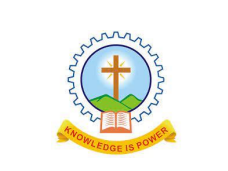
****

**Mar Athanasius College of Engineering**

**Kothamangalam**

**Initial Project Report**

**STAR-GALAXY CLASSIFICATION USING DEEP LEARNING**

Done by

**AJAY DAS M**

Reg No: MAC23MCA-2008

Under the guidance of

Prof. Nisha Markose

**Abstract**

The challenge of accurately classifying astronomical objects as stars or galaxies has been a fundamental task in astrophysics for centuries. Traditional methods relied heavily on visual inspection and morphological analysis, which were labour-intensive and limited by human subjectivity and the capacity to process large data volumes. With the advent of modern sky surveys like the Sloan Digital Sky Survey (SDSS), the volume of astronomical data has grown exponentially, rendering manual classification impractical.

The literature survey across the reviewed papers highlights three algorithms Convolution Neural Network (CNN), deep convolutional neural networks (ConvNets), ContextNet where taken into consideration.

The performance of deep learning architecture Convolution Neural Network (CNN) is used to classify stars and galaxies. Steps include rejecting data with errors, correcting for extinction, aligning images, and centring objects using nMontage and SExtractor.

The Dataset is taken from the Kaggle repository, the dataset contains 3986 data which 942 galaxy 3044 Star data.

Among the three Architecture, the Convolution Neural Network (CNN) is found to be best in terms of model building and computation. Thus, Star-Galaxy Classification Using Deep learning offers significant benefits for star-galaxy classification, including reduced human error, increased scalability, and efficient handling of vast data quantities.

**References:**

* Ganesh Ranganath Chandrasekar Iyer Krishna Chaithanya Vastare (2017). Deep Learning for Star-Galaxy Classification
* Kim EJ, Brunner RJ. Star-galaxy classification using deep convolutional neural networks. Monthly Notices of the Royal Astronomical Society. 2016 Oct 17:stw2672.
* Kennamer N, Kirkby D, Ihler A, Sanchez-Lopez FJ. ContextNet: Deep learning for star galaxy classification. In International conference on machine learning 2018 Jul 3 (pp. 2582-2590). PMLR.

**Submitted By: Faculty Guide:**

Ajay Das M Prof. Nisha Markose

MAC23MCA-2008 Associate Professor

MCA Dept, MACE

**INTRODUCTION**

The challenge of accurately classifying astronomical objects as stars or galaxies has been a fundamental task in astrophysics for centuries. Traditional methods relied heavily on visual inspection and morphological analysis, which were labour-intensive and limited by human subjectivity and the capacity to process large data volumes. With the advent of modern sky surveys like the Sloan Digital Sky Survey (SDSS), the volume of astronomical data has grown exponentially, rendering manual classification impractical.

The literature survey across the reviewed papers highlights three algorithms Convolution Neural Network (CNN), deep convolutional neural networks (ConvNets), ContextNet where taken into consideration.

The performance of deep learning architecture Convolution Neural Network (CNN) is used to classify stars and galaxies. Steps include rejecting data with errors, correcting for extinction, aligning images, and centring objects using nMontage and SExtractor.

The Dataset is taken from the Kaggle repository, the dataset contains 3986 data which 942 galaxy 3044 Star data.

Among the three Architecture, the Convolution Neural Network (CNN) is found to be best in terms of model building and computation. Thus, Star-Galaxy Classification Using Deep learning offers significant benefits for star-galaxy classification, including reduced human error, increased scalability, and efficient handling of vast data quantities.

**LITERATURE SUMMARY**

**Paper 1**

This project explores a CNN-based classifier to address these limitations. The paper "Deep Learning for Star-Galaxy Classification" (2017) demonstrates that Convolutional Neural Networks (CNNs) can effectively distinguish between stars and galaxies in astronomical images, achieving higher accuracy than traditional methods.

|  |  |
| --- | --- |
| **Title of the paper** | Ganesh Ranganath Chandrasekar Iyer Krishna Chaithanya Vastare (2017). Deep Learning for Star-Galaxy Classification |
| **Area of work** | Using deep learning, specifically Convolutional Neural Networks (CNNs), for classifying stars or galaxies. |
| **Dataset** | Dataset was taken from the Sloan Digital Sky Survey (SDSS). The dataset contains 30 million images. |
| **Methodology / Strategy** | CNN-based binary star-galaxy classifier involves collecting labelled image data from sources like the SDSS, pre-processing the data by normalizing and resizing images, and splitting it into training, validation, and test sets. A CNN is designed with convolutional and pooling layers for feature extraction, followed by fully connected layers for classification, with a sigmoid output layer for binary classification. The model is trained using binary cross-entropy loss and the Adam optimizer, then evaluated using accuracy, precision, recall, and F1-score metrics. Finally, the trained model is deployed to classify new astronomical data. |
| **Architecture** | Convolutional Neural Networks(CNN) |
| **Result/Accuracy** | CNN(Convolutional Neural Networks) – 99.19 |

**Page 2**

Kim and Brunner (2016) developed a deep CNN approach for classifying stars and galaxies in astronomical images. Their method improves accuracy by effectively learning from the features in the images, outperforming traditional classification techniques.

|  |  |
| --- | --- |
| **Title of the paper** | Kim EJ, Brunner RJ. Star-galaxy classification using deep convolutional neural networks. Monthly Notices of the Royal Astronomical Society. 2016 Oct 17:stw2672. |
| **Area of work** | Star-galaxy classification using deep convolutional neural networks. |
| **Dataset** | photometric and spectroscopic data sets with different characteristics and compositions.  data sets and the image pre-processing steps for retrieving cutout images |
| **Methodology / Strategy** | The research uses deep convolutional neural networks (ConvNets) to classify astronomical objects from SDSS and CFHTLenS survey data. The ConvNet, with several convolutional and fully connected layers, employs data augmentation and dropout to reduce over fitting. The study compares ConvNet performance to the Trees for Probabilistic Classifications (TPC) algorithm, focusing on accuracy and probabilistic calibration. |
| **Architecture** | Convolutional Neural Networks (ConvNets) |
| **Result/Accuracy** | ConvNet - 99.48 |

**Page 3**

The paper titled "ContextNet: Deep Learning for Star Galaxy Classification" presents a framework for classifying stars and galaxies in astronomical images, specifically for data from the Large Synoptic Survey Telescope (LSST)

|  |  |
| --- | --- |
| **Title of the paper** | Kennamer N, Kirkby D, Ihler A, Sanchez-Lopez FJ. ContextNet: Deep learning for star galaxy classification. In International conference on machine learning 2018 Jul 3 (pp. 2582-2590). PMLR. |
| **Area of work** | The work applies ContextNet Architecture to classify stars and galaxies in astronomical images from ground-based surveys like the LSST |
| **Dataset** | The dataset used in the work consists of simulated images from the Large Synoptic Survey Telescope (LSST) observations, generated using the GalSim image simulation package. |
| **Methodology / Strategy** | The methodology uses ContextNet, a three-step neural network framework. It includes a local network for individual object features, a global network for comparing features across objects to capture context, and a prediction network that combines these features for classification. This approach handles non-IID data and improves accuracy by leveraging neural network weight replication for variable object numbers in each exposure. |
| **Architecture** | **Local Network**: Convolutional Neural Networks (CNNs)  **Global Network**: Recurrent Neural Networks (RNNs)  **Prediction Network**: Fully Connected Neural Networks (FCNs) |
| **Result/Accuracy** | ContextNet - 95% |

**PAPER SUMMARY**

From the above three papers, we get to know that different models were used for the classification of Stars and Galaxies. The initial project report on star-galaxy classification using deep learning explores three key research papers that leverage different neural network architectures for this task. The first paper, "Deep Learning for Star-Galaxy Classification" (2017), focuses on using Convolutional Neural Networks (CNNs) to classify astronomical objects. This study utilized a large dataset from the Sloan Digital Sky Survey (SDSS) and demonstrated that CNNs could achieve high accuracy in distinguishing between stars and galaxies, with the model reaching an accuracy of 99.19%. The CNN-based approach was found to be highly effective, emphasizing the strength of deep learning in handling complex classification tasks.

The second paper, "Star-Galaxy Classification Using Deep Convolutional Neural Networks" (2016), by Kim EJ and Brunner RJ, further advanced the use of deep learning by employing deep convolutional neural networks (ConvNets). This study worked with photometric and spectroscopic datasets from the SDSS and CFHTLenS surveys and incorporated techniques like data augmentation and dropout to enhance model performance. The ConvNet model outperformed traditional classification methods, achieving a remarkable accuracy of 99.48%. This research highlighted the potential of deep learning to improve the precision and reliability of astronomical classifications.

The third paper, "ContextNet: Deep Learning for Star-Galaxy Classification" (2018), introduced a more complex architecture known as ContextNet, designed to handle data from the Large Synoptic Survey Telescope (LSST). ContextNet integrates CNNs, Recurrent Neural Networks (RNNs), and Fully Connected Neural Networks (FCNs) to capture both local and global features of astronomical images. Although this model achieved a slightly lower accuracy of 95%, it offered a sophisticated approach to addressing the challenges posed by non-independent and identically distributed (non-IID) data in astronomical surveys. Together, these studies underscore the effectiveness of deep learning, particularly CNNs, in advancing star-galaxy classification.

**PROJECT PROPOSAL**

Astronomical object classification, particularly distinguishing between stars and galaxies, has long been a fundamental challenge in astrophysics. Traditional methods, such as visual inspection and morphological analysis, are labour-intensive and limited in handling the growing volume of astronomical data generated by modern sky surveys. With the advent of deep learning, specifically Convolutional Neural Networks (CNNs), there is now a significant opportunity to automate and improve the accuracy of star-galaxy classification. This project aims to implement a CNN-based classifier, building on the work presented in the paper "Deep Learning for Star-Galaxy Classification" (2017), which demonstrated the effectiveness of CNNs in this domain.

The model will be trained with the help of a dataset taken from kagil repository. Kaggle repository. The dataset contains 3986 sample observations with 942 Galaxies and 3044 Stars photometric data.

The proposed system uses Convolution Neural Network (CNN). The models will classify photometric data under two classes Star and Galaxy. An automated system can be very helpful to offers significant benefits for star-galaxy classification, including reduced human error, increased scalability, and efficient handling of vast data quantities.

**DATASET**

The dataset is taken from the Kaggle repository. The dataset contains 3986 sample observations with 942 Galaxies and 3044 Stars photometric data

The dataset contains a collection of astronomical images captured using a 1.3-meter telescope located in Nainital, India. These images feature stars, galaxies, and other celestial objects. Researchers and data scientists can utilize this dataset for various tasks, including star-galaxy classification, object detection, and image analysis. The dataset provides a valuable resource for exploring the cosmos through machine learning and computer vision techniques.

**Dataset**: <https://www.kaggle.com/datasets/divyansh22/dummy-astronomy-data>