Classification of Morphologies of Galaxies

Alvin Dey, Pranay Raj Anand, Swagatam Chakraborti

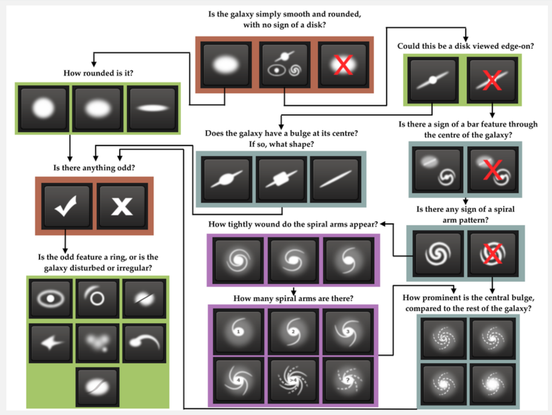
Indraprastha Institute of Information and Technology, Delhi

Objectives

Galaxy morphological classification is a system used by astronomers to divide galaxies into groups based on their visual appearance. This has been a manual process till date and extensive research is being made to automate this process.

- Our aim is to predict the morphology of a galaxy by answering 11 questions in terms of probability.

Galaxy Zoo Decision Trees



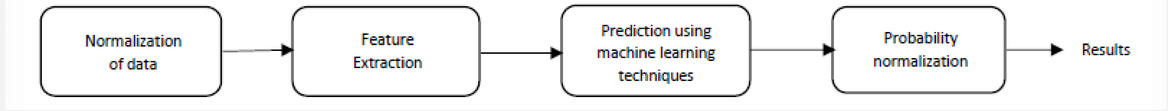
Dataset

All images has been annotated by human volunteers by answering to the 11 questions and accordingly, probabilities are provided for a total of 37 answers. Due to computational limitations, we have used 10,000 images.

The Galaxy Zoo Dataset used is taken from the Kaggle challenge.It consists of 61,578 RGB galaxy images of 424 x 424 pixels.

Proposed Methods

Regression: Linear, Lasso and Ridge Regression has been implemented for the purpose of prediction.



Convolution Neural Network: Alexnet architecture, with 5 convolution layers and 3 fully connected layers has been used. It provides probabilistic estimation per answer for each of the 11 questions.

Feature Extraction

We have used 3 features to describe the image: Histogram Of Gradient(HOG), Local Binary Pattern(LBP) and Image Moments.

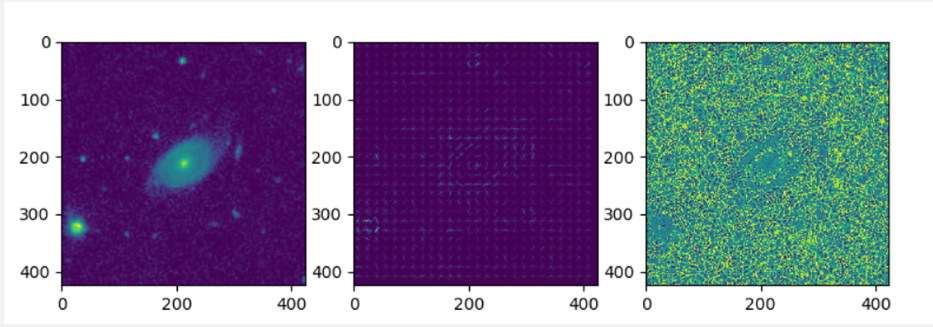


Fig : original image,HOG transformed image,LBP transformed image

Feature Anaysis

Comparative analysis of the effect of different features over each questions separately have been made.

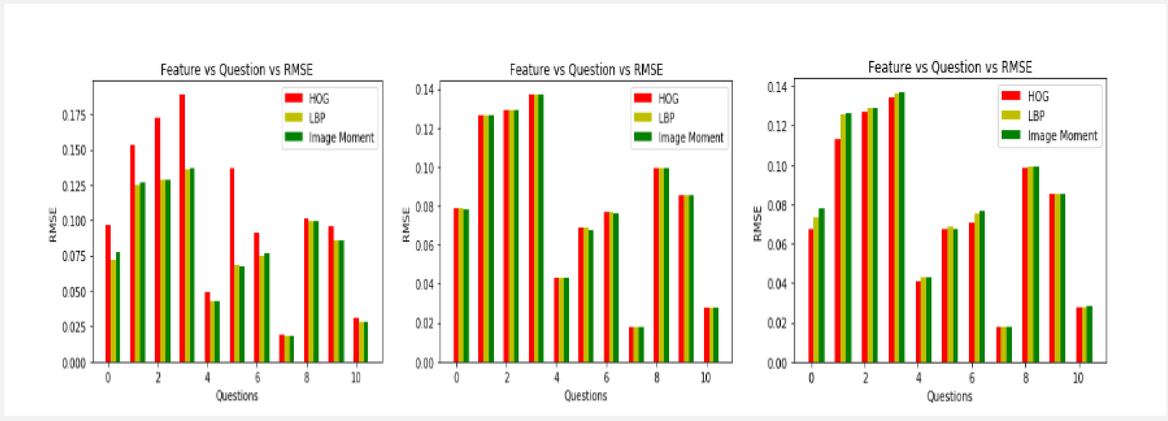


Fig : Analysis of different features over different algorithms

Results

The evaluation metric used for experiments was Root Mean Square Error(RMSE). We used the optimal feature for each question and applied regression technique and compared them with each other and also with results given by CNN.

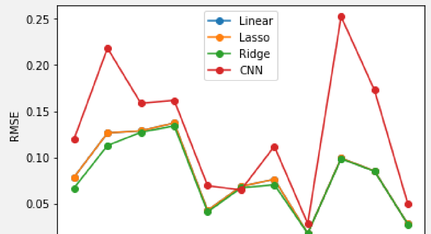


Fig : Comparison of rmse score across algorithms

Conclusion

We find that Ridge and Lasso regression perform better than both Linear regression and CNN. Image moment as a feature can be seen as the best for most of the questions. Primarily because morphology depends on shapial features and image moment describes that.

References

[1] Goderya, S.N., Lolling, S.M., 'Morphological Classification of Galaxies Using Computer Vision and Artificial Neural Networks: A Computational Scheme.' Astrophysics and Space Science. Vol. 279.

[2] Calleja, J., Fuentes, O., 'Machine Learning and Image Analysis for Morphological classification of galaxies.' Monthly Notices of the Royal Astronomical Society,Vol. 24, 2004.

[3] Calleja, J., Fuentes, O., 'Automated Classification of Galaxy Images', Lecture Notes in Computer Science, Vol. 3215, 2004.