1D Separable Convolutional Neural Network Architecture for Stellar Classification Based on Spectral Characteristics

Jorge Felix Martínez Pazos¹, Jorge Gulín Gonzales², David Batard Lorenzo², Ariel Ramirez Alvarez³, and Yunwei Chen⁴,⁵

¹Affiliation not available
²Study Center on Computational Mathematics (CEMC), University of Informatics Science
³University of Informatics Science
⁴Scientometrics & Evaluation Research Center (SERC), National Science Library (Chengdu), Chinese Academy of Science, Sichuan
⁵Department of Information Resources Management, School of Economics and Management, University of Chinese Academy of Sciences

June 14, 2024
1D Separable Convolutional Neural Network Architecture for Stellar Classification Based on Spectral Characteristics

Jorge Felix Martínez Pazos¹, Jorge Gulín Gonzales¹, David Batard Lorenzo¹, Ariel Ramirez Alvarez², Yunwei Chen³,4

¹ Study Center on Computational Mathematics (CEMC). University of Informatics Science, Havana, Cuba.
² University of Informatics Science, Havana, Cuba.
³ Scientometrics & Evaluation Research Center (SERC), National Science Library (Chengdu), Chinese Academy of Science, Sichuan, China.
⁴ Department of Information Resources Management, School of Economics and Management, University of Chinese Academy of Sciences, Beijing, China.

*Address correspondence to: jorgefmp.mle@gmail.com

Abstract
Artificial Intelligence has enabled scientists to rapidly sift through and analyze massive collections of images, helping to identify objects worthy of closer study such as supernovae, pulsars, and quasars, and allowing us to classify stars, label galaxies, and other astronomical systems. Stellar classification serves as a cornerstone of astronomy, providing a framework for understanding and characterizing the diversity of celestial objects. This classification is based on the spectral properties. AI algorithms, specifically machine learning and deep learning models, have been designed to learn from data and make predictions or classifications based on patterns and trends. These models of spectral and stellar classification are trained on large datasets of known celestial objects, enabling them to recognize and categorize new objects based on their spectral signatures. Here we present a novel Fine Tuned Deep Convolutional Neural Network of 1D Separable Convolutional blocks for stellar classification based on spectral characteristics using SSDS-17 data captured by the Sloan Digital Sky Survey, where the class imbalance is assessed using the SMOTE balancing technique. The results obtained during the performance evaluation confirmed the reliability of the proposed architecture of StellarNet in multi-class stellar classification, obtaining remarkable values of around 97% and 99% for accuracy and AUC score respectively. The proposed StellarNet architecture can be applied to enable real-time classification of captured spectral characteristic data, facilitating automation of the task and providing labeled data for study and future research.

Keywords: Artificial Intelligence, Spectral Characteristics, Stellar classification, Deep Convolutional Neural Network

Figure 1: Workflow of the Proposed Approach for Addressing Data Imbalance Using Min Class Balance and SMOTE Balance, and Comparative Evaluation of the Designed, Trained, and Fine-Tuned StellarNet Architectures
1 Introduction

The advent of artificial intelligence (AI) has been a game changer in astronomy and astrophysics [1,2]. AI has enabled scientists to rapidly sift through and analyze vast collections of images, helping to identify objects worthy of closer study such as supernovae, pulsars, and quasars, and allowing us to classify stars, label galaxies, and evaluate redshifts [3]. One of the most important contributions of AI to this field is to support the development and use of tools such as the James Webb Space Telescope (JWST). The JWST, the largest and most powerful space telescope ever built, will break new ground in many areas of astronomical research [4]. Blackbody emitters output a pattern of electromagnetic waves with an uneven distribution of intensities for different wavelengths, called spectra. Within this pattern there is much information: the wavelength of its peak informs of the body’s temperature by Wien’s law; absorption lines that are unique for each element can show the composition of the blackbody and the position of these lines within the pattern can be shifted depending on the relative velocity of the object [5,6]. This shift (usually towards infrared wavelengths) is caused by the Doppler’s effect where velocity affects the frequency of waves [7]. Stellar classification serves as a cornerstone of astronomy, providing a framework for understanding and characterizing the diversity of celestial objects. This classification is based on the spectral properties which refer to the lines and bands that appear in the spectrum of an astronomical object, and is produced by the absorption or emission of light at specific wavelengths by the chemical elements present in the object [8,9]. Photometric filters, on the other hand, are tools used in astronomical observations to isolate and measure the intensity of light in specific wavelength ranges. Photometric filters commonly used in stellar classification include i, r, z, k, and u [10]. Redshift is a measure of how much the light from an astronomical object has shifted toward the red due to the expansion of the universe. This phenomenon is crucial in determining the distance of distant astronomical objects [11]. The terms “alpha” and “delta” are used in astronomy to indicate right ascension and declination, respectively, which are the two coordinates used to specify the position of an object in the sky in the equatorial coordinate system [12].

When studying stars, the Harvard spectral classification scheme, developed in the late 1800s and refined by Annie Jump Cannon in 1924, is one such system that classifies stars based on their temperature, which is determined by the strength of the hydrogen lines present in their spectra [13]. This classification scheme has been instrumental in our understanding of stellar evolution [14]. Quasars are a class of active galactic nuclei (AGN). They are powered by the accretion of matter onto a supermassive black hole surrounded by an accretion disk [15]. Most quasar spectra from ultraviolet to optical wavelengths can be characterized by a featureless continuum and a series of mostly broad emission line features; compared with galaxies or stars, these spectra are remarkably similar from one quasar to another. A quasar’s spectra have a larger redshift than stars. Galaxies are formed by stars, interstellar gas, and dust, which can occupy hundreds of thousands of light-years. Its spectral characteristics will be a result of the combination of the spectral information of millions of stars and the absorption lines from its gas and dust [16,17]. The importance of stellar classification and spectral characteristics in astronomy cannot be overstated, since they serve as a primary task for the study of celestial objects [18]. The spectral characteristics of stars, galaxies, and quasars provide a wealth of information about their physical properties and the large-scale structure of the Universe [18]. However, with the advent of large sky survey projects such as the Sloan Digital Sky Survey (SDSS) and the Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST), vast amounts of spectral data are being generated [19]. Manual analysis of this data is not feasible due to its sheer volume, hence the need for Artificial Intelligence (AI) [17].

AI algorithms, specifically machine learning and deep learning models, are designed to learn from data and make predictions or classifications based on patterns and trends. In the context of spectral and stellar classification, these models are trained on large datasets of known celestial objects, enabling them to recognize and categorize new objects based on their spectral signatures. Zhao et al. (2022) [19] developed a robust ensemble convolutional neural network (ECNN) for the classification of massive stellar spectra. The methodology involved the design of six distinct convolutional neural networks (CNNs) as classifiers to recognize the spectra in DR16. These classifiers were then integrated in an ensemble learning manner, based on the cross-entropy testing error of the spectra at different signal-to-noise ratios. This innovative approach resulted in a one-dimensional ECNN strategy that achieved a classification accuracy of 95.0% for stellar spectra, surpassing the accuracy of traditional methods such as principal component analysis and support vector machine models. A study conducted by Li, Lin, and Qiu (2019) [20] addressed the challenge of efficiently and accurately handling large amounts of spectral data by focusing on classifying stellar spectra, assuming the absence of perfect absolute flux calibration, a scenario that arises when considering spectra from the Guo Shou Jing Telescope (also known as the Large Sky Area Multi-Object Fiber
Spectroscopic Telescope, LAMOST). Their proposal involves two key techniques: first, spectrum normalization based on a seventeenth-order polynomial fit; second, a random forest (RF) classification of the stellar spectra. The experiments conducted on four stellar spectral libraries demonstrated the effectiveness of the RF in classifying stellar spectra. Tao et al. (2018) [21] introduced an automated method for galaxy spectral classification through machine learning. Utilizing a dataset of 10,000 galaxy spectra from SDSS DR14, they applied algorithms such as logistic regression, random forest, and linear SVM to improve the efficiency of galaxy spectral classification and aid in the study of galaxy properties and evolution.

Spectral characteristics hold a central position in the realm of astronomical research, serving as a crucial factor in the classification of celestial entities such as stars, galaxies, and quasars. The integration of Artificial Intelligence (AI) has instigated a significant transformation in this field. It has automated the classification process, thereby facilitating the efficient management of the extensive data generated by contemporary astronomical surveys. This synergy between astrophysics and AI is paving the way for a more streamlined exploration of our universe.

The following is an outline of the contributions that have been made to draw attention to the relevance of the work that will be presented in this study:

- StellarNet is the main contribution of this proposal, a novel Deep Convolutional Neural Network of 1D Separable Convolutional blocks for stellar classification based on spectral characteristics.

- A comparative analysis was conducted between Minority Class Balance and the Synthetic Minority Over-sampling Technique (SMOTE), confirming the superiority of SMOTE for tasks similar to the one evaluated, thereby highlighting its effectiveness in addressing class imbalance issues.

- The results obtained during the performance evaluation confirmed the reliability of the proposed architecture of StellarNet in multi-class stellar classification, obtaining remarkable values of around 97% for each metric evaluated.

- The proposed StellarNet architecture can be applied to enable real-time classification of captured spectral characteristic data, facilitating automation of the task and providing labeled data for study and future research.

2 Materials & Methods

2.1 Dataset

In this study was used the Stellar Classification Dataset - SDSS17, which comprises 100,000 space observations captured by the Sloan Digital Sky Survey (SDSS). Each observation is characterized by 17 feature columns and one class column that classifies it as a star, galaxy, or quasar. The Stellar Classification Dataset - SDSS17, utilized in the project, is a comprehensive collection of data that provides a wealth of information about celestial objects. This dataset is derived from the Sloan Digital Sky Survey (SDSS), which has made use of a dedicated 2.5m wide-angle optical telescope to capture images of more than a quarter of the sky. The dataset comprises several features, each providing unique insights into the objects being studied. The pobj_ID serves as a unique identifier for each object in the image catalog used by the Catalog Archive Server (CAS). The alpha and delta represent the Right Ascension and Declination angles respectively, providing the precise location of the object in the celestial sphere. The dataset also includes photometric data captured through various filters: u (Ultraviolet), g (Green), r (Red), i (Near Infrared), and z (Infrared) [22, 23]. These photometric measurements provide a spectrum of light intensities, which are crucial in determining the physical properties of these celestial objects. The redshift value is another significant feature in this dataset. It measures the shift in wavelength due to the Doppler effect, which can be used to calculate the distance and velocity of an object relative to the observer. For this particular study, only spectral characteristics (alpha, delta, u, g, r, i, z, and redshift) were considered. Other attributes such as run_ID, rerun_ID, cam_col, field_ID, spec_obj_ID are dismissed since they are only identifiers and don’t add domain knowledge.

2.2 Exploratory Data Analysis

Analysis of the distribution of the dataset reveals a significant imbalance between classes. Specifically, the dataset contains 59,445 tuples for GALAXY, 21,594 for STAR, and 18,961 for QSO. These numbers represent 59.4% of
the dataset for GALAXY and only 19.0% for QSO, as shown in Figure 2. Training a model on such a highly unbalanced dataset could lead to biased results. Therefore, data balancing techniques are employed to mitigate this problem. Two specific approaches are investigated to determine which provides superior results: simple minimum class balance and the synthetic minority oversampling technique (SMOTE). These methods are used to ensure a more balanced representation of classes in the dataset, thereby increasing the reliability of the model’s predictions.

Figure 2: SSDS17 Classes Distribution

When analyzing the correlation of the spectral characteristics, it was observed that there are several instances of correlation among the spectral characteristics. These correlations can be divided into two different domains, the first domain is characterized by a maximum positive correlation with a value of 1, which includes the spectral features $g$, $u$, and $z$. The second domain is characterized by a high positive correlation, with values ranging from 0.43 to 0.96, encompassing the spectral features $i$, $r$, and redshift. These correlations are visually represented in Figure 3, where they can be observed in the correlation heatmap, providing a comprehensive overview of the relationships between these spectral features, thereby facilitating a deeper understanding of their interactions in the context of the study.

Figure 3: SSDS17 Correlational Heatmap of Spectral Characteristics
2.3 Balancing Dataset

In the pursuit of a balanced dataset, the minimum class balance approach was adopted to limit the size of all classes larger than 20,000 tuples to the size of the smallest class containing over 18,000 tuples. While this approach ensures a high degree of balance across all classes, it is not without its drawbacks. In particular, it results in the loss of approximately 50% of the data set. This significant reduction in data could potentially have a negative impact on the robustness and generalizability of models trained on this dataset. Therefore, it is crucial to weigh the benefits of class balancing against the potential loss of information in this context.

In response to the data reduction strategy previously discussed, an alternative approach was implemented to balance the dataset without discarding data. The Synthetic Minority Oversampling Technique (SMOTE) was applied, which is a widely recognized method for addressing class imbalance in machine learning [24]. This imbalance can significantly degrade the performance of predictive models and is characterized by a severe disproportion in the distribution of classes within a dataset.

The SMOTE algorithm addresses this issue by generating synthetic instances of the minority class. The procedure involves selecting a vector from the minority class and identifying its k nearest neighbors. A synthetic point is then interpolated along the line segment connecting the chosen vector and one of its neighbors. This process is iteratively performed until a balance between the classes is achieved. By applying SMOTE, the dataset is balanced not by losing data but by generating more synthetic data. This approach ensures that no information is lost, potentially enhancing the robustness and generalizability of models trained on this balanced dataset.

The following Table 1 provides a detailed comparison of the datasets balanced using the Minimum Class Balance method and the Synthetic Minority Oversampling Technique (SMOTE). Further research and experimentation will be conducted on both datasets to determine which approach produces more favorable results.

<table>
<thead>
<tr>
<th>Class</th>
<th>Amount</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>GALAXY</td>
<td>20000</td>
<td>33.9</td>
</tr>
<tr>
<td>QSO</td>
<td>18961</td>
<td>32.2</td>
</tr>
<tr>
<td>STAR</td>
<td>20000</td>
<td>33.9</td>
</tr>
<tr>
<td>Total</td>
<td>58961</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Class</th>
<th>Amount</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>GALAXY</td>
<td>59445</td>
<td>33.3</td>
</tr>
<tr>
<td>QSO</td>
<td>59445</td>
<td>33.3</td>
</tr>
<tr>
<td>STAR</td>
<td>59445</td>
<td>33.3</td>
</tr>
<tr>
<td>Total</td>
<td>178335</td>
<td></td>
</tr>
</tbody>
</table>

2.4 StellarNet Architecture

The features of the dataset (alpha, delta, u, g, r, i, z, redshift) were subjected to min-max normalization, transforming them to a scale of 0-1. This process improves computational processing and prevents model bias due to significant numerical differences. Meanwhile, the target (class) was processed using one hot encode to convert classes to vectors, which is beneficial for predictive models because it allows the model to understand categorical data as a form of numerical input. Applying min-max normalization and one-hot encoding techniques ensures efficient computational processing, prevents model bias, and allows the model to better understand the input data.

The data sets were divided into three subsets: 40% for training, 30% for testing, and the remaining 30% for validation. The distribution of these subsets is shown in Table 2 for both the min-class balanced dataset and the smote balanced dataset. A significant portion of the data was reserved for testing and validation to evaluate the model’s performance on unseen data, thereby providing an estimate of its potential real-world applicability and effectiveness. This approach underscores the importance of rigorous model validation in predictive analytics to ensure that the models developed are robust and reliable in different scenarios.
Table 2: Train, Test & Validation Distribution for Both Datasets

<table>
<thead>
<tr>
<th>Min. Class Balance</th>
<th>Set</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Train</td>
<td>35376</td>
</tr>
<tr>
<td></td>
<td>Test</td>
<td>11793</td>
</tr>
<tr>
<td></td>
<td>Validation</td>
<td>11792</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SMOTE Balance</th>
<th>Set</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Train</td>
<td>107001</td>
</tr>
<tr>
<td></td>
<td>Test</td>
<td>35667</td>
</tr>
<tr>
<td></td>
<td>Validation</td>
<td>35667</td>
</tr>
</tbody>
</table>

A variety of approaches have been adopted in star classification, including machine learning techniques such as decision trees, support vector machines, and random forests, all of which are widely used. At the same time, dense neural networks within deep learning strategies have been used to address similar challenges. The goal of this research is to design a 1D Deep Convolutional Neural Network, consisting of separable convolutional blocks, to perform the task of stellar classification.

The StellarNet architecture, a 1D Convolutional Neural Network with Separable Convolutions, is systematically organized into three main flows: the input flow, mid flow, and output flow these are detailed in Figure 4.

**Input Flow:** The architecture begins with an input layer designed to accommodate data that matches the shape of the feature set. This is followed by a SeparableConv1D layer, which is configured with 32 filters, a kernel size of 2, an 'elu' activation function, and padding set to 'same'. The output from this layer is then processed through a MaxPooling1D layer with a pool size of 2, followed by a BatchNormalization layer.

**Mid Flow:** The mid flow consists of two distinct blocks of layers. Each block incorporates a SeparableConv1D layer (with 64 and 128 filters respectively, a kernel size of 2 for the first block and 1 for the second block, and an elu activation function), succeeded by a MaxPooling1D layer (pool size 2), and a BatchNormalization layer. Notably, the second block also integrates a Dropout layer with a dropout rate selected from [0.1, 0.2, 0.25].

**Output Flow:** The final phase, or the output flow, initiates with a GlobalAveragePooling1D layer. This is followed by two Dense layers: the first with units selected from [256, 512, 1024, 2048] and an elu activation function; the second with units selected from [32, 64, 128] and also an elu activation function. A Dropout layer with a rate selected from [0.7, 0.8] is applied preceding the final Dense output layer. This output layer consists of 3 units (corresponding to the three classes: GALAXY, STAR, QSO) and uses a softmax activation function to return the class probability vector.
The model is compiled using the Adam optimizer with a learning rate selected from $[0.003, 0.0003]$, categorical cross-entropy as the loss function, and accuracy as the evaluation metric. StellarNet has a total of 52,869 parameters, of which 52,421 are trainable and 448 are not. Optimal values for dropouts, dense units, and learning rate of the architecture are determined by hyper-parameter tuning using Keras Tuner. The random search of hyperparameters was performed using both training set and validation set, with a search space of 5 due to each defined variable, the max trials parameters is set to 25, even if this value does not match the total parameters combination is a very good value to try parameter combinations without such a large processing time. The hyperparameter search is performed for 30 training epochs, table 3 summarizes hyperparameter search.

<table>
<thead>
<tr>
<th>Hyperparameter</th>
<th>Best Value</th>
<th>Search Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dense Unit</td>
<td>1024</td>
<td>[256, 512, 1024, 2048]</td>
</tr>
<tr>
<td>FDense Unit</td>
<td>64</td>
<td>[32, 64, 128]</td>
</tr>
<tr>
<td>Dropout</td>
<td>0.25</td>
<td>[0.1, 0.2, 0.25]</td>
</tr>
<tr>
<td>FDropout</td>
<td>0.7</td>
<td>[0.7, 0.8]</td>
</tr>
<tr>
<td>lr</td>
<td>0.00003</td>
<td>[0.003, 0.0003]</td>
</tr>
</tbody>
</table>

Total Search Space: 144

For both approaches of StellarNet four callbacks were used during the training process: early stop, learning rate reduction, tensorboard, and checkpoint. The early stop callback stops the training process when no further improvement in the validation accuracy metric is observed with a patience value of 25. The learning rate reduction callback reduces the model's learning rate, initially set at 3e-4, by monitoring the loss on validation data, thereby improving the model's optimization and accuracy on validation data. Tensorboard, a tool included in the Tensorflow framework, provides real-time visualization of all variables and model behavior, facilitating hyperparameter optimization. Checkpoint callback allows the preservation of the model's weights and biases at a given state during training; in this case, it saves each time the model improves based on the validation accuracy metric.

The training process was performed over 250 epochs using both the training and validation sets with a batch size of 1024, which greatly accelerated the process. Techniques such as learning rate reduction and dropout layer were used.
in conjunction with real-time hyperparameter monitoring via Tensorboard to prevent overfitting of the proposed model. The process was conducted on an Intel(R) Core(TM) i7-12700H computer with a 2.30GHz CPU, an NVIDIA GeForce RTX 3060 8GB GPU, and 16GB RAM DDR5.

3 Results & Discussion

The early stop callback stopped the process at epoch 76 for each model after both models had been trained for 250 epochs. This occurred because the set patience value was exceeded without any improvement in the model based on the defined metrics. At this point, both models reported fairly similar accuracy values of approximately 0.97 for training and validation, with the Min Class approach slightly lower by 0.01. Figures 4 and 5 illustrate the behavior of the metrics during each epoch for both approaches, specifically Loss, Accuracy, Validation Loss, Validation Accuracy, and Learning Rate Reduction.

![Figure 4: Metrics behavior throughout the training process of StellarNet with SMOTE Balance.](image1)

![Figure 5: Metrics behavior throughout the training process of StellarNet with Min Class Balance.](image2)
Both StellarNets models exhibit robust generalization capabilities, as evidenced by their consistent performance across the training and validation datasets. The model using Min Class Balance is slightly less accurate, but given that the other approach, SMOTE Balance, is nearly twice as accurate and achieves slightly better results, it is considered superior between the two. This comparative analysis highlights the effectiveness of the SMOTE Balance approach in improving model accuracy and providing robust performance. After training is complete, the significant impact of the learning rate reduction callback on the model’s performance on the validation data is evident. It also illustrates how the learning rate reduction stabilizes the model, keeping it close to the optimal metrics. Reducing the learning rate also results in a smaller amplitude between the training and validation curves. This is because a lower learning rate in the gradient descent implemented by the Adam optimizer results in smaller shifts in the search for the global minimum. As a result, it can locate this minimum or values close to it more effectively, with less chance of deviating to less favorable situations. The optimal values for each metric are reached at epoch 76, which is the model configuration saved by the checkpoint callback. This highlights the effectiveness of our approach, which combines advanced techniques such as learning rate reduction and early stopping to fine-tune the model and prevent overfitting. This ensures that StellarNets are not only accurate, but also robust and able to generalize well to unseen data. In addition, StellarNets performance was evaluated and compared across the three sets using accuracy, loss, confusion matrix, AUC score, and ROC curve.

The confusion matrices derived from both approaches show high-performance values. The diagonal entries of these matrices correspond to correctly predicted values, while the off-diagonal elements represent instances of incorrect predictions. As shown in the following figures, both models show a remarkably low misclassification rate. However, there is a notable difference when comparing the two models and the significant difference in data length. The difference in the amount of misclassified data between the two approaches is almost similar, although there is a significant gap in the amount of data from the SMOTE approach compared to the Min Class Balance approach, as the SMOTE sets are more than 3 times larger. This supports the claim that the SMOTE model is more effective and robust than the Min Class Balance model.

On the other hand, both models have certain peculiarities. They are more sensitive to the correct classification of the star class. It seems that the spectral properties of stars differ more from galaxies and quasars than from each other, allowing the models to identify them more effectively.

Furthermore, an examination of the heatmaps of the confusion matrices reveals notable differences in the classification performance of the StellarNet model under the two strategies. The quantity of correctly predicted values for the QSO class is more discernible in the heatmap due to the slightly lighter shade of the corresponding square, indicating that the Minimum Class Balance strategy exhibits a lower classification performance for the QSO class. This is not the case for the Synthetic Minority Over-sampling Technique (SMOTE) strategy, where all squares corresponding to correctly predicted classes exhibit a considerably darker shade, signifying near-maximum quantities. The aforementioned analysis can be substantiated by referring to Figures 6 and 7, which provide visual representations of the discussed findings.

Figure 6: Confusion Matrices for SMOTE StellarNet
To delve into the intricacies of the StellarNet architecture, the Receiver Operating Characteristic (ROC) curve for One vs. Rest is shown in figures 8 and 9, accompanied by the Area Under the Curve (AUC) score for both asset strategies. The AUC score serves as a performance metric for the model per domain class, assessing the trade-off between true positives and false positives. The One vs. Rest (OvR) method is used to adapt the ROC curve and AUC score metrics, originally designed for binary classification, to multi-class classification scenarios [25, 26].

As shown in the figures, StellarNet exhibits superior AUC values for the STAR class across the training, validation, and test sets, with an approximate value of 0.9980, approaching the maximum achievable value of 1. Overall, the model exhibits exceptional performance, with the lowest AUC values of 0.9908 and 0.9886 for the OvR in the test set for the Synthetic Minority Over-sampling Technique (SMOTE) and Minimum Class Balance strategies, respectively. When examining the ROC curve and the AUC value, it was found that the results derived from the two strategies did not show significant differences. However, the STAR class performed slightly better in both cases, with the difference being slightly noticeable when comparing the AUC and ROC curves, as well as when examining the confusion matrix.

Figure 7: Confusion Matrices for Min Class StellarNet

Figure 8: One-vs-Rest (OvR) Receiver Operating Characteristic (ROC) curves, and corresponding Area Under the Curve (AUC) scores for SMOTE strategy.
Figure 9: One-vs-Rest (OvR) Receiver Operating Characteristic (ROC) curves, and corresponding Area Under the Curve (AUC) scores for Min Class Balance strategy.

Table 3 provides a detailed analysis and comparison of the performance of the StellarNet model using two different strategies: Synthetic Minority Over-sampling Technique (SMOTE) and Minimum Class Balance. A cursory examination of the results reveals a slight superiority of the SMOTE approach in terms of performance. However, when considering the significant differences in data volume between the two strategies, this marginal superiority is amplified. This observation underscores the effectiveness of the SMOTE approach in handling unbalanced data sets, thereby improving the overall performance of the StellarNet model. In Table 3, the columns labeled (TP+TN) represent the sum of true positives and true negatives, which together represent the total number of correct classifications from the entire dataset. Conversely, the columns labeled (FP+FN) represent the sum of false positives and false negatives, which together denote the total number of incorrect classifications from the entire dataset, providing a comprehensive overview of the classification performance of the model.

<table>
<thead>
<tr>
<th>Min. Class Balance</th>
<th>Set</th>
<th>Accuracy</th>
<th>AUC</th>
<th>TP+TN</th>
<th>FP+FN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train</td>
<td>0.968</td>
<td>0.995</td>
<td>34242</td>
<td>1134</td>
<td></td>
</tr>
<tr>
<td>Test</td>
<td>0.965</td>
<td>0.993</td>
<td>11376</td>
<td>417</td>
<td></td>
</tr>
<tr>
<td>Validation</td>
<td>0.965</td>
<td>0.993</td>
<td>11373</td>
<td>419</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SMOTE Balance</th>
<th>Set</th>
<th>Accuracy</th>
<th>AUC</th>
<th>TP+TN</th>
<th>FP+FN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train</td>
<td>0.970</td>
<td>0.995</td>
<td>104243</td>
<td>2758</td>
<td></td>
</tr>
<tr>
<td>Test</td>
<td>0.973</td>
<td>0.995</td>
<td>34719</td>
<td>948</td>
<td></td>
</tr>
<tr>
<td>Validation</td>
<td>0.974</td>
<td>0.995</td>
<td>34733</td>
<td>934</td>
<td></td>
</tr>
</tbody>
</table>

The values are highlighted in bold to emphasize superior performance.

The proposed architecture of the 1D separable convolutional neural network, combined with its regularization techniques, allows StellarNet to achieve robustness in various class balance scenarios. This architecture, combined with an extensive hyperparameter search space of 144 and the strategic use of callbacks during the training phase, significantly increases the learning efficiency of the network, largely due to the ability to define, customize, adjust, and fine-tune the learning process to meet the specific requirements of the task at hand; such a level of specificity in the learning process is a key factor in optimizing the network's performance.

4 Conclusions

The application of StellarNet may contribute to the field of computational astrophysics by automating the classification of celestial objects as stars, galaxies, or quasars (QSOs) with an impressive accuracy of 97%. It uses spectral properties such as right ascension (alpha), declination (delta), photometric system filters (u, g, r, i, z), and redshift to perform classifications that can be used to integrate real-time automated data collection and labeling.
systems. By simplifying the identification and categorization of celestial objects, the proposed model allows researchers to focus more on data interpretation and hypothesis testing, thereby accelerating the pace of astronomical discovery. Its potential extends beyond research, demonstrating the power of artificial intelligence to transform space exploration. StellarNet is a testament to the synergy between astrophysics, astronomy, and artificial intelligence, providing a valuable resource for large sky surveys and researchers based on the income of automated labeling data without the need for manual sorting or annotations.

References


