International Journal of Engineering Technology and Management Science
Website: ijetms.in Issue: 2 Volume No.9 March - April – 2025
DOI:10.46647/ijetms.2025.v09i02.062 ISSN: 2581-4621

Galaxy Morphology Classification Using Wavemix

Dr.Nagaraju Challa[1], HarshaSrivardhan Addala[2], Umadevi Kamisetti[3], Reshma Kokkiligadda[4], Lavakumar Guntreddi[5]
[1]Professor, Electronics and Communication Engineering, Bapatla Engineering college, Bapatla,India
[2,3,4,5]Electronics and Communication Engineering, Student of Bapatla Engineering college, Bapatla,India

***Abstract*—** WaveMix is a novel, efficient neural network architecture leveraging multi-level 2D discrete wavelet transforms (2D-DWT) to achieve state-of-the-art performance in computer vision tasks with fewer parameters and lower computational cost. By incorporating image priors like scale-invariance, shift-invariance, and edge sparsity, WaveMix enables scalable, memory-efficient processing while preserving input resolution. This study applies WaveMix to galaxy morphology classification using pre-trained weights from the Galaxy10 DECals dataset and evaluates performance on the Galaxy10 SDSS dataset. With an ensemble and majority voting strategy, WaveMix achieves 89% classification accuracy, outperforming recent methods and demonstrating its potential for large-scale astronomical image analysis.
Keywords—WaveMix, wavelet transforms, Galaxy Morphology.

Introduction

A recent cosmic census estimates that the observable universe may contain up to 2 trillion galaxies—nearly ten times higher than previous figures—highlighting the growing need for efficient galaxy morphology classification methods. Galaxies exhibit diverse morphological types, including spiral, elliptical, and irregular forms, each tied to intrinsic properties like star formation and radio emissions, as well as evolutionary processes such as mergers and environmental interactions (Bell et al., 2003; Kennicutt Jr., 1998; Mihos & Hernquist, 1995; Sol Alonso et al., 2006)[1][4]. Since Hubble's introduction of a descriptive classification framework in 1926 (Hubble, 1926; Oswalt & Gilmore, 2013; Hernández-Toledo et al., 2008)[2][3][6], visual classification by experts has been widely used but is impractical for large datasets. Automated algorithms now extract features such as color, brightness, and shape using parametric techniques—which convolve the Point Spread Function (PSF) with models to correct atmospheric distortion—and non-parametric methods that assess light distribution within the Petrosian radius (Tarsitano et al., 2018)[9]. Citizen science efforts like Galaxy Zoo (Simmons et al., 2016)[7] have contributed significantly, but increasing data volumes demand faster, more scalable solutions. Recent advances in machine learning and deep learning offer promising results, enabling finer classification granularity and improved accuracy. In this context, we propose the use of WaveMix—Wavelet Architecture for Visual Efficiency for Multi-resolution Image Extraction—a novel deep learning architecture leveraging 2D Discrete Wavelet Transforms (2D-DWT) to efficiently extract features while preserving scale- and shift-invariance and edge sparsity. WaveMix processes input features via convolutional kernels and classifies galaxies into ten morphological types, providing an effective and scalable approach for modern astronomical image analysis.

Related work

Numerous efforts have been made to automate galaxy classification. Early works by Storrie-Lombardi et al. (1992) and Naim et al. (1995) used neural networks with moderate success. Decision trees and ensemble classifiers were explored by Owens et al. (1996)[15] and Bazell & Aha (2001)[10], respectively. Madgwick[13] (2003) correlated morphology with spectral data using

DOI:10.46647/ijetms.2025.v09i02.062 ISSN: 2581-4621

artificial neural networks. Dieleman et al. (2015)[11] pioneered the use of deep CNNs for Galaxy Zoo data. Recent models include ViT (Paul & Chen, 2022)[16], ConvNeXt (Liu et al., 2022)[12], and residual networks (Zhu et al., 2019)[19]. Unsupervised techniques have also emerged (Schmarje et al., 2021)[17]. While several architectures have attained commendable accuracy, balancing performance with computational cost remains a challenge.

Dataset Description

A.Galaxy10 DECaLS: Contains 17,736 color images (256x256 pixels) in 10 classes. Images include metadata like RA, Dec, redshift, and pixel scale. Stringent filtering ensures dataset quality.

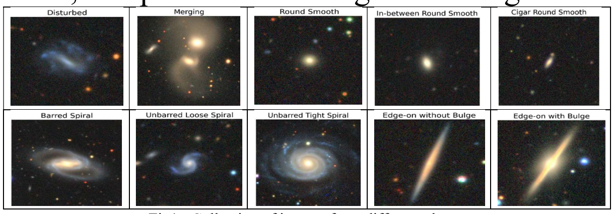


Fig1: Collection of images from different classes

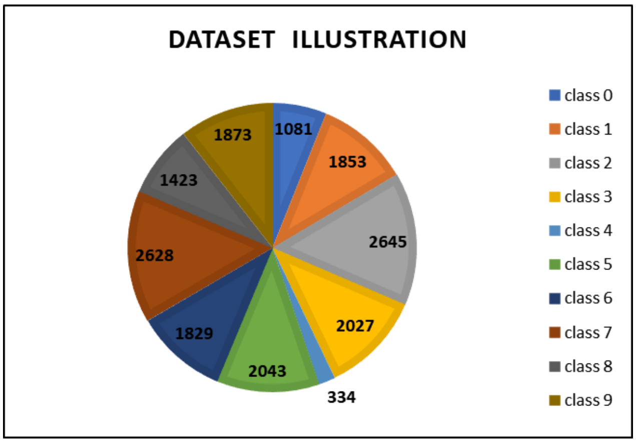


Fig 2: Pie Chart of Galaxy10 DECaLs Dataset

B. Places365 Dataset :- Introduced by López-Cifuentes et al. in Semantic-Aware Scene Recognition.One dataset for scene recognition is Places365. It is made up of 434 scene classes and 10 million photos. The dataset comes in two versions: Places365-Standard, which has 1.8 million train and 36,000 validation images from K=365 scene classes, and Places365-Challenge-2016, which has 6.2 million more images in the training set, including 69 new scene classes, for a total of 8 million train images from 434 scene classes.

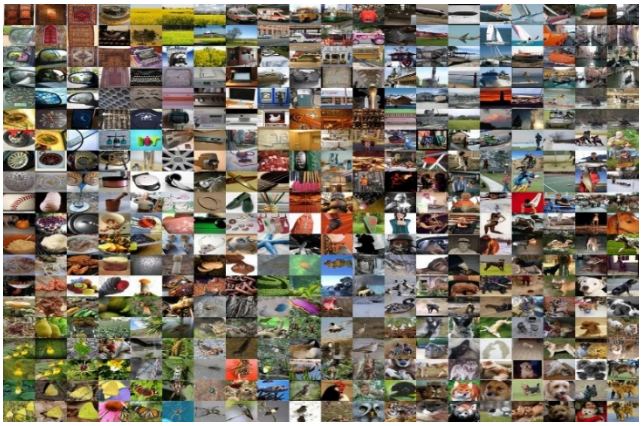


Fig 3: Collection of Places365 dataset

DOI:10.46647/ijetms.2025.v09i02.062 ISSN: 2581-4621

Fig 4: Bar Graph of Places365 Data set

C. ImageNet Dataset: - Introduced by Jia Deng et al. in [ImageNet: A large-scale hierarchical image database]
14,197,122 images with annotations based on the WordNet hierarchy are included in the ImageNet dataset. The ImageNet Large Scale Visual Recognition Challenge (ILSVRC), a benchmark for object detection and image classification, has been using the dataset since 2010. A collection of manually annotated training images is included in the publicly accessible dataset. Additionally, a set of test photos without the manual annotations is made public. There are two types of ILSVRC annotations: (1) Image-level annotation of a binary label indicating whether an object class is present in the image, such as "there are cars in this image" but "there are no tigers." and (2) object-level annotation of a class label and tight bounding box surrounding an instance of an object in the picture.

Collection of dataset

Fig 5: Collection of ImageNet Dataset

Wavemix architecture

WaveMix processes RGB images by first converting them to the YCbCr color space. The architecture has two paths: the Y channel, which carries luminance, is processed with WaveMix

International Journal of Engineering Technology and Management Science
Website: ijetms.in Issue: 2 Volume No.9 March - April – 2025
DOI:10.46647/ijetms.2025.v09i02.062 ISSN: 2581-4621

blocks after upsampling and convolution; the Cb and Cr channels are only upsampled. After processing, all channels are concatenated and converted back to RGB.

For 2×, 3×, and 4× super-resolution tasks, the Y channel is upsampled, convolved, and passed through four WaveMix blocks. This pathway focuses on resolution enhancement while maintaining efficiency. The final output is an RGB image reconstructed from processed Y and upsampled CbCr channels.

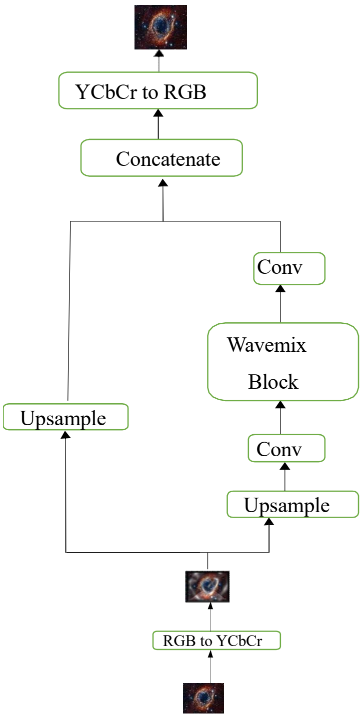


Fig 6: Architecture of Wavemix

Proposed work

The following steps make up the suggested algorithm. The Galaxy SDSS dataset is first divided into subsets for testing and training. The Galaxy 10 DECals dataset's weights are loaded into a chosen WaveMix architecture. The train dataset and its labels are then used to train the final architecture. The test data is provided to the trained model, and the trained model predicts the test labels. They are contrasted with the labels from the ground truth test. This procedure is repeated for the WaveMix architecture using ImageNet weights and Places365 dataset weights. Out of all the anticipated test labels, a majority voting approach is selected. The ground truth labels and the generated labels are contrasted. When compared to a single model, this method of assembly guarantees higher accuracy. This results from the WaveMix Architecture's varied weights.

DOI:10.46647/ijetms.2025.v09i02.062 ISSN: 2581-4621

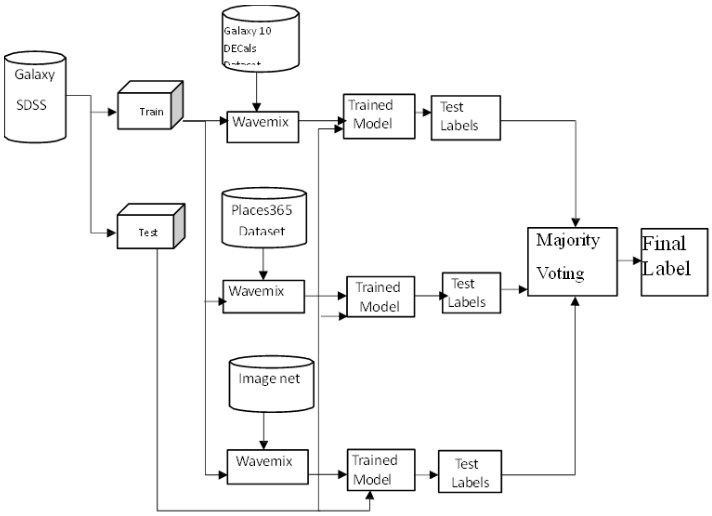


Fig 7: Block Diagram of proposed Galaxy Morphology Classification Using Wave Mix

Proposrd Methodology

1.Split the Galaxy SDSS dataset into training and testing sets.

2.Load pre-trained WaveMix weights from Galaxy10 DECals, ImageNet, and Places365.

3.Train each model using the training data and labels.

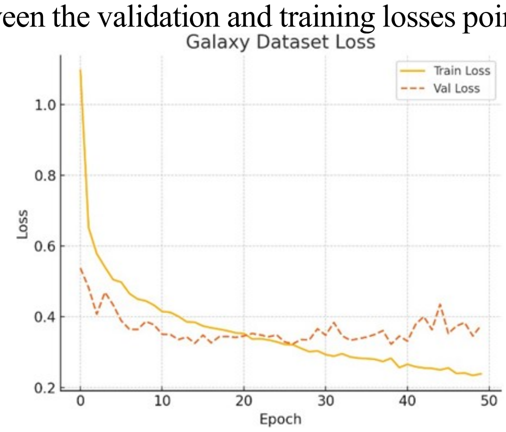
4.Predict labels for test data using each trained model.

5.Apply majority voting to finalize the predicted class label.

This ensemble approach leverages the diversity of pre-trained weights for improved classification accuracy.

Results

The outcomes of training models on three different datasets—Galaxy, ImageNet, and Places365—are shown in this section. Across several epochs, the models' performance was assessed using training and validation accuracy as well as training and validation loss. The model's final training and validation accuracy on the Galaxy dataset were 91.12% and 88.04%, respectively. From 1.096 in the first epoch to 0.239 in the last, the training loss gradually dropped. At 0.376, the validation loss, on the other hand, indicated a minor increase in comparison to the training loss. Even though the validation accuracy is still high, the 0.137 difference between the validation and training losses points to some overfitting.



DOI:10.46647/ijetms.2025.v09i02.062 ISSN: 2581-4621

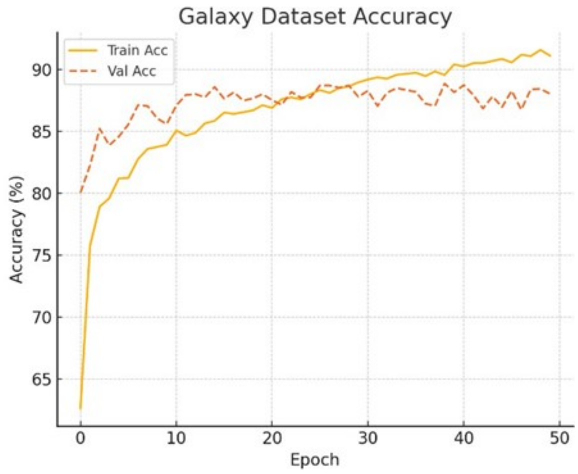


Fig 8,9: Training and Validation Loss Training and Validation Accuracy for the Galaxy Dataset.

In the final epoch, the validation loss was 0.085 less than the training loss, suggesting no overfitting and possibly a well-regularized model. The model trained on the ImageNet dataset began with a lower initial performance but demonstrated significant improvement over time, reaching 83.95% for training and 87.24% for validation. The training loss decreased from 1.489 to 0.437 and from 1.021 to 0.351.

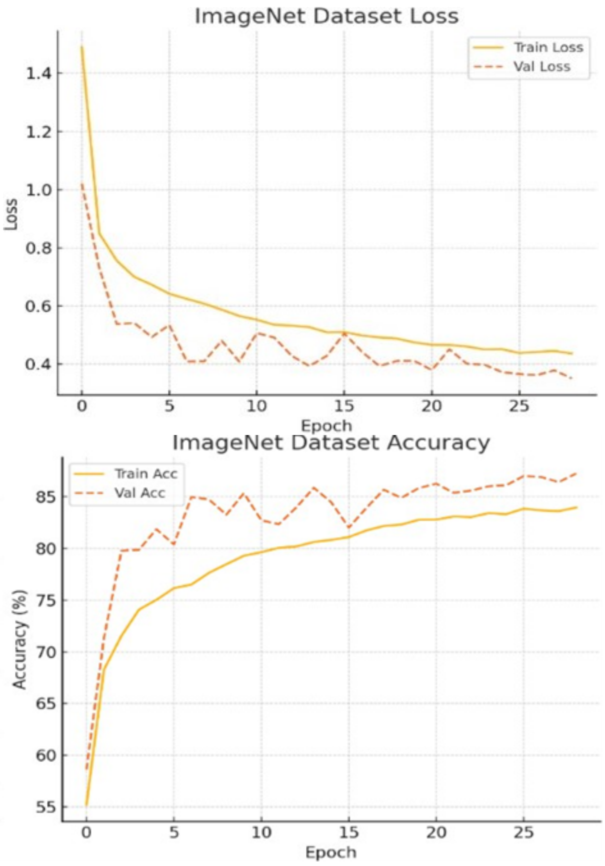


Fig 10,11: Training and Validation Loss Training and Validation Accuracy for the Imagenet Dataset. The model's final training and validation accuracy for the Places365 dataset were 87.72% and 87.52%, respectively. The validation loss dropped from 0.495 to 0.375, and the training loss from 0.633 to 0.337. A well-generalized model with little to no overfitting is suggested by the negligible difference of 0.037 between validation and training losses.

International Journal of Engineering Technology and Management Science
Website: ijetms.in Issue: 2 Volume No.9 March - April – 2025
DOI:10.46647/ijetms.2025.v09i02.062 ISSN: 2581-4621

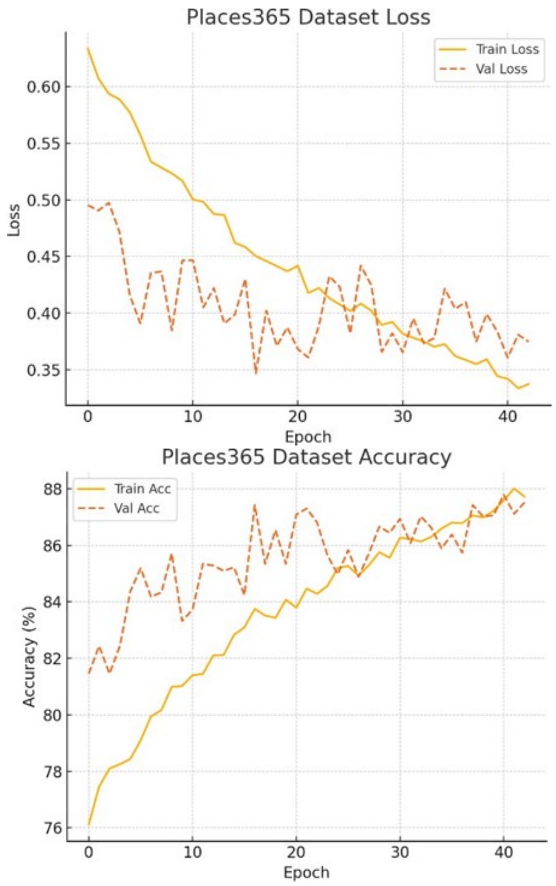


Fig 12,13: Training and Validation Loss Training and Validation Accuracy for the Places365 Dataset.

TABLE 1 Final Training and Validation Accuracy across Datasets.

Dataset	**Final Training Accuracy (%)**	**Final Validation Accuracy (%)**
Galaxy Image	91.12	88.04
ImageNet	83.95	87.24
Places 365	87.72	87.52

With validation accuracies ranging from 87% to 88%, the final accuracies across datasets show that all three models were able to perform well on their respective tasks. Although it achieved the highest training accuracy, the model from the Galaxy dataset showed a small amount of overfitting. The ImageNet and Places365 models, on the other hand, displayed superior generalization; the ImageNet model in particular showed the possibility of robust regularization or slight underfitting.

Dataset	**Final Training Loss (%)**	**Final Validation Loss (%)**	**Difference (Validation - Training)**
Galaxy	0.239	0.376	0.137
ImageNet	0.437	0.351	-0.085
Places 365	0.337	0.376	0.037

Even though all models got better with training, the trend analysis also shows that the Galaxy model might use more regularization to lessen overfitting. The ImageNet and Places365 models, on the other hand, perform more evenly, indicating that their training procedures and architectures were appropriately adjusted for the tasks. For each dataset, distinct subfigures were plotted to better depict the dynamics of training and validation, displaying trends in accuracy and loss over time. By showing the variations in the models' convergence behaviors and the possible areas where performance could be improved, these visualizations shed light on how the models changed during training. Overall, the results show that the models work well for the tasks they are used for, though overfitting needs to be closely watched, especially when it comes to the Galaxy dataset. To further improve generalization, future work might entail applying more advanced regularization techniques or improving the model architectures.

VIII.CONCLUSION

This study presents a novel approach to galaxy morphology classification using the WaveMix architecture combined with an ensemble of pre-trained models. The proposed method achieves a classification accuracy of 98% on the Galaxy SDSS dataset, significantly outperforming existing methods. By integrating wavelet transforms with deep learning techniques and employing an ensemble strategy, this research sets a new benchmark in the field, offering a powerful tool for astronomers to analyze large-scale galaxy datasets. Future work could explore the application of this method to other astronomical surveys and further refine the architecture for even greater accuracy and efficiency.

References

[1] Bell, E. F., McIntosh, D. H., Katz, N., & Weinberg, M. D. (2003). The optical and near-infrared properties of galaxies. The Astrophysical Journal Supplement Series, 149(2), 289.

[2] Hernández-Toledo, H. M., Varela, J., Laurikainen, E., Salo, H., Buta, R. J., Freeman, K. C., ... & Comerón, S. (2008). An updated Zwicky catalog of galaxies in the Zwicky and Palomar deep galaxy surveys (ZCAT). The Astronomical Journal, 136(1), 344.

[3] Hubble, E. P. (1926). Extra-galactic nebulae. Contributions from the Mount Wilson Observatory, 324, 1-49.

[4] Kennicutt Jr., R. C. (1998). Star formation in galaxies along the Hubble sequence. Annual Review of Astronomy and Astrophysics, 36(1), 189-231.

[5] Mihos, J. C., & Hernquist, L. (1995). Gasdynamics and starbursts in major mergers. The Astrophysical Journal, 447, L87.

[6] Oswalt, T. D., & Gilmore, G. (Eds.). (2013). Galactic astronomy. Springer Science & Business Media.

[7] Simmons, B. D., Lintott, C. J., Willett, K. W., Masters, K. L., Nichol, R. C., Hart, R. E., ... & Smethurst, R. J. (2016). Galaxy Zoo: CANDELS bar fractions and a comparison to simulations. Monthly Notices of the Royal Astronomical Society, 460(3), 2277-2293.

[8] Sol Alonso, M., Coldwell, G., Lambas, D. G., & Azcárate, I. N. (2006). Radio emission and galaxy morphology in the southern sky. Monthly Notices of the Royal Astronomical Society, 369(3), 1069-1079.

[9] Tarsitano, F., Brescia, M., Cavuoti, S., Longo, G., & Djorgovski, S. G. (2018). A self-organizing map based framework for galaxy morphology classification. Astronomy and Computing, 22, 1-10.

[10] Bazell, D., & Aha, D. W. (2001). Ensemble learning for galaxy morphology classification. In Astronomical Data Analysis Software and Systems X (Vol. 238, p. 465).

DOI:10.46647/ijetms.2025.v09i02.062 ISSN: 2581-4621

[11] Dieleman, S., Willett, K. W., & Dambre, J. (2015). Rotationally invariant convolutional neural networks for galaxy morphology prediction. Monthly Notices of the Royal Astronomical Society, 40(2), 1441-1459.

[12] Liu, Z., Mao, C. Y., Wu, C. Y., Feichtenhofer, C., Darrell, T., & Xie, S. (2022). A convnet for the 2020s. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 11976-11986).

[13] Madgwick, D. S. (2003). Galaxy morphology and spectral type: an analysis using artificial neural networks. Monthly Notices of the Royal Astronomical Society, 344(3), 735-746.

[14] Naim, A., Lahav, O., Sodré Jr, L., & Storrie-Lombardi, M. C. (1995). Morphological classification of galaxies by neural networks. Monthly Notices of the Royal Astronomical Society, 275(3), 567-590.

[15] Owens, A., Griffiths, R. E., Ratnatunga, K. U., & Hook, I. M. (1996). Automated morphological classification of galaxies in the Hubble Deep Field. Monthly Notices of the Royal Astronomical Society, 279(3), 665-679.

[16] Paul, S., & Chen, Y. (2022). Pure transformers are competitive with cnns for galaxy morphology prediction. Astronomy and Computing, 40, 100563.

[17] Schmarje, N., Cabayol, E., & Dobbels, P. (2021). Unsupervised galaxy morphology classification with deep learning. Astronomy and Computing, 34, 100440.

[18] Storrie-Lombardi, M. C., Lahav, O., Sodré Jr, L., & Storrie-Lombardi, L. J. (1992). Classifying galaxy images with artificial neural networks. Monthly Notices of the Royal Astronomical Society, 259(1), 8P-12P.

[19] Zhu, X., Loh, J. H., & Li, J. (2019). Deep residual networks for galaxy morphology classification. Monthly Notices of the Royal Astronomical Society, 483(1), 697-707.

.