each other due to differing definitions. In this work, we study the applicability of supervised machine-learning techniques to systematically classify emission-line regions from the ratios of certain emission lines. Using the Million Mexican Model database, which contains information from grids of photoionization models using `cloudy`, and from shock models, we develop training and test sets of emission line fluxes for three key diagnostic ratios. The sets are created for three classifications: classic H II regions, planetary nebulae, and supernova remnants. We train a neural network to classify a region as one of the three classes defined above given three key line ratios that are present both in the SITES and MUSE instruments' band-passes: $[\text{O III}]\lambda 5007/\text{H}\beta$, $[\text{N II}]\lambda 6583/\text{H}\alpha$, $([\text{S II}]\lambda 6717 + [\text{S II}]\lambda 6731)/\text{H}\alpha$. We also tested the impact of the addition of the $[\text{O II}]\lambda 3726$, $3729/[\text{O III}]\lambda 5007$ line ratio when available for the classification. A maximum luminosity limit is introduced to improve the classification of the planetary nebulae. Furthermore, the network is applied to SITES observations of a prominent field of M33. We discuss where the network succeeds and why it fails in certain cases. Our results provide a framework for the use of machine learning as a tool for the classification of extragalactic emission regions. Further work is needed to build more comprehensive training sets and adapt the method to additional observational constraints.

Key words: Machine Learning – Data Methods – Planetary Nebulae – Supernova Remnants – Galactic H II regions.

1 INTRODUCTION

Extragalactic emission line regions have been extensively studied over the last hundred years. These studies highlight the different

feedback mechanisms responsible for injecting significant amounts of energy in the interstellar medium (ISM) and for ionizing it (i.e. Veilleux & Osterbrock 1987; Osterbrock & Ferland 1989). Presently, a vast number of optical observations of extragalactic ionized gas in emission-line regions are available for which we have accurate measurements of the intensity of strong emission lines such as $\text{H}\alpha$, $\text{H}\beta$, $[\text{N II}]\lambda 6583$, $[\text{S II}]\lambda 6717$, $[\text{S II}]\lambda 6731$, $[\text{O III}]\lambda 5007$, and $[\text{O II}]\lambda 3726$, 3729 (e.g. Baldwin et al. 1981; MacAlpine & Williams

* E-mail: carter.rhea@umontreal.ca

† ARC DECRA Fellow.

1981; Ivezic et al. 2002; Kewley & Dopita 2002; Salzer et al. 2005; Moustakas et al. 2010; Kewley et al. 2019). These regions are generally categorized based on their morphological features (e.g. compact or extended, slope of the luminosity profile, general shape, etc.) or excitation mechanisms (radiative or mechanic). These can be segregated into three main classes for the most common bright emission line regions: H II regions photoionized by young hot stars or star clusters (e.g. Viallefond 1985; Melnick et al. 1987), supernova remnants and other shock induced emission regions (e.g. Danziger & Dennefeld 1976; Fesen et al. 1985), and planetary nebulae (e.g. Osterbrock 1964; Miller 1974). We do not consider activate galactic nuclei since this work is aimed at spatially resolved observations.

Initially categorized in Sérsic (1960) as luminous extragalactic emission, H II regions represent an important class of objects (e.g. Kennicutt 1984; Kennicutt et al. 1989). These regions emerge from giant molecular clouds where a young stellar cluster containing at least one ionizing O or B star is formed through the gravitational collapse of the cloud (e.g. Osterbrock & Ferland 1989). Depending on the properties of the ionizing sources, the region can vary significantly in size, luminosity, and morphology. Recombination lines from hydrogen and helium are predominant in the nebula's optical spectra. Numerous collisionally excited lines, CELs, emitted by different ions of metals such as oxygen, sulphur, and nitrogen are also present (e.g. Baldwin et al. 1981; Kewley et al. 2001; Kewley et al. 2006).

Meanwhile, planetary nebulae (PNe) are relatively compact objects formed by the gaseous ejecta (stellar envelope) from an evolving low-mass star which is later ionized by the star as it evolves to higher temperatures. Similar to classic H II regions, their optical spectra are dominated by strong recombination and collision lines (e.g. Osterbrock 1964; Miller 1974; Osterbrock & Ferland 1989). Supernovae remnants (SNRs) are formed from the gaseous debris scattered following the explosive death of a massive star (e.g. Fesen et al. 1985) or thermonuclear runaway in a white dwarf (Iben & Tutukov 1984). They are ionized both by high-velocity shocks in the ISM and the stellar remnant. Their size, surface brightness, and morphology evolve quickly through time until they blend with the surrounding diffuse ionized gas medium of their host galaxy (e.g. Woltjer 1972; Smith et al. 1993; Moumen et al. 2019).

The differing ionization mechanisms and underlying physics of these three main classes of emission-line regions manifest themselves distinctly in the relative intensity of the lines of their optical spectra. Although the original Baldwin–Phillips–Terlevich (BPT; Baldwin et al. 1981) diagnostic diagrams have been revised several times, they continue to represent the main emission mechanism characterization tools (e.g. Veilleux & Osterbrock 1987; Kewley et al. 2013). Several prominent theoretical diagnostic lines for these diagrams exist; the most notable are the Kewley et al. (2001) and Kauffmann et al. (2003) diagnostics. Using these diagnostic line formulas, observers are able to classify excitation mechanisms. However, discrepancies between the two diagnostic lines, along with issues segregating shocked gas emission from AGN emission, have led to the creation of new classification schemes (Constantin & Vogeley 2006; Kewley et al. 2006; de Souza et al. 2017; D'Agostino et al. 2019). Moreover, the data used to create the diagnostic lines were determined from integrated spectra of galaxies rather than resolved, parsec-scale objects.

New integral field units (IFUs) are revolutionizing the way in which extragalactic emission-line regions are studied by providing both the spectral and spatial information at the same time (e.g. Hénault et al. 2003; Kreckel et al. 2017; Della Bruna et al. 2020; McLeod et al. 2021). Using these instruments, several observatories

are deploying large legacy surveys; SIGNALS is the Star formation, Ionized Gas, and Nebular Abundances Legacy Survey currently being conducted at the Canada–France–Hawaii Telescope (CFHT; Rousseau-Nepton et al. 2019). The program uses over 350 h of observing time on the CFHT's new imaging Fourier Transform Spectrometer, SITELLE (e.g. Baril et al. 2016; Martin & Drissen 2017; Drissen et al. 2019). SITELLE produces spectral cubes that contain more than 4 million spaxels with a varying spectral resolution ($R \sim 1\text{--}10\,000$). This enables a detailed study of extragalactic emission regions in which we can resolve structures spatially and obtain tight constraints on their emission-line ratios.

In this paper, we explore the use of artificial neural networks (ANNs) to categorize extragalactic emission-line objects into H II regions, SNRs, and PNe. In Section 2.1, we describe the synthetic data set and the neural network architecture used in our analysis. In Section 3, we share the results of our network and compare them with traditional classification techniques. We apply the network to a SITELLE field of M33 in Section 4. In Section 5, we present the conclusions of our work.

2 METHODOLOGY, OBSERVATIONS, AND SIMULATIONS

2.1 Methodology

In the following section, we outline the databases and methods used to create synthetic emission-line ratios and the machine learning algorithm utilized to classify the ratios into different ionization mechanisms.

2.2 Synthetic data

In order to train and test our classification methodology, a set of emission-line ratios labelled by the region's type (i.e. H II, PNe, or SNR) is required. We use the Million Mexican Model Database (3MdB; Morisset et al. 2015; Alarie & Morisset 2019) and its ancillary databases to construct the training, validation, and test sets). 3MdB contains several grids of simulations that use the photoionization code `CLOUDY` (v.17) to emulate the expected emission from different ionizing sources and their surrounding ISM (Ferland et al. 2017). We take the line intensity values of lines of interest (discussed in detail below), which are quoted in ergs/s in the 3MdB. We use the 3MdB project entitled BOND (Asari et al. 2016) to obtain lines ratios from classic H II regions; similarly, we use the PNe (associated with Delgado-Inglada et al. 2014) project to obtain line ratios for planetary nebulae: the data can be found under PNe_2021 in the 3MdB. Since these models were run expressly for this project, we include a brief description here. The primary difference between the previous version, PNe_2020, and the updated version is that the O/H grid now covers a wider range of values (from -5.46 to -2.96) with finer sampling, and the N/O ratio was left free during the simulations. Together with other minor changes, the new models are more representative of physical planetary nebulae expected to be seen by the SIGNALS collaboration. We further constrain the planetary nebulae sample following the methodology outlined in Delgado-Inglada et al. (2014).

Regarding the H II regions, we only consider a subset of the entire BOND simulation set in 3MdB since the original set contains models that are not likely to represent H II regions that can be found in nature. A method to select a subset of the model database, along with a detailed discussion on the necessity to subsample the database, is presented for the case of giant H II regions in Amayo et al. (2021).

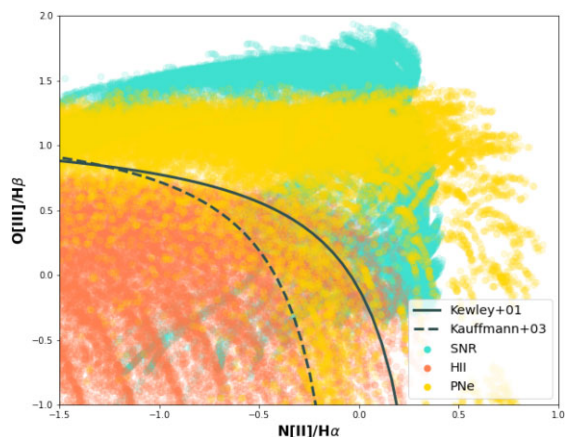


Figure 1. $[\text{N II}] \lambda 6583/\text{H}\alpha$ versus $[\text{O III}] \lambda 5007/\text{H}\beta$ diagnostic plots. The diagnostic lines from Kewley et al. (2001) and Kauffmann et al. (2003) are plotted in black. Values beneath the Kauffmann et al. (2003) line are interpreted as photo-ionized while regions above the Kewley et al. (2001) line are considered as shock-ionized regions. Points that lie between the two curves are classified as composite regions. The points represent our training, validation, and test sets. Fig. B1 shows this BPT diagram broken down into each region.

In this study, we created a subset of the model that remains broader, constraining only the physical parameters within the range expected in local galaxies; we retain only the completely filled geometry (i.e. the gas fills the entire volume), which allows us to focus on a younger population of H II regions (the ages are between 1 and 6 Myr; Cedrés et al. 2013; Stasinska et al. 2015; Asari et al. 2016). Moreover, we constrain the ionization parameter, $\log(U)$, between -3.5 and -2.5 , the metallicity proxy, $12+\log(\text{O}/\text{H})$, between 7.4 and 9.0 , and the nitrogen-to-oxygen ratio, $\log(\text{N}/\text{O})$, between -2 and 0 (Kashino & Inoue 2019; Pérez-Montero et al. 2019; Rousseau-Nepton et al. 2019; Zinchenko et al. 2019; Rhea et al. 2020). This selection ensured that we kept models that could represent the space of parameters for the giant H II regions, faint H II regions, and those lying in uncommon environments. It also includes indirect regions composed of one or a few ionizing stars (O or B stars) since the ionizing spectrum of a stellar population is dominated by the emission of the most massive stars; as the age increases, it ultimately becomes dominated by the late B stars.

The supernova remnant emission lines were taken from the 3MdB¹ table described in Alarie & Morisset (2019). Although 3MdB is not specifically for supernova remnants, we interpret the simulated values as coming from such objects. We note this may affect subsequent classifications of real data. Information pertaining to each grid can be found in the respective project’s reference paper. The data of the subgrids are shown in Figs 1 and 2 plotted on characteristic BPT diagrams with diagnostic lines overlaid.

Each simulation contains information for thousands of emission lines. This paper focuses on the emission lines available as part of the SIGNALS program. Although the classic BPT diagrams require four line ratios ($[\text{O III}] \lambda 5007/\text{H}\beta$, $[\text{N II}] \lambda 6583/\text{H}\alpha$, $([\text{S II}] \lambda 6717 + [\text{S II}] \lambda 6731)/\text{H}\alpha$, $[\text{O I}] \lambda 6300/\text{H}\alpha$), we excluded $[\text{O I}] \lambda 6300/\text{H}\alpha$ since $[\text{O I}] \lambda 6300$ is outside of the standard SITES filters (e.g. Baldwin et al. 1981; Kewley et al. 2006; Martin & Drissen 2017). We thus are left with three strong line ratios which are all attainable using the SITES filters SN2 (480–520 nm) and SN3 (651–685 nm).

¹3MdB shock.

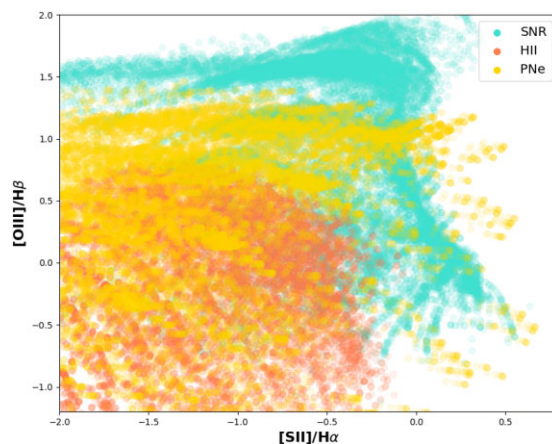


Figure 2. $[\text{S II}] \lambda(6731 + 6716)/\text{H}\alpha$ versus $[\text{O III}] \lambda 5007/\text{H}\beta$ diagnostic plots. The theoretical AGN and LINER/Seyfert 2 lines are plotted in black and are taken from Kewley et al. (2006). The points represent our training, validation, and test sets. Fig. B2 shows this BPT diagram broken down into each region.

We retained only the models for which flux of the three strong lines used in classification is higher than 1 per cent that of the $\text{H}\alpha$ emission. This ensures that all lines are visible in the original emission spectra and can be detected in the SIGNALS program (e.g. Rousseau-Nepton et al. 2019). There are 31 911 H II regions, 266 238 PNe, and 355 683 SNR. We randomly sampled 30 000 models from each grid to reduce training time and have consistently sized samples. To avoid a bias in the network towards data with more training set data, we have set the number of PNe and SNR samples to be the same order of magnitude as the H II sample; our testing revealed that 30 000 samples is sufficient to well-classify both the PNe and SNR regions without artificially injecting bias. The reduction in sample size does not affect results since the grid sampling remains mostly uniform over the target parameter space. We further tested the effect of random sampling by rerunning our analysis ten times with ten differently sampled data sets.

2.3 Artificial neural network

Artificial neural networks, and their closely related counterpart, convolutional neural networks, are becoming ubiquitous in astronomical applications due to their versatility and speed (e.g. Bertin 1994; Biswas & Adlak 2018; Baron 2019; Shatskiy & Evgeniev 2019; Uzeirbegovic et al. 2020). In this paper, we explore the use of an artificial neural network to classify extragalactic emission regions into H II regions, PNe, and SNRs. The ANN takes the $[\text{O III}] \lambda 5007/\text{H}\beta$, $[\text{N II}] \lambda 6583/\text{H}\alpha$, $([\text{S II}] \lambda 6717 + [\text{S II}] \lambda 6731)/\text{H}\alpha$, ratios as inputs and outputs the most likely categorization and its corresponding probability. Following standard methods, we use 70 per cent of the synthetic data for the training set, 20 per cent for the validation set, and 10 per cent for the test set (e.g. Breiman 2001).

The neural network was built using `tensorflow` (v2.4.0) (Abadi et al. 2016) implemented in `python` (v3.6.0). A standard grid-search algorithm to determine the number of layers and nodes within each layer, implemented in `sklearn` (v0.23), revealed three layers with 126, 256, and 128 nodes, respectively, to be optimal. In each layer, the nodes are subject to the standard `relu` activation function; we use the categorical cross-entropy loss function. We implement two dropout layers of 25 per cent in between the first and second layers and the second and third layers. Since the `softmax` activation function is used on the last layer (i.e. the output

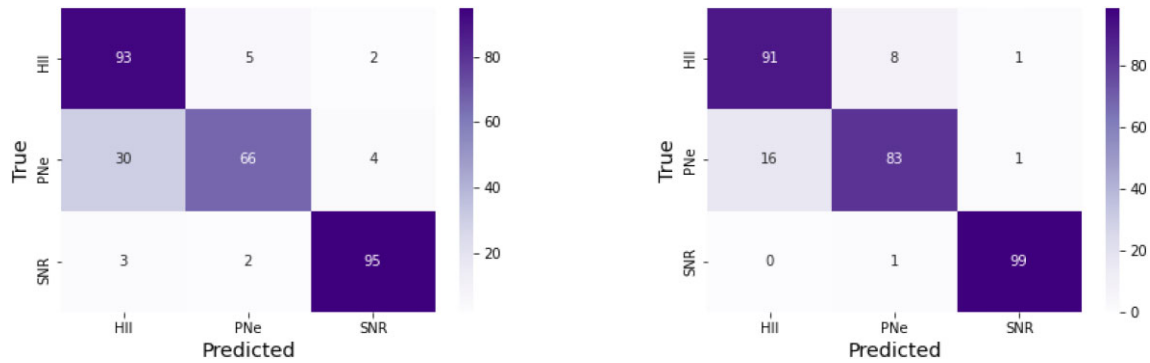


Figure 3. Confusion matrices for the artificial neural networks applied to the test set using three (left-hand panel) and four (right-hand panel) line ratios as input parameters to distinguish between H II regions, supernova remnants, and planetary nebulae. The left-hand side confusion matrix is based on the neural network trained on $[\text{O III}] \lambda 5007/\text{H}\beta$, $[\text{N II}] \lambda 6583/\text{H}\alpha$, $([\text{S II}] \lambda 6717 + [\text{S II}] \lambda 6731)/\text{H}\alpha$ while the right contains these three ratios and additionally the $[\text{O II}] \lambda 3726, 3729/[\text{O III}] \lambda 5007$ line ratio.

layer), each node of the output layer has a final logit value associated with the probability of the classification being an H II region (0), a planetary nebula (1), or a supernova remnant (2). We also employ the Adam optimizer algorithm described in Kingma & Ba (2017).

Further network details can be seen in the demo code at <https://github.com/sitelle-signals/Pamplemousse>.

3 RESULTS AND DISCUSSION

3.1 Confusion matrix

Fig. 3 (left-hand panel) visualizes the ability of the network to accurately categorize emission regions into either classic H II regions, planetary nebulae, or supernova remnants using the three-line ratios ($[\text{O III}] \lambda 5007/\text{H}\beta$, $[\text{N II}] \lambda 6583/\text{H}\alpha$, $([\text{S II}] \lambda 6717 + [\text{S II}] \lambda 6731)/\text{H}\alpha$) as inputs to the network. We stress that the data used for these line ratios (and thus the confusion matrix results) are valid only for data well-sampled by the test set. If, in reality, the test set is not representative of observation, then the results are expected to be worse than what is reported here. Each element of the confusion matrix can be interpreted as the efficiency (percentage of accurate classifications) at which the neural network categorizes the emission region y to their true type x . For example, the first row indicates that the network correctly categorizes H II regions 93 per cent of the time but that it misclassifies H II as either PNe or SNR 5 per cent and 2 per cent of the time, respectively. Similarly, PNe regions are correctly categorized 66 per cent of the time and are misclassified as either H II regions or supernova remnants 30 per cent and 4 per cent of the time, respectively. Finally, supernova remnants are correctly classified 95 per cent of the time and are misclassified as H II regions and planetary nebulae 3 per cent and 2 per cent of the time, respectively. A diagonal confusion matrix, such as we have, indicates that the network is correctly classifying the emission regions the overwhelming majority of the time.

Additionally, we report the confusion matrix (Fig. 3, right-hand panel) when we incorporate a fourth line ratio, $[\text{O II}] \lambda 3726, 3729/[\text{O III}] \lambda 5007$, into the input vector of the network. As evidenced by the figure, the classification accuracy for planetary remnants increases significantly (from 66 per cent to 83 per cent), while the classification accuracy for the other two region types remains consistent.

However, we draw attention to the high level of planetary nebulae misclassified as H II regions; this incorrect categorization is likely due

to the confusion in this restricted three parameters input space of line ratios between H II regions and planetary nebulae. A possible avenue to solve this entanglement is integrating a luminosity threshold for the PNe. This will be discussed further in Section 3.3.

3.2 Portability to other instruments

Although the methodology described in this paper has been applied only to line ratios calculated from SITELLE data cubes, the algorithm is instrument-agnostic. Since the method requires three-line ratios, the only requirement is that the instrument captures a signal from 4861 to 6731 Å. Since several instruments do not capture the $[\text{O II}] \lambda 3726/3729$ lines (such as the Multi-Unit Spectroscopic Explorer, MUSE, instrument; Bacon et al. 2010), we opted not to use them as inputs for our classifier. None the less, we have demonstrated that the network obtains better classification accuracies using the $[\text{O II}] \lambda 3726, 3729/[\text{O III}] \lambda 5007$ line ratios. With this in mind, adding additional-relevant line ratios and/or other observational constraints (such as the physical relationship between luminosity and line ratios, size and luminosity, reducing when known the range of metallicity of the training set, etc.) could further increase the performance of the network.

3.3 Inclusion of information in the classification input vector

In this work, we only considered the three or four-line ratios present in standard BPT plots for our input vector; however, we note that the efficacy of the classifications may be improved by including other line ratios or equivalent widths. Since this work focuses on improving the existing BPT structure, we chose not to include other line ratios. We note, though, that work is currently being done to use unsupervised machine learning algorithms to explore other line ratios (Moumen et al., in preparation). We note that $\text{H}\alpha$ luminosity can be used as a prior to putting additional constrain on the planetary nebulae. Although, at the moment, this is not possible with the current parameters included in 3MdB, we propose the following avenue: incorporate a luminosity threshold above which a region could not be a planetary nebula. As stated by Delgado-Inglada et al. (2020), the dust-corrected $\text{H}\alpha$ (or $\text{H}\beta$) maximum luminosity can be used as a relevant threshold since it does not depend on the PNe metallicities. We define the $\text{H}\alpha$ log-luminosity limit for the PNe at 36.0 ± 0.1 , the value derived by Delgado-Inglada et al. (2020) using a sample of 500 extragalactic PNe covering various environments. It is also consistent with values reported by others in the literature, including

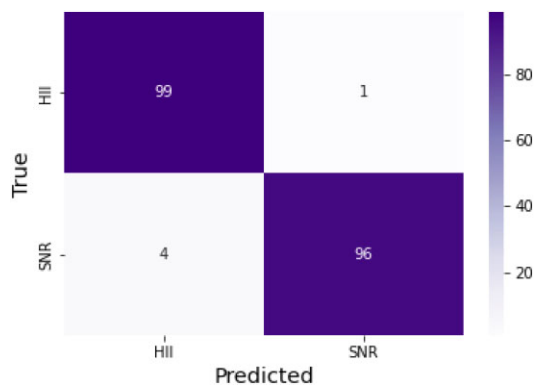


Figure 4. Confusion Matrix for the artificial neural network applied to the test set using the three line ratios ($[\text{O III}] \lambda 5007/\text{H} \beta$, $[\text{N II}] \lambda 6583/\text{H} \alpha$, $([\text{S II}] \lambda 6717 + [\text{S II}] \lambda 6731)/\text{H} \alpha$) as input parameters to distinguish between only H II regions and supernova remnants. The values quoted represent the classification percentage.

Martin et al. (2018) and Braun & Walterbos (1992). Any region with a $\text{H} \alpha$ luminosity above this limit cannot be a planetary nebula. If the region is below this limit, we use the classifier that distinguishes between planetary nebulae, supernova remnants, and H II regions. Otherwise, we use a classifier trained only to distinguish SNRs and H II regions. Therefore, we train an additional network using the same architecture described above to distinguish only H II regions and SNR. Fig. 4 demonstrates that the network excels as distinguishing between SNRs and H II regions.

Similar cuts could be applied for supernova remnants (e.g. low luminosity cut in Delgado-Inglada et al. 2020) or high luminosity cut using the brightest SNR observed in the local universe. We do not believe a low-luminosity cut should be used since SNR fade in intensity with time until their luminosity passes the detection limit of the observational configuration. On the other hand, a maximum luminosity could be considered, as it is for the PNe. From the literature, different values can be found: $4.96 \times 10^{37} \text{ erg s}^{-1}$ in Winkler et al. (2017). Nevertheless, these values are observational and not theoretical; one could expect an even higher maximum luminosity for supernovae surrounded by a denser ISM. Since SNRs are not a point-like source compared to extragalactic unresolved PNe, their emission is often embedded in H II regions that can be significantly brighter. In these cases, their estimated luminosity is bias towards higher values. Therefore, we decided not to use a maximum luminosity threshold for the SNR in this work. This decision is based on the possibility of evaluating SNR contamination in H II regions using the method. We will see later that three of the observed and confirmed SNRs used to test our method are above these observational thresholds and also coincide with objects embedded in H II regions. Additional elements regarding this aspect will be present in Section 4.2. It is also important to mention, as investigated in Vale Asari (2021), that line widths could also be included to help further distinguish regions. The addition of these supplementary input parameters is the subject of future work.

3.4 Addition of noise in the training set to fill in modelling gaps

Although the models obtained from 3MdB are extensive and include various scenarios, they are built on grids of model parameters. Therefore, gaps may exist in the distribution of parameters; these gaps can be seen in Figs 2 and 1. In order to explore the effect of these gaps on our results, we add random noise to the synthetic data.

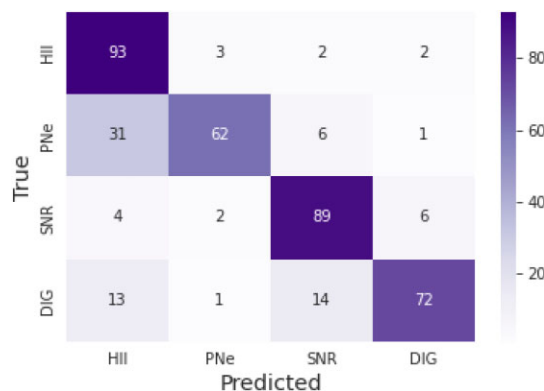


Figure 5. Confusion matrix including DIG as a fourth option using the three standard line ratios as input parameters.

In doing so, we draw the line ratio values randomly from a Gaussian distribution centred on the line ratio value from 3MdB with sigma values of 1 per cent, 2 per cent, and 5 per cent the line ratio value. Upon retraining the algorithm for each level of noise, we obtain similar accuracy values for each category.

3.5 Application to diffuse ionized gas

In addition to classical H II regions, supernova remnants, and planetary nebula, diffuse ionized gas (DIG) is an important feature of the ISM in many galaxies (e.g. Reynolds 1984; Walterbos & Braun 1994; Haffner et al. 2009). Therefore, we constructed a fourth classification set based on DIG simulations from 3MdB using the same methodology discussed in Section 2.2. We use the DIG_HR database and filter out H II regions by only retaining regions for which $\text{phi}_{\text{OB}} < 4.5$.² We retrain the artificial neural network developed in Section 2.3 with DIG regions as a fourth classification. We provide the confusion matrix in Fig. 5; the figure reveals that the network retains its accuracy of determining H II regions, planetary nebulae, and supernova remnants while achieving an accuracy of over 70 per cent in classifying DIGs. Incorrectly classified DIGs are equally split between supernova remnant and H II region classifications. While this level of accuracy is acceptable for many cases, we do not use this network in future sections. The authors also note that extensive work has been done to successfully distinguish DIG regions from H II regions using the equivalent width of the $\text{H} \alpha$ emission (e.g. Lacerda et al. 2018). Additionally, we supply users with a decision tree (see Fig. 6), which can be applied to determine which network to use.

4 APPLICATION TO M33

In order to test the method with observations, we use a field from the SIELLE instrument on the galaxy M33. These observations were conducted for the SIGNALS program and are ideal for this study since previous identification of SNR and PNe sources abounds in the literature and the high spatial resolution enables a precise selection of the emission area of the candidates.

4.1 The data

We extracted the PNe and SNR sources from the Ciardullo (2004) and Long et al. (2018) catalogues, respectively. In order to also

² phi_{OB} is the surface flux of the OB star.

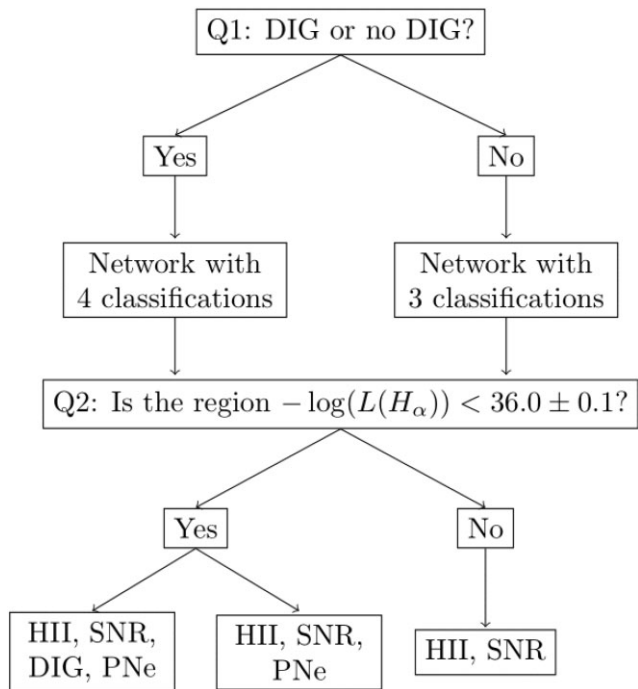


Figure 6. Tree diagram representing the possible classification routes based on the region's luminosity and the chosen classification scheme.

include a significant fraction of the H II regions visible in the field, we used a combination of the information provided from an emission map produced by adding all emission line maps together and the known position of the PNe and SNR. Using the ds9 tools, we handpicked a large fraction of the H II regions while keeping only the good candidates, excluding regions with not well-defined contours or obvious contamination. Fig. 7 shows all regions selected in the final test set. Some H II regions harbour a small overlap with a previously identified SNR. We tried to minimize the impact of these overlaps and avoid selecting H II regions that were particularly well blended with the SNRs. It is also important to note that the SNR radii were selected using the visible morphology of the SNR optical component (expanding shell). Nevertheless, we did not exclude known SNRs that are embedded in H II regions emission. We will discuss later in this section the impact of this inherent contamination.

Our literature review revealed 84 H II regions, 6 planetary nebulae, and 24 supernova remnants (see Fig. 7). However, upon further investigation of the spectra captured by SITELE of these regions, we rejected eight H II regions, four planetary nebulae, and two supernova remnants as having insufficient signal-to-noise³ to properly fit the emission lines. In order to extract the line ratios for these regions, we used the spectral analysis software LUCI. Using SN1 ($R \sim 1800$), SN2 ($R \sim 1800$), and SN3 ($R \sim 5000$) observations of the M33 field (PI: Laurie Rousseau-Nepton), we fit the following strong emission-lines using a `singgauss` function: H α , H β , [N II] $\lambda 6583$, [S II] $\lambda 6717$, [S II] $\lambda 6731$, [O III] $\lambda 5007$, [O III] $\lambda 4959$, [O II] $\lambda 3726$, and [O II] $\lambda 3729$. After calculating the flux of each line, we applied dereddening by calculating the Balmer decrement and using it in conjunction with a dereddening

³A signal-to-noise as calculated by LUCI below 2 is considered insufficient.

law.

$$F_{0,\lambda} = F_{\text{obs},\lambda} e^{\tau_\lambda} = F_{\text{obs},\lambda} e^{\tau_V q_\lambda}, \quad (1)$$

where $F_{\text{obs},\lambda}$ is the observed flux, τ_λ is the optical depth at a given wavelength, τ_V is the optical depth in the V-band, and the shape of the dust attenuation curve is parametrized by $q_\lambda \equiv \tau_\lambda/\tau_V$.

We use the Cardelli, Clayton & Mathis (1989) attenuation law with $R_V = 3.1$. We use the Balmer decrement, defined as $B_d = F_{\text{obs},\text{H}\alpha}/F_{\text{obs},\text{H}\beta}$, to calculate τ_V :

$$\tau_V = \frac{1}{q_{\text{H}\beta} - q_{\text{H}\alpha}} \ln \frac{B_d}{B_{d,\text{in}}}, \quad (2)$$

where $B_{d,\text{in}}$ is the intrinsic Balmer decrement which is assumed to be 2.87.

4.2 The results

In order to demonstrate the applicability of our two network frameworks described previously, we first apply the threshold on planetary nebulae luminosities derived above $-\log(L_{\text{H}\alpha}) = 36.0 \pm 0.1$. Regions with an H α luminosity below this threshold were passed to the NN including planetary nebula as an output classification; regions with a higher value were passed the NN restricted to H II regions and supernova remnants. There were 93 regions above this cut and 7 below. After experimentation, we found that on real data it is best to use all four input parameters if we have three classification options and only three input parameters (i.e. we do not use O II] $\lambda 3726$, 3729/[O III] $\lambda 5007$) if we have two classification options.

Fig. 8 shows the confusion matrix after running the network developed in Section 2.1 on the 100 zones in Field 7 of M33. As illustrated in the figure, the network excels at categorizing H II regions – as expected by the results reported in Section 3.1. The confusion matrix reveals that one planetary nebula is correctly categorized while the other is categorized as a supernova remnant; the spectra for these PNe are shown in Fig. C2. Furthermore, the network correctly classifies supernova remnants for approximately 65 per cent of the regions (i.e. 14 regions of 22 total regions). The remaining supernova remnants are incorrectly classified as H II regions. Therefore, we explore in detail the reasoning behind the network's miscategorizations of the supernova remnants.

Figs 9 and 10 show the placement of the M33 emission regions on standard BPT diagrams. From the figures, it is clear that the supernova remnants lie within the same regions as the H II regions in these diagrams. More importantly, by comparing Figs 9 and 10 with 1 and 2, we note that the supernova remnants in M33 do not fall within the same regions of line-ratio space as the training set; this explains why the network does a poor job of accurately classifying these regions. Fig. 11 shows similarly the placements of the regions identified using a network trained with the four input parameters over the [O II] $\lambda 3726$, 3729/H β versus [S II] $\lambda 6717$ /H α diagram. We note that the H II regions with a high [S II] $\lambda 6717$ /H α ratio are wrongfully identified as SNR, and the SNR with a high [O II] $\lambda 3726$, 3729/H β ratio are also generally wrongfully identified as H II regions.

Recent studies show that the parameter space indicated by the M33 supernovae may in fact be spanned by old supernova remnants (e.g. Moumen et al. 2019). Fig. 7 shows the location of the supernova remnants in cyan, the planetary nebulae in green, and the H II regions in yellow. As we can see, the regions identified as supernova remnants in the literature do not correspond with regions of strong emission; rather, they correspond with regions that would be considered as diffuse H II emission which is how our network classifies them. Therefore, we conclude that the misclassification

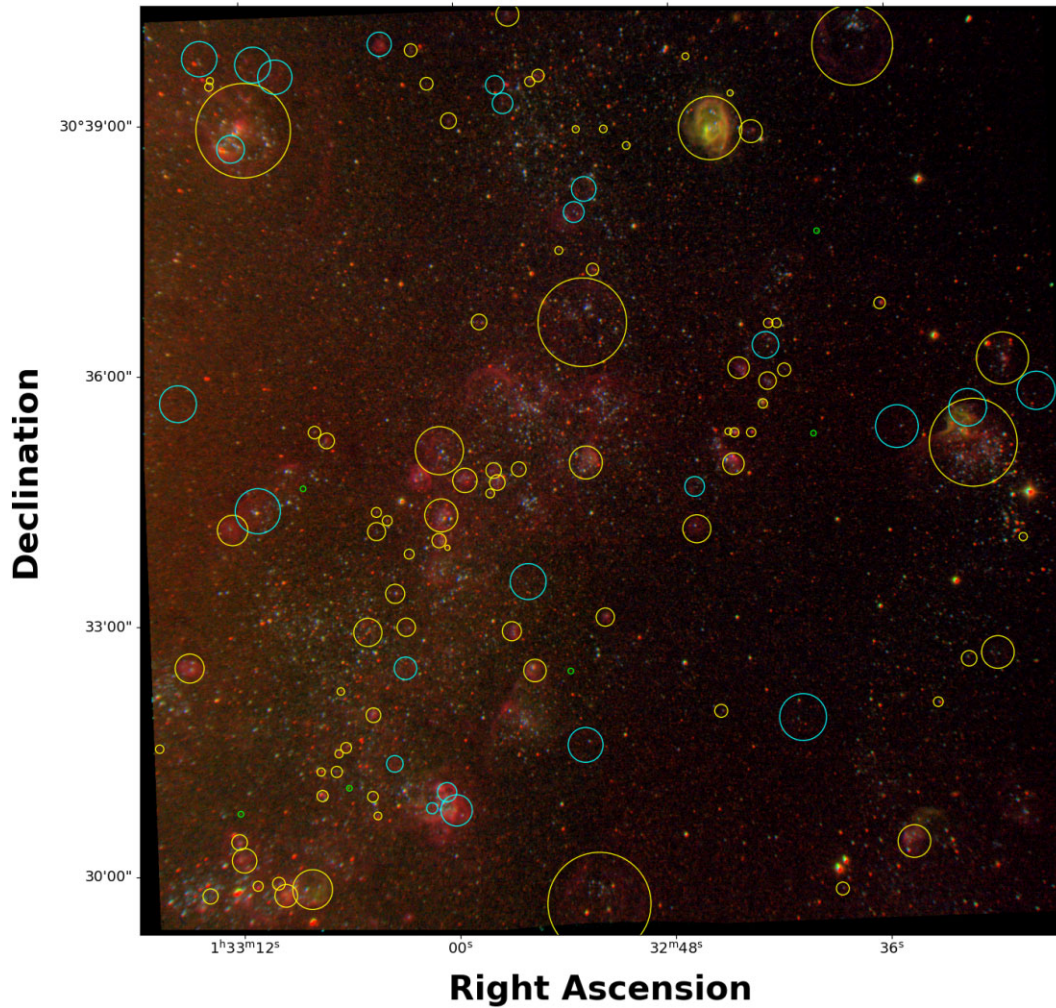


Figure 7. SITELE Field 7 of M33. This image is a combination of the SN1 deep frame (blue), the SN2 deep frame (green), and the SN3 deep frame (red). Here, we can see the location and size (corresponding to the circle size) of H II regions (yellow), planetary nebulae (green), and supernova remnants (cyan). This figure was made by using the `make_luption_rgb` function implemented in `astropy` (Lupton et al. 2004; Robitaille et al. 2013).

True	H II	62	3	11
	PNe	0	1	1
	SNR	8	0	14
		H II	PNe	SNR
		Predicted		

Figure 8. Confusion matrix for M33 Field 7 using a neural network trained with three input parameters: $[\text{O III}] \lambda 5007/\text{H}\beta$, $[\text{N II}] \lambda 6583/\text{H}\alpha$, $([\text{S II}] \lambda 6717 + [\text{S II}] \lambda 6731)/\text{H}\alpha$. Values quoted are the number of regions. We only retained regions for which the spectra in all three filters were adequate.

of SNRs is either due to contamination from nearby H II regions or DIGs (e.g. Cid Fernandes et al. 2021; see appendices E and F1 for an in-depth discussion), misclassification of these regions in the literature, or possibly an incomplete training set for supernova remnants (i.e. since our training set is a grid of shock models instead of supernova remnant models). We stress that the most likely culprit for the misclassification is the contamination from surrounding regions or incorrect classifications in the literature. However, a more comprehensive suite of SNR simulations under these conditions would be useful in improving the accuracy of the network on real data. Additionally, comparing the true and false classifications in these plots reveals that the network is automatically learning similar diagnostics as the ones presented by Kewley et al. (2001).

As noted in Section 3.4, there exist gaps in the supernova remnant data owing to the metallicity grid employed when creating the data set. Coincidentally, several misclassified supernova remnants lie in these gaps. Therefore, in order to ensure that these misclassifications are not due to the gaps in the training set, we interpolate the grid in metallicity in order to construct a new data set that does not present gaps (see appendix F2). Using this complete data set, we reconstruct our training, validation, and test sets. We then retrain our network and apply it to the M33 sample. The results indicate that

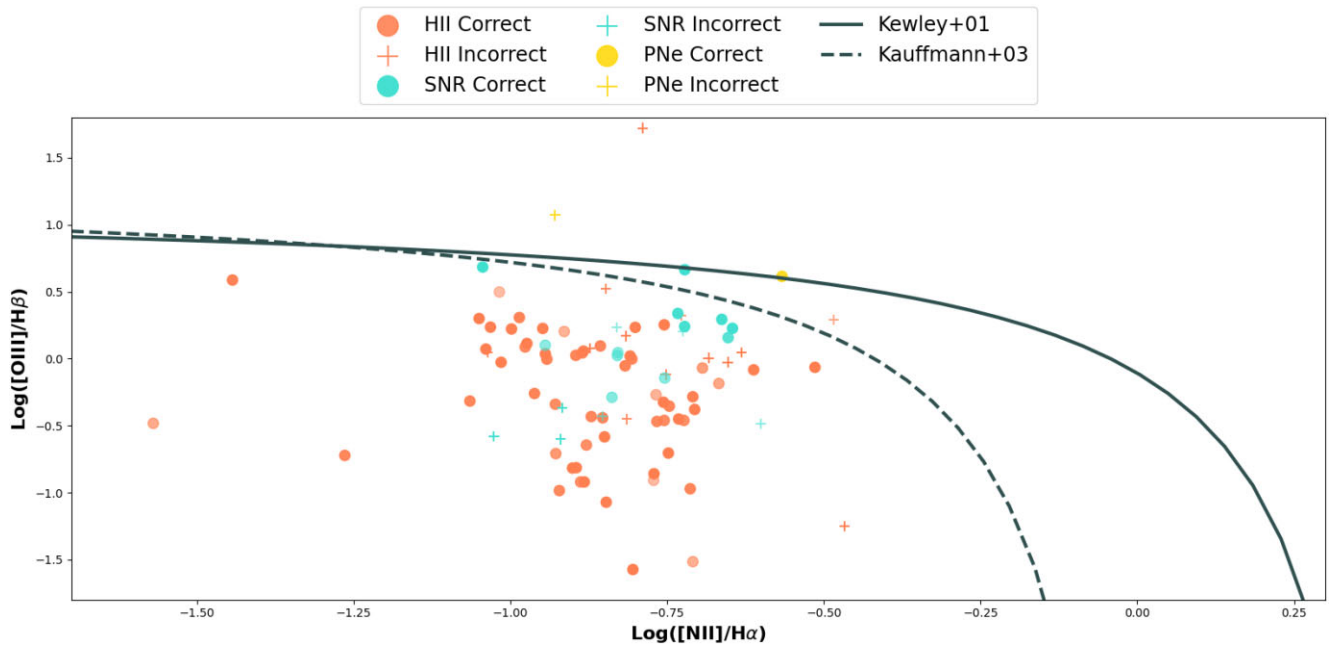


Figure 9. The BPT diagram of $\log([\text{N II}]/\text{H}\alpha)$ versus $\log([\text{O III}]/\text{H}\beta)$ for the regions identified in a SITELLE field of M33. The dashed and solid lines represent the standard Kewley et al. (2001) and Kauffmann et al. (2003) delineations. The emission region type is designated by differing colours. Additionally, the correct and incorrect classification (True and False in the legend) are indicated by circles and crosses, respectively. The points’ opacity corresponds to the probability of the classification. The lowest probability value is approximately 60 per cent. We stress that this plot is not to be taken as a comparison between the standard BPT diagnostic cuts and our methodology.

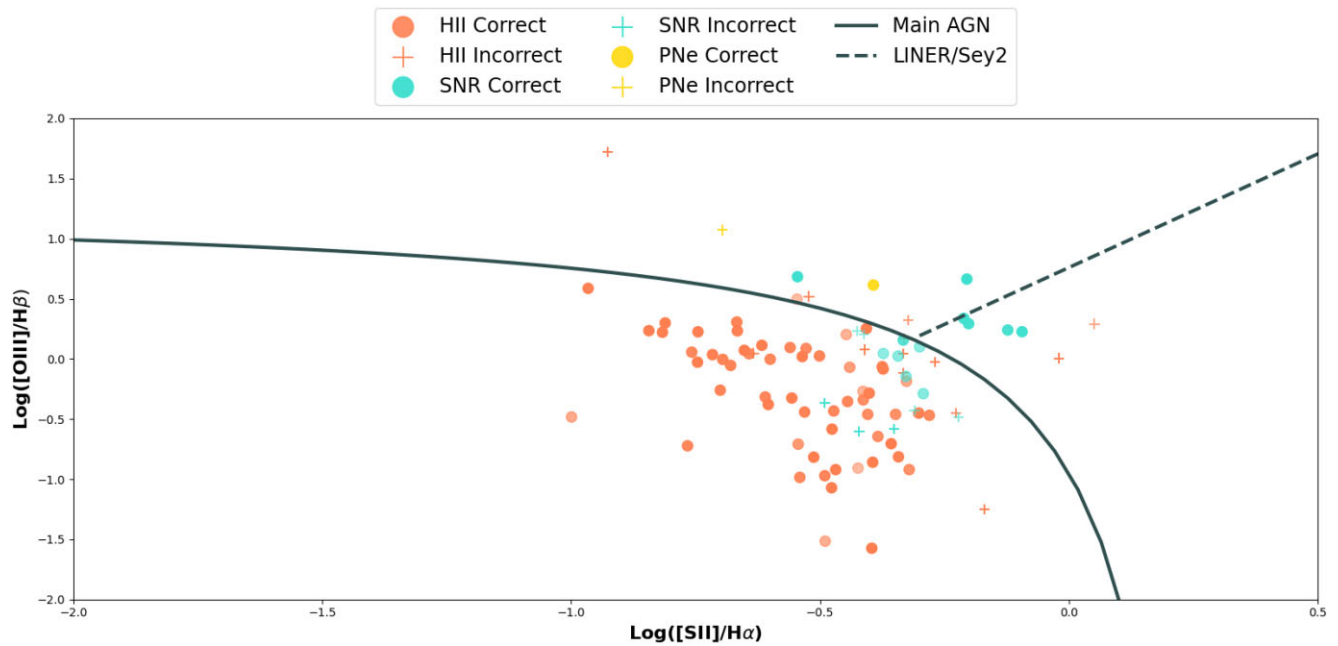


Figure 10. The BPT diagram of $\log([\text{S II}]/\text{H}\alpha)$ versus $\log([\text{O III}]/\text{H}\beta)$ for the regions identified in a SITELLE field of M33. The dashed and solid lines represent the standard AGN and LINER delineations. The emission region type is designated by differing colours. Additionally, the correct and incorrect classification (True and False in the legend) are indicated by circles and crosses, respectively. The points’ opacity corresponds to the probability of the classification. The lowest probability value is approximately 60 per cent. We stress that this plot is not to be taken as a comparison between the standard BPT diagnostic cuts and our methodology.

while supernova remnant classification slightly increases, several H II regions are now misclassified as supernova remnants. These results highlight the importance of the training set when applying machine learning algorithms to real data.

5 CONCLUSIONS

In this work, we have demonstrated the feasibility of using a popular machine learning paradigm, artificial neural networks, to

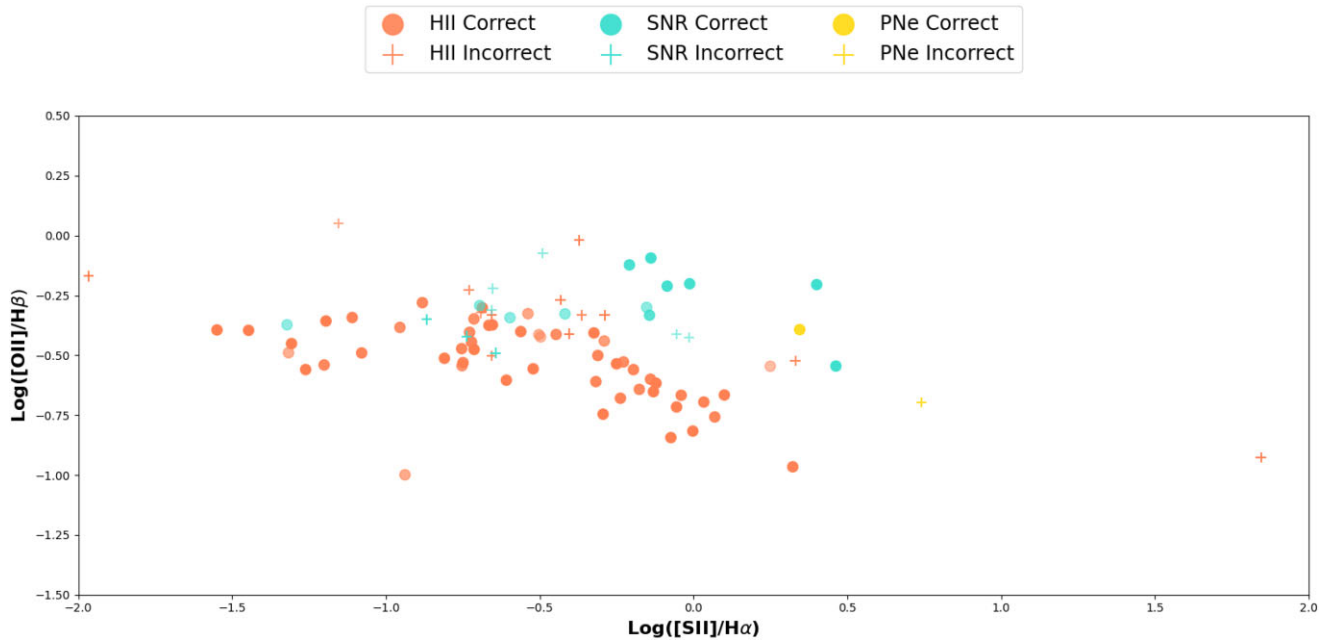


Figure 11. The line ratio diagram of $\log([\text{O III}]/\text{H}\beta)$ versus $\log([\text{S II}]/\text{H}\alpha)$ for the regions identified in a SITELLE field of M33. The emission region type is designated by differing colours. Additionally, the correct and incorrect classification (True and False in the legend) are indicated by circles and crosses, respectively. Note these classifications were made with the network taking four input variables instead of three. The points' opacity corresponds to the probability of the classification. The lowest probability value is approximately 60 per cent.

classify extragalactic emission regions into three categories: classic H II regions, planetary nebulae, and supernova remnants. Synthetic line ratios for each emission type are generated using line amplitudes taken from 3MdB. We train, validate, and test our network on approximately 90 000 synthetic data containing three key line ratios: $[\text{O III}] \lambda 5007/\text{H}\beta$, $[\text{N II}] \lambda 6583/\text{H}\alpha$, $([\text{S II}] \lambda 6717 + [\text{S II}] \lambda 6731)/\text{H}\alpha$. We also test the addition of a fourth line ratio: $[\text{O II}] \lambda 3726/\text{H}\beta$. We demonstrate the observational identification efficiency of this method by applying it to the Southwest field of M33; our results corroborate with existing literature with the exception of supernova remnants. We stress that these results are either due to an inconsistency between the training set and the real data, a common occurrence in supervised machine learning problems, incorrect identifications in the literature, or contamination from surrounding H II or DIG. In order to resolve this inconsistency, a training set that explores the same region of parameter space for SNRs is required; we note this currently does not exist. Although the method was created with SITELLE in mind, the results can readily be ported to other instruments as long as the line fluxes are recovered.

Additionally, this work highlights an important caveat that must be taken into account when applying machine learning algorithms to real astronomical data: without rigorous testing on real data, the results of the algorithm should be taken lightly. That is to say that it is inadvisable to use a machine learning algorithm that has not been verified thoroughly on real data to a new data set. As demonstrated here, even though the network does an excellent job (95 per cent) classifying supernova remnants, it fails to do so correctly in real data due to the issues described above.

Moreover, this work serves to expand the usage of machine learning in astronomy for classification purposes. Although we focused on three key categories of extragalactic emission line regions, this work can be extended to classify other objects in large catalogues of emission line regions (see Fig. 6).

Another important conclusion of this work remains the inherent dependence of our method to the photoionization models' ability to reproduce properly the natural properties and physics of the PNe, SNR, and H II regions. The photoionization conditions, including the wide varieties of ionizing sources, the morphology and characteristics of their surrounding ISM, and the accurate modelling of the physical interactions between them, need to be fully understood while constraining their space of parameters covered in Nature. Surveys of ionized regions in all conditions and at a high spatial resolution that include a good characterization of ionizing sources can help us to constrain the parameters while testing the photoionization models' performances.

Moreover, these large surveys, such as SIGNALS (Rousseau-Nepton et al. 2019) will enable us to amass a large quantity of well-resolved H II and supernova remnants. Since we will have an adequate spectral resolution from these surveys, precise flux measurements can be made. Overall, these surveys will enable the creation of a training set based on real observations instead of simulations. Although bias will always remain in any training set, whether with simulated or real data, this approach is expected to yield more accurate results. Furthermore, combining these large surveys with multiwavelength observations can ensure the correct classification of regions.

ACKNOWLEDGEMENTS

The authors would like to thank the Canada–France–Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii. The observations at the CFHT were performed with care and respect from the summit of Maunakea which is a significant cultural and historic site. CLR acknowledges financial support from the physics department of the Université de Montréal, the MITACS scholarship program, and the IVADO

doctoral excellence scholarship. JH-L acknowledges support from NSERC via the Discovery grant program, as well as the Canada Research Chair program. NVA acknowledges the support of the Royal Society and the Newton Fund via the award of a Royal Society–Newton Advanced Fellowship (grant NAF\R1\180403), and of Fundação de Amparo à Pesquisa e Inovação de Santa Catarina (FAPESC) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). CR is grateful to the Fonds de recherche du Québec – Nature et Technologies (FRQNT), for SIGNALS team financial support, and to the Natural Sciences and Engineering Research Council of Canada (NSERC) KG is supported by the Australian Research Council through the Discovery Early Career Researcher Award (DECRA) Fellowship DE220100766 funded by the Australian Government. K.G. is supported by the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project number CE170100013. The research of L. Chemin is funded by the Fondecyt Regular project 1210992 from Agencia Nacional de Investigación y Desarrollo de Chile.

Software: The following softwares were utilized in this work:

PYTHON (Van Rossum & Drake 2009), NUMPY (van der Walt et al. 2011), SCIPY (Virtanen et al. 2020), MATPLOTLIB (Hunter 2007), PANDAS (McKinney 2010), SEABORN (Waskom et al. 2017), (Robitaille et al. 2013), TENSORFLOW (Abadi et al. 2016), KERAS (Chollet 2015), LUCI (Rhea et al. 2021).

DATA AVAILABILITY

All data and methods used in this paper are available at [sitelinks:Pamplemousse](https://github.com/rastrai/signals:Pamplemousse).

REFERENCES

- Abadi M. et al., 2016, preprint ([arXiv:1603.04467](https://arxiv.org/abs/1603.04467))
- Alarie A., Morisset C., 2019, *Rev. Mex. Astron. Astrofis.*, 55, 377
- Amayo A., Delgado-Inglada G., Stasinska G., 2021, *MNRAS*, 505, 2361
- Asari N. V., Stasinska G., Morisset C., Fernandes R. C., 2016, *MNRAS*, 460, 1739
- Bacon R. et al., 2010, in McLean I. S., Ramsay S. K., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 7735, Ground-based and Airborne Instrumentation for Astronomy III. SPIE, Bellingham, p. 773508
- Baldwin J. A., Phillips M. M., Terlevich R., 1981, *PASP*, 93, 5
- Baril M., Grandmont F., Mandar J., Drissen L., Martin T., Rousseau-Nepton L., 2016, in Evans C. J., Simard L., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI. SPIE, Bellingham, p. 990829
- Baron D., 2019, preprint ([arXiv:astro-ph/1904.07248](https://arxiv.org/abs/1904.07248))
- Bertin E., 1994, in Epchtein N., Omont A., Burton B., Persi P., eds, Science with Astronomical Near-Infrared Sky Surveys. Springer, Netherlands, Dordrecht, p. 49
- Biswas M., Adlak R., 2018, in 2018 4th International Conference for Convergence in Technology (I2CT). IEEE, New York, p. 1
- Braun R., Walterbos R. A. M., 1992, *ApJ*, 386, 120
- Breiman L., 2001, *Mach. Learn.*, 45, 5
- Cardelli J. A., Clayton G. C., Mathis J. S., 1989, *ApJ*, 345, 245
- Cedr s B., Beckman J. E., Bongiovanni A., Cepa J., Ramos A. A., Giammanco C., Cabrera-Lavers A., Alfaro E., 2013, *ApJ*, 765, L24
- Chollet F., 2015, Keras. Available at: <https://keras.io>
- Ciardullo R., 2004, preprint ([arXiv:astro-ph/0407290](https://arxiv.org/abs/astro-ph/0407290))
- Cid Fernandes R., Carvalho M. S., S nchez S. F., de Amorim A., Ruschel-Dutra D., 2021, *MNRAS*, 502, 1386
- Constantin A., Vogeley M. S., 2006, *ApJ*, 650, 727
- D’Agostino J. J., Kewley L. J., Groves B. A., Medling A., Dopita M. A., Thomas A. D., 2019, *MNRAS*, 485, L38
- Danziger I. J., Dennefeld M., 1976, *PASP*, 88, 44
- Delgado-Inglada G., Morisset C., Stasinska G., 2014, *MNRAS*, 440, 536
- Delgado-Inglada G., Garc a-Rojas J., Stasinska G., Rechy-Garc a J. S., 2020, *MNRAS*, 498, 5367
- Della Bruna L. et al., 2020, *A&A*, 635, A134
- de Souza R. S. et al., 2017, *MNRAS*, 472, 2808
- Drissen L. et al., 2019, *MNRAS*, 485, 3930
- Ferland G. J. et al., 2017, *Rev. Mex. Astron. Astrofis.*, 53, 385
- Fesen R. A., Blair W. P., Kirshner R. P., 1985, *ApJ*, 292, 29
- Haffner L. M. et al., 2009, *Rev. Mod. Phys.*, 81, 969
- H nault F. et al., 2003, in Iye M., Moorwood A. F. M., eds, Proc. SPIE Conf. Ser. Vol. 4841, Instrument Design and Performance for Optical/Infrared Ground-based. SPIE, Bellingham, p. 12
- Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90
- Iben I., Jr, Tutukov A. V., 1984, *ApJS*, 54, 335
- Ivezic Z. et al., 2002, *AJ*, 124, 2364
- Kashino D., Inoue A. K., 2019, *MNRAS*, 486, 1053
- Kauffmann G. et al., 2003, *MNRAS*, 346, 1055
- Kennicutt R. C., Jr, 1984, *ApJ*, 287, 116
- Kennicutt R. C., Jr, Edgar B. K., Hodge P. W., 1989, *ApJ*, 337, 761
- Kewley L. J., Dopita M. A., 2002, *ApJS*, 142, 35
- Kewley L., Heisler C., Dopita M., 2001, *ApJS*, 132, 37
- Kewley L. J., Groves B., Kauffmann G., Heckman T., 2006, *MNRAS*, 372, 961
- Kewley L. J., Maier C., Yabe K., Ohta K., Akiyama M., Dopita M. A., Yuan T., 2013, *ApJ*, 774, L10
- Kewley L. J., Nicholls D. C., Sutherland R. S., 2019, *ARA&A*, 57, 511
- Kingma D. P., Ba J., 2017, preprint ([arXiv:1412.6980](https://arxiv.org/abs/1412.6980))
- Kreckel K., Groves B., Bigiel F., Blanc G. A., Kruijssen J. M. D., Hughes A., Schruha A., Schinnerer E., 2017, *ApJ*, 834, 174
- Lacerda E. A. D. et al., 2018, *MNRAS*, 474, 3727
- Long K. S., Blair W. P., Milisavljevic D., Raymond J. C., Winkler P. F., 2018, *ApJ*, 855, 140
- Lupton R., Blanton M. R., Fekete G., Hogg D. W., O’Mullane W., Szalay A., Wherry N., 2004, *PASP*, 116, 133
- MacAlpine G., Williams G., 1981, *ApJS*, 45, 113
- Martin T., Drissen L., 2017, preprint ([arXiv:astro-ph/1706.03230](https://arxiv.org/abs/astro-ph/1706.03230))
- Martin T. B., Drissen L., Melchior A.-L., 2018, *MNRAS*, 473, 4130
- McKinney W., 2010, in van der Walt S., Millman J., eds, Proceedings of the 9th Python in Science Conference. Scipy Publishing, Austin, Texas, p. 56. Available at: <http://conference.scipy.org/proceedings/scipy2010/mckinney.html>
- McLeod A. F. et al., 2021, *MNRAS*, 508, 5425
- Melnick J., Terlevich R., Moles M., 1987, *Rev. Mex. Astron. Astrofis.*, 14, 158
- Miller J., 1974, *ARA&A*, 12, 332
- Morisset C., Delgado-Inglada G., Flores-Fajardo N., 2015, *Rev. Mex. Astron. Astrofis.*, 51, 19
- Moumen I., Robert C., Devost D., Martin R. P., Rousseau-Nepton L., Drissen L., Martin T., 2019, *MNRAS*, 488, 803
- Moustakas J., Kennicutt R. C., Tremonti C. A., Dale D. A., Smith J.-D. T., Calzetti D., 2010, *ApJS*, 190, 233
- Oserbrock D., 1964, *ARA&A*, 2, 95
- Osterbrock D., Ferland G., 1989, *Astrophysics of Gaseous Nebulae and Active Galactic Nuclei*, 1st edn. Vol. 1. University Science Books, Sausalito, CA
- P rez-Montero E., Garc a-Benito R., V lchez J. M., 2019, *MNRAS*, 483, 3322
- Reynolds R. J., 1984, *ApJ*, 282, 191
- Rhea C. L., Rousseau-Nepton L., Prunet S., Hlavacek-Larrondo J., Fabbro S., 2020, *ApJ*, 901, 152
- Rhea C., Hlavacek-Larrondo J., Rousseau-Nepton L., Vigneron B., Guit  L.-S., 2021, *Res. Notes Am. Astron. Soc.*, 5, 208
- Robitaille T. P. et al., 2013, *A&A*, 558, A33
- Rousseau-Nepton L. et al., 2019, *MNRAS*, 489, 5530
- Salzer J. J., Jangren A., Gronwall C., Werk J. K., Chomiuk L. B., Caperton K. A., Melbourne J., McKinstry K., 2005, *AJ*, 130, 2584
- S rsic J. L., 1960, *Z. Astrophys.*, 50, 168

- Shatskiy A., Evgeniev I., 2019, *J. Exp. Theor. Phys.*, 128, 592
- Smith R. C., Kirshner R. P., Blair W. P., Long K. S., Winkler P. F., 1993, *ApJ*, 407, 564
- Stasinska G., Izotov Y., Morisset C., Guseva N., 2015, *A&A*, 576, A83
- Uzeirbegovic E., Geach J. E., Kaviraj S., 2020, preprint ([arXiv:astro-ph/2004.06734](https://arxiv.org/abs/2004.06734))
- Vale Asari N., 2021, Technical report, The role of the diffuse ionized gas in metallicity calibrations. Available at: <https://ui.adsabs.harvard.edu/abs/2021arXiv210800076V>.
- van der Walt S., Colbert S. C., Varoquaux G., 2011, *Comput. Sci. Eng.*, 13, 22
- Van Rossum G., Drake F. L., 2009, Python 3 Reference Manual. Create Soace, Scott's Valley, CA
- Veilleux S., Osterbrock D. E., 1987, *ApJS*, 63, 295
- Viallefond F., 1985, PhD Dissertation, Université de Paris, Paris
- Virtanen P. et al., 2020, *Nat. Methods*, 17, 261
- Walterbos R. A. M., Braun R., 1994, *ApJ*, 431, 156
- Waskom M. et al., 2017, *mwaskom/seaborn: v0.8.1* (September 2017). Available at: <https://zenodo.org/record/883859#.XvjkqXVKjJU>
- Winkler P. F., Blair W. P., Long K. S., 2017, VizieR Online Data Catalog, J/ApJ/839/83
- Woltjer L., 1972, *ARA&A*, 10, 129

- Zinchenko I. A., Dors O. L., Hagele G. F., Cardaci M. V., Krabbe A. C., 2019, *MNRAS*, 483, 1901

APPENDIX A: CRITERION FOR FILTERING DATA

In this section, we describe the filters applied to the 3MdB data set for reproducibility's sake. The python application of these filters can be found in the notebook entitled `generating_data.ipynb`. To filter the HII regions, we keep `lU_mean` values equal to `-2.5`, `-3.0`, and `-3.5`. We only retain `fr==3.0`. We only keep `ab_0` values between `-5.4` and `-3.0`. The only additional filtering we apply on the PNe data set is `com6==1`. For the shock models, we only take `emis_VI.model_type='shock'` and `abundances.name='Allen2008.Solar'`.

APPENDIX B: DIAGNOSTIC PLOTS OF TEST SET

In this section, we present the diagnostic plots of the test set data (as shown in Figs 1 and 2) broken into each type.

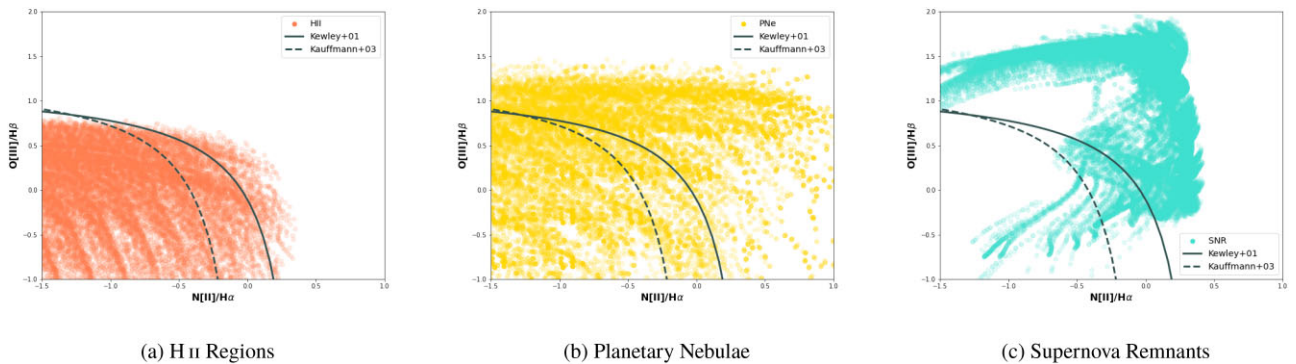


Figure B1. $[N\ II]/H\ \alpha$ versus $[O\ III]/H\ \beta$ Diagnostic plot broken down by emission region type.

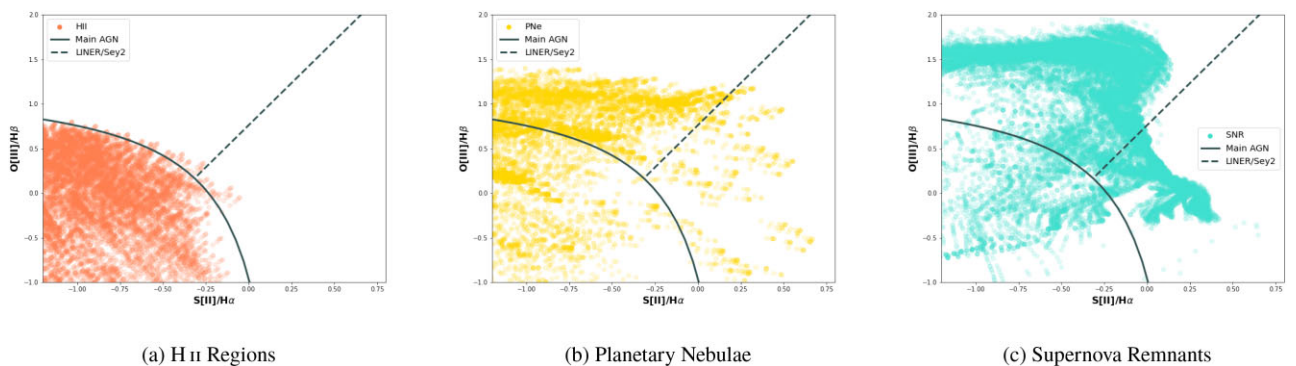
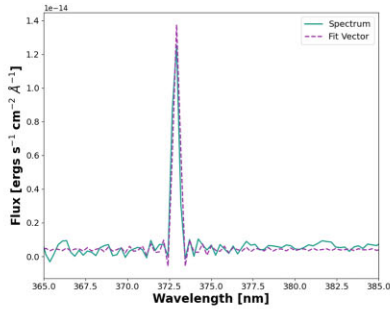


Figure B2. $[S\ II]/H\ \alpha$ versus $[O\ III]/H\ \beta$ Diagnostic plot broken down by emission region type.

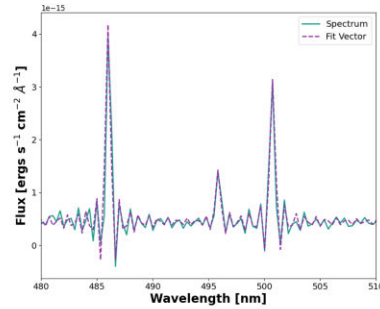
APPENDIX C: FIT PLOTS

This section contains three plots families (Figs C1, C2, and C3). Each family consists of the SN1, SN2, and SN3 fits of a emission

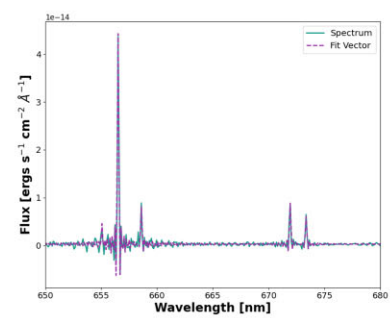
region type (i.e. H II, PNe, or SNR). Additionally, each family is further broken into the spectra which we correctly categorized and those incorrectly categorized.



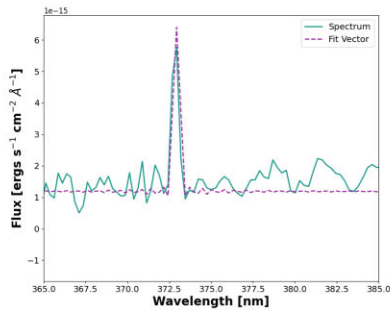
(a) SN1 spectrum and fitted vector for correctly categorized H II region



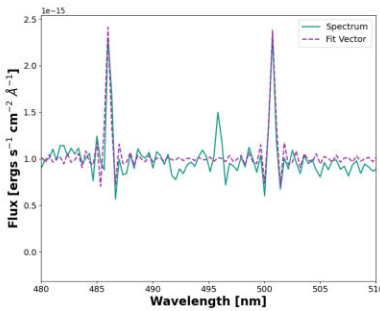
(b) SN2 spectrum and fitted vector for correctly categorized H II region



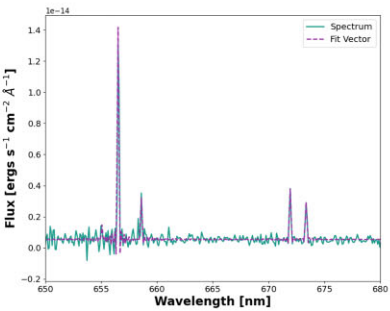
(c) SN3 spectrum and fitted vector for correctly categorized H II region



(d) SN1 spectrum and fitted vector for incorrectly categorized H II region

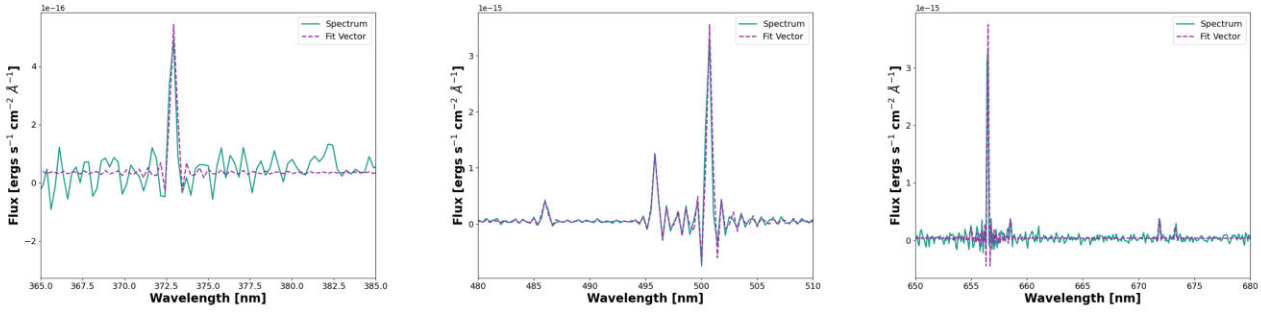


(e) SN2 spectrum and fitted vector for incorrectly categorized H II region

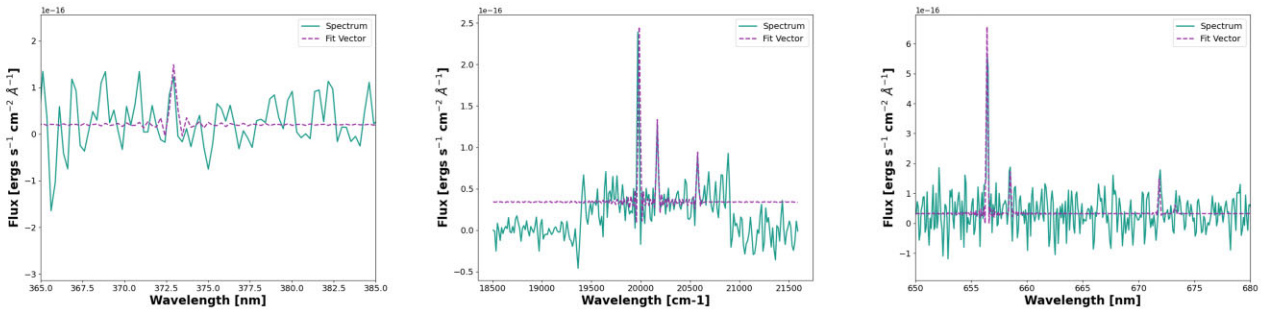


(f) SN3 spectrum and fitted vector for incorrectly categorized H II region

Figure C1. SN1, SN2, and SN3 of correctly categorized H II regions.

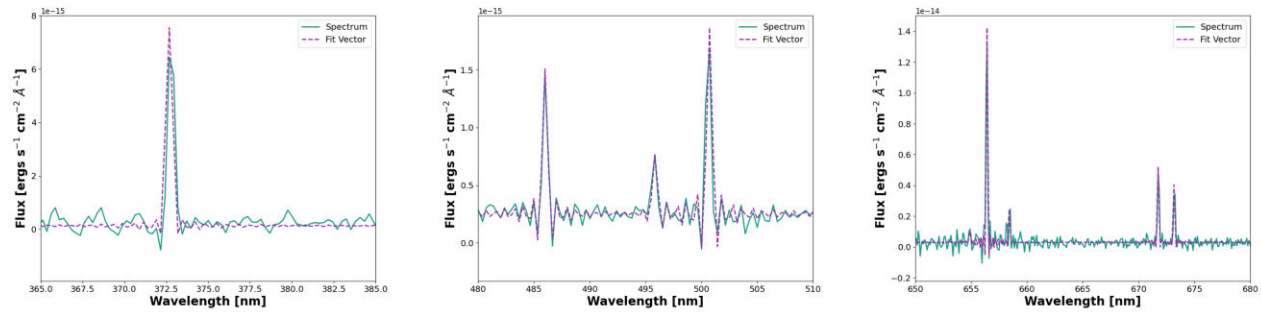


(a) SN1 spectrum and fitted vector for correctly categorized PNe region (b) SN2 spectrum and fitted vector for correctly categorized PNe region (c) SN3 spectrum and fitted vector for correctly categorized PNe region

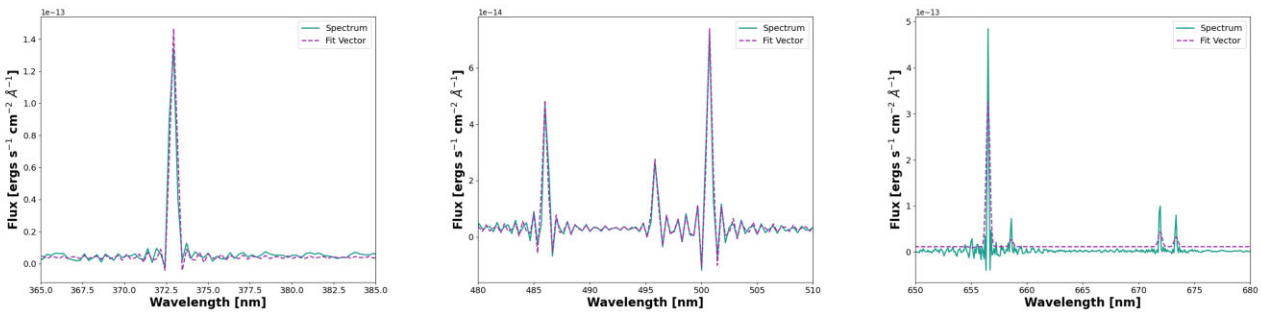


(d) SN1 spectrum and fitted vector for incorrectly categorized PNe region (e) SN2 spectrum and fitted vector for incorrectly categorized PNe region (f) SN3 spectrum and fitted vector for incorrectly categorized PNe region

Figure C2. SN1, SN2, and SN3 of the correctly and incorrectly categorized PNe regions. The region incorrectly classified was classified as a supernova remnant.



(a) SN1 spectrum and fitted vector for correctly categorized SNR region (b) SN2 spectrum and fitted vector for correctly categorized SNR region (c) SN3 spectrum and fitted vector for correctly categorized SNR region



(d) SN1 spectrum and fitted vector for incorrectly categorized SNR region (e) SN2 spectrum and fitted vector for incorrectly categorized SNR region (f) SN3 spectrum and fitted vector for incorrectly categorized SNR region

Figure C3. SN1, SN2, and SN3 of correctly and incorrectly categorized SNR regions. The region incorrectly classified was classified as an H II region.

APPENDIX D: LIST OF REGIONS

This section contains a table (Table D1) documenting the regions of M33 used in this study. The table includes the region type, the RA, the DEC, and radius. The table only reports regions that were not rejected on the basis of insufficient signal-to-noise of their spectra.

Table D1. Table of regions, their RA, DEC, and radius in arcseconds. The regions are shown in Fig. 7.

Region ID	RA	DEC	Radius
0	1:33:14.9493	+30:32:29.863	10.440 arcsec
2	1:33:12.2492	+30:30:24.375	5.606 arcsec
3	1:33:11.9795	+30:30:11.208	8.822 arcsec
4	1:33:13.8888	+30:29:45.757	5.438 arcsec
5	1:33:11.2340	+30:29:52.776	3.603 arcsec
6	1:33:09.6747	+30:29:45.723	8.170 arcsec
7	1:33:10.0807	+30:29:54.507	4.545 arcsec
8	1:33:12.4954	+30:34:08.952	10.957 arcsec
9	1:33:11.7219	+30:38:56.571	34.026 arcsec
10	1:33:13.5530	+30:39:32.587	2.502 arcsec
11	1:33:13.6216	+30:39:28.201	2.947 arcsec
15	1:33:06.1099	+30:31:02.774	2.000 arcsec
16	1:33:08.5500	+30:34:38.500	2.000 arcsec
22	1:33:11.0818	+30:34:22.629	16.226 arcsec
24	1:33:12.4386	+30:38:43.268	10.000 arcsec
25	1:33:14.1437	+30:39:48.311	12.684 arcsec
26	1:32:31.4769	+30:35:32.901	13.528 arcsec
27	1:32:46.7300	+30:34:37.800	7.099 arcsec
29	1:32:53.3331	+30:37:56.037	7.449 arcsec
30	1:32:57.6721	+30:39:27.962	6.705 arcsec
31	1:33:00.1500	+30:30:46.200	11.272 arcsec
32	1:33:00.6700	+30:30:59.300	6.911 arcsec
33	1:33:01.5100	+30:30:47.874	3.977 arcsec
34	1:33:02.9300	+30:32:28.737	8.102 arcsec
35	1:33:03.5700	+30:31:20.037	5.971 arcsec
36	1:33:04.0971	+30:39:58.016	8.697 arcsec
37	1:33:09.9369	+30:39:34.899	12.389 arcsec
38	1:33:11.1669	+30:39:43.699	12.952 arcsec
39	1:32:27.6494	+30:35:44.598	13.786 arcsec
40	1:32:35.4269	+30:35:19.800	15.386 arcsec
41	1:32:40.8064	+30:31:51.099	16.877 arcsec
42	1:32:42.7100	+30:36:19.237	9.503 arcsec
43	1:32:52.9336	+30:31:32.474	12.540 arcsec
44	1:32:56.0532	+30:33:30.400	12.923 arcsec
45	1:32:57.2469	+30:39:14.700	7.378 arcsec
54	1:32:45.6879	+30:38:55.637	22.822 arcsec
55	1:32:43.4136	+30:38:53.039	8.238 arcsec
56	1:32:56.9272	+30:40:18.512	8.238 arcsec
57	1:33:02.3461	+30:39:53.458	4.602 arcsec
58	1:33:01.4757	+30:39:29.292	4.602 arcsec
59	1:33:00.2709	+30:39:02.536	5.676 arcsec
60	1:33:07.2179	+30:35:12.876	5.583 arcsec
61	1:33:07.8867	+30:35:18.914	4.481 arcsec
62	1:32:58.6626	+30:36:37.518	5.583 arcsec
63	1:32:52.3103	+30:37:14.648	4.428 arcsec
64	1:32:54.1829	+30:37:28.459	2.852 arcsec
65	1:32:57.9253	+30:34:50.479	5.583 arcsec
66	1:32:57.7246	+30:34:41.847	5.583 arcsec
67	1:32:58.1255	+30:34:34.077	3.210 arcsec
68	1:32:52.7781	+30:34:55.667	11.849 arcsec
69	1:32:56.5215	+30:34:51.346	5.276 arcsec
70	1:32:52.9121	+30:36:36.665	31.885 arcsec
71	1:32:59.5295	+30:34:43.568	8.761 arcsec
72	1:33:00.8657	+30:34:18.529	11.849 arcsec

Table D1 – *continued*

Region ID	RA	DEC	Radius
73	1:33:00.9989	+30:34:00.400	5.126 arcsec
74	1:33:00.5310	+30:33:55.223	1.825 arcsec
75	1:33:04.4752	+30:34:21.100	3.567 arcsec
76	1:33:03.8734	+30:34:15.061	3.567 arcsec
77	1:33:04.4747	+30:34:07.288	6.463 arcsec
78	1:33:02.6696	+30:33:50.897	3.567 arcsec
79	1:33:03.4707	+30:33:22.406	6.862 arcsec
80	1:33:05.0067	+30:32:54.773	10.142 arcsec
81	1:33:02.8685	+30:32:58.238	6.560 arcsec
82	1:32:56.9878	+30:32:54.808	6.862 arcsec
83	1:32:51.7755	+30:33:04.309	6.553 arcsec
84	1:32:55.7178	+30:32:26.324	8.094 arcsec
86	1:33:04.7373	+30:31:55.212	5.268 arcsec
87	1:33:06.2730	+30:31:31.894	3.843 arcsec
88	1:33:06.6737	+30:31:27.575	2.922 arcsec
90	1:33:07.6752	+30:31:14.620	2.922 arcsec
91	1:33:07.6076	+30:30:57.355	4.015 arcsec
92	1:33:06.8067	+30:31:14.626	4.015 arcsec
93	1:33:04.8019	+30:30:56.511	3.832 arcsec
94	1:33:04.5342	+30:30:42.701	2.770 arcsec
95	1:33:08.2056	+30:29:50.051	14.323 arcsec
96	1:32:52.2431	+30:29:37.996	37.088 arcsec
97	1:32:45.3615	+30:31:56.103	4.646 arcsec
98	1:32:38.6850	+30:29:47.449	4.646 arcsec
99	1:32:34.6761	+30:30:21.086	11.742 arcsec
101	1:32:29.9250	+30:32:36.569	11.742 arcsec
102	1:32:31.5290	+30:32:32.269	5.541 arcsec
103	1:32:46.6289	+30:34:07.319	10.078 arcsec
104	1:32:44.5559	+30:34:53.927	7.709 arcsec
105	1:32:44.4886	+30:35:16.371	3.318 arcsec
106	1:32:43.5526	+30:35:16.367	3.341 arcsec
107	1:32:44.8228	+30:35:17.235	2.441 arcsec
109	1:32:44.2201	+30:36:02.985	7.709 arcsec
110	1:32:41.6795	+30:36:01.247	4.735 arcsec
112	1:32:29.5105	+30:36:08.058	18.767 arcsec
113	1:32:31.1856	+30:35:07.648	31.644 arcsec
115	1:32:42.5477	+30:36:34.918	3.341 arcsec
116	1:32:42.0797	+30:36:34.916	3.341 arcsec
117	1:32:55.7223	+30:39:30.173	3.590 arcsec
118	1:32:55.2541	+30:39:34.490	4.428 arcsec
119	1:32:47.0250	+30:39:47.435	2.393 arcsec
120	1:32:44.5502	+30:39:20.667	2.166 arcsec
123	1:32:53.1800	+30:38:55.646	2.517 arcsec
124	1:32:33.2680	+30:32:01.209	3.485 arcsec
125	1:32:28.4494	+30:33:59.424	2.922 arcsec
126	1:32:37.7140	+30:39:54.586	29.031 arcsec
128	1:33:00.9338	+30:35:05.144	17.306 arcsec

APPENDIX E: THE STRANGE CASE OF ID 27

Region 27, the supernova remnant, was misclassified despite the fact that the combined SIELLE deep image clearly shows an SNR. While the other misclassifications of supernova remnants are attributable to other causes (see discussion in Section 4), the reason for this incorrect categorization is not clear without examining the spectrum in SN3; the [S II] doublet and H α emission clearly demonstrate the presence of multiple components. Thus, the fluxes for these lines are incorrectly estimated which leads the network to classify it as an H II region. The presence of multiple lines is likely due to contamination from the surrounding DIG or H II region. We do not see these multiple components in the other filters since their spectral resolution is insufficient to resolve them. Although an in

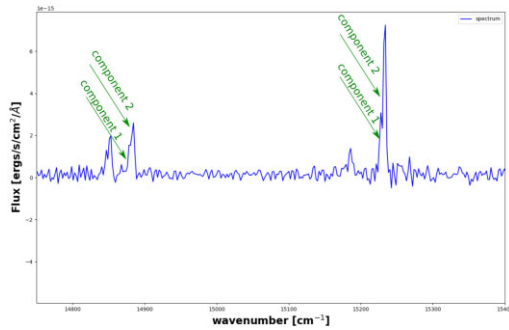


Figure E1. Zoom-in on the SN3 spectrum of region 27. We have noted the location of the double components with green arrows.

depth examination of this region is beyond the scope of this paper, we have attempted to quantify this effect in the following manner: (1) we fit two $H\alpha$ and $[S\text{II}]$ lines and (2) we artificially reduce the spectral resolution to 3000 (Fig. E1). Although we do not repeat this process for the seven incorrectly categorized SNRs that exhibit multiple components, region 27 serves as an example.

In fitting the two $H\alpha$ components and two $[S\text{II}]$ doublet components, we were able to distinguish emission lines from each component. We note that the flux of the first component of the $[S\text{II}]$ doublet and $H\alpha$ emission line is less than the second component. Moreover, the combined flux of the first $[S\text{II}]$ doublet component is greater than the $H\alpha$ emission line which indicates component 1 represents the SNR emission while component 2 represents the $H\text{II}$ region emission. This finding confirms our hypothesis that $H\text{II}$ region or DIG contamination is responsible for the misclassifications.

In artificially reducing the spectral resolution to 3000, we smear the double components into a single component. Although this clearly ignores the underlying astrophysics of the region, it allows for a more succinct fit of the region. We apply the network assuming these updated values for the SN3 emission lines; however, the network still classifies the region as an $H\text{II}$ region. This fortifies our position that contaminant emission from an overlapping $H\text{II}$ region or DIG is causing a misclassification of the supernova remnant.

The careful review of our results can help appreciate the level of contamination of the emission line spectra in galaxies even in the case where spatial resolution enables us to identify the complex structures in the ISM. Moreover, it demonstrates that the decontamination of spectra is not trivial and using dynamical properties such as Full Width Half Maximum (FWHM) and decomposition of the different velocity components of the ionized gas offer a new avenue to identify the well blended SNR emission nebulae.

APPENDIX F: SNR MISCLASSIFICATIONS

F1 Explanation of misclassifications

In addition to misclassifications of supernova remnants due to $H\text{II}$ or DIG contamination, a careful examination of their spectra revealed

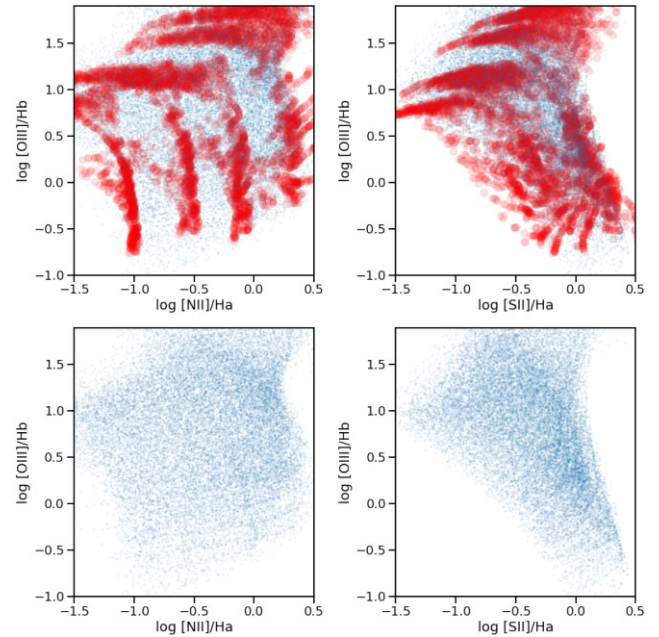


Figure F1. Line-ratio parameter space showing the initial data (in red) and the interpolated data (in blue).

additional reasons for the incorrect classifications. Two SN2 spectra (IDs 102 and 125) demonstrate a high level of noise not present in the SN3 filter (where our signal-to-noise ratio threshold was applied); the misclassifications are likely due to the poor constraints on $[O\text{III}]$ and $H\beta$ due to the noise. An additional three SN2 spectra (IDs 58, 75, and 98) reveal a complete lack of $[O\text{III}]\lambda 5007$ emission; again, this explains the misclassification by the network. Therefore, if we apply an additional signal-to-noise threshold to SN1 and SN2, we have only three misclassified supernova remnants instead of 8. This brings the accuracy from 64 per cent to 82 per cent. Similarly, the misclassified planetary nebula would be removed.

F2 Interpolating over the SNR grid

As discussed in Section 4, the SNR grid from the 3MdB shock extension has gaps in the line-ratio parameter space. In Fig. F1, we show the initial data (in red) and the interpolated data (in blue). The top two figures juxtapose the data initially collected from the 3MdB shock extension in red and the interpolated data in blue in the two standard BPT diagrams. The bottom panels show only the interpolated data points. The figure demonstrates that our linear interpolation scheme covers the entire parameter space.

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.