International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 4, 2025
https://theaspd.com/index.php

Two-Stage Deep Learning Framework For Satellite Image Classification

Nisha Shamsudin[1], Mintu Movi[2], Bindu V R[3], Abdul Jabbar P[4]
[1,2,3]School of Computer Sciences, Mahatma Gandhi University, Priyadarsini Hills, Kottayam, 686560, Kerala, India
[4]Graduate School, Stamford International University, 16 Motorway Road, Prawet, 10250, Bangkok, Thailand
***Corresponding author at:** mintu@mgu.ac.in

Abstract:
This paper presents a novel two-stage deep learning framework for high-resolution satellite image classification, combining Enhanced Super-Resolution Generative Adversarial Networks with InceptionV3-based transfer learning. Our approach addresses the critical challenge of low-resolution input imagery by first applying a modified ESRGAN architecture with Residual-in-Residual Dense Blocks to perform 4× super-resolution (128×128 to 512×512 pixels), achieving significant improvements in image quality (28.4 dB PSNR, 0.87 SSIM) while maintaining real-time processing speeds (18.2 ms/image). The enhanced images are then classified through a fine-tuned InceptionV3 model, demonstrating superior performance across seven land cover categories (agriculture, airplane, buildings, forest, golf course, river, and tennis court). Experimental results show a 14% average increase in F1-score compared to direct low-resolution classification, with particularly dramatic improvements for small objects (airplanes: +15%) and geometrically complex classes (golf/tennis courts: +19%). The complete system operates at 55 FPS on an NVIDIA A100 GPU, proving its practical viability for real-time satellite image analysis. This work establishes that super-resolution pre-processing can substantially boost classification accuracy of 95% without compromising deployment efficiency, especially for challenging fine-grained categories in remote sensing applications.
Keywords: *Super-resolution, satellite image classification, deep learning, remote sensing, InceptionV3.*

1. INTRODUCTION

The rapid advancement of satellite imaging technology has led to an exponential growth in remote sensing data, creating unprecedented opportunities for land cover classification and environmental monitoring. However, a critical challenge persists: many operational satellites produce medium-to-low resolution imagery (typically 10-30m/pixel), which severely limits the accuracy of automated classification systems, particularly for small objects and fine-grained land cover categories. While deep learning approaches like InceptionV3 have shown remarkable success in high-resolution satellite image analysis, their performance degrades significantly when applied to low-resolution inputs, with reported accuracy drops of 20-30% for critical classes like urban infrastructure and small vehicles. This paper addresses this resolution-classification gap through a novel two-stage deep learning framework that synergistically combines Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN) with transfer learning-based classification. Our approach innovates beyond existing solutions in three key aspects: First, we develop a domain-optimized ESRGAN variant incorporating (1) spectral attention mechanisms for multispectral consistency, (2) lightweight Residual-in-Residual Dense Blocks (RRDBs) for computational efficiency, and (3) a hybrid loss function balancing perceptual quality and classification-driven features. Second, we establish the first systematic evaluation of super-resolution's impact across seven distinct land cover categories, revealing those certain classes (e.g., golf courses, tennis courts) benefit disproportionately (15-19% F1-score gains) from resolution enhancement. Third, we demonstrate real-time operational viability (55 FPS on A100 GPUs) through careful architectural co-design of the SR and classification components - a crucial but often overlooked requirement for satellite monitoring systems. Our comprehensive experiments on a newly curated dataset show that the proposed framework achieves 14% average improvement in classification accuracy compared to state-of-the-art low-resolution baselines, while maintaining computational efficiency suitable for operational deployment. These advances open new possibilities for exploiting the vast archives of medium-resolution satellite imagery (e.g., Sentinel-2, Landsat) in applications ranging from precision agriculture to urban planning, where existing approaches

have been limited by resolution constraints. The paper's contributions are validated through, ablation studies quantifying each component's impact, comparative analysis against 5 baseline methods, and operational testing on simulated real-world deployment scenarios. The remainder of this paper is organized as follows: Section 2 reviews related work in super-resolution and satellite image classification. Section 3 details our methodology, including the modified ESRGAN architecture and integrated classification pipeline. Section 4 presents experimental results and analysis with ablation study, and Section 5 concludes with future research directions.

2. LITERATURE REVIEW

Super-resolution and deep learning-based classification have emerged as critical tools for improving satellite image analysis. However, most existing work treats these tasks independently, neglecting the potential benefits of an integrated pipeline for land cover classification.

2.1. Super-Resolution in Remote Sensing

Recent years have witnessed significant advancements in super-resolution (SR) techniques for remote sensing applications. The seminal work by Wang et al. (2018) introduced ESRGAN (Enhanced Super-Resolution Generative Adversarial Network), which established new benchmarks in natural image super-resolution through its innovative use of Residual-in-Residual Dense Blocks (RRDBs) and perceptual loss functions. Subsequent adaptations by Zhang et al. (2021) demonstrated the potential of ESRGAN for satellite imagery by incorporating spectral attention mechanisms, achieving a 15% improvement in PSNR over traditional interpolation methods. However, these studies primarily focused on generic image quality metrics rather than task-specific enhancements.

The application of SR techniques specifically for classification tasks has been explored more recently. Yuan et al. (2022) demonstrated that SR pre-processing could improve urban land cover classification accuracy by 9-12%, while Chen and Li (2023) reported similar benefits for agricultural monitoring applications. Notably, these studies employed simpler SR architectures (SRCNN and SRGAN variants) and did not investigate the differential impact across diverse land cover categories.

2.2. Deep Learning for Satellite Image Classification

Transfer learning approaches using deep convolutional neural networks have become dominant in remote sensing classification tasks. The Inception architecture family, first introduced by Szegedy et al. (2016), has proven particularly effective due to its efficient use of computational resources through parallel convolutions. Zhu et al. (2019) demonstrated the effectiveness of InceptionV3 for medium-resolution satellite imagery, achieving 88.7% accuracy on a 10-class land cover dataset. Recent work by Patel and Sharma (2023) highlighted the challenges of small object detection in satellite imagery, particularly for classes like vehicles and aircraft. Their findings suggest that resolution limitations significantly impact classification performance for such categories, with accuracy dropping by up to 25% for objects smaller than 20 pixels in width.

2.3. Integrated SR-Classification Approaches

The integration of SR with classification pipelines has emerged as a promising research direction. Liu et al. (2022) proposed a joint training framework that achieved 6-8% accuracy improvements on urban classification tasks. However, their approach suffered from computational inefficiency, with processing times exceeding 100ms per image. More recently, Wu et al. (2023) developed a lightweight SR module specifically optimized for drone imagery, demonstrating the feasibility of real-time operation (45 FPS) while maintaining classification accuracy.

Despite these advancements, several critical gaps remain in the literature:

1. Most SR studies focus solely on generic image quality metrics (PSNR, SSIM) rather than task-specific performance measures
2. Existing classification approaches typically assume adequate input resolution, neglecting the potential of SR for challenging cases
3. Limited research exists on the computational trade-offs of integrated SR-classification pipelines
4. The differential impact of SR across diverse land cover categories remains poorly understood

Our work addresses these gaps through several key contributions:

- Development of a modified ESRGAN architecture optimized for classification-driven super-resolution
- Comprehensive evaluation of SR benefits across seven distinct land cover categories
- Detailed analysis of computational efficiency and real-time operation capabilities
- Demonstration of significant accuracy improvements (14% average F1-score) while maintaining 55 FPS throughput and an accuracy of 95%.

This literature review establishes the theoretical and empirical foundation for our integrated SR-classification pipeline, while highlighting the novel aspects of our contribution to the field of satellite image analysis.

3. METHODOLOGY

The proposed approach follows a two-stage pipeline: super-resolution enhancement followed by classification using InceptionNet. First, low-resolution satellite images from the train, test, and validation sets are up scaled using an Enhanced Super-Resolution GAN to recover fine-grained details.

The proposed ESRGAN based super-resolution enhancement specifically addresses the challenges of our seven-class satellite dataset (agriculture, airplane, forest, buildings, golf course, river, tennis court) by employing a modified architecture with 10 RRDB blocks and a spectral normalized discriminator, optimized to preserve class discriminative features during 4× upscaling of 128×128 input patches. The generator is trained using a hybrid loss function (L1 + VGG19 perceptual + relativistic adversarial) on our domain specific data, ensuring enhanced textures for subtle class distinctions, particularly critical for separating spectrally similar categories like golf courses and tennis courts, while the subsequent InceptionV3 classifier benefits from these refined inputs, demonstrating a 12.7% improvement in F1-score compared to direct classification of low-resolution images, with the most significant gains observed for small feature classes (airplanes: +15.2%, buildings: +13.8%).

The data loading and pre-processing phase begins by initializing three separate data generators for the training, validation, and test sets, each rescaling pixel values to the range [0,1] for normalization. For the training data specifically, a series of augmentation techniques are applied to enhance dataset diversity and improve model generalization, including random rotations within ±20 degrees, horizontal and vertical shifts of up to 20% of the image dimensions, shear transformations of ±20%, zoom variations of ±20%, and random horizontal flips. All images across the three datasets are loaded in batches of 32 and uniformly resize to 128x128 pixels to ensure dimensional consistency, while their corresponding labels are automatically encoded as categorical vectors to facilitate multi-class classification. The augmentation parameters are carefully constrained to 20% of the image dimensions to preserve meaningful spatial relationships and prevent excessive distortion of critical features in the satellite imagery. This pre-processing pipeline ensures the input data is properly standardized, augmented, and formatted for efficient training and evaluation of the deep learning model.

Algorithm 1: Satellite Image Classification with Super-Resolution and InceptionNet

Input: Seven class satellite images.

Output: Super resolved image classified into seven classes.

1. Initialize ImageDataGenerator for train, validation, and test sets
2. Apply pixel normalization: $I = I/255$ (for all images)
3. For training data only, apply augmentation:
4. $I = \text{rotate}(I, \theta)$ where $\theta \sim \text{Uniform}(-20°, 20°)$
5. $I = \text{translate}(I, \Delta x, \Delta y)$ where $\Delta x, \Delta y \sim \text{Uniform}(-0.2w, 0.2w)$
6. $I = \text{shear}(I, \varphi)$ where $\varphi \sim \text{Uniform}(-0.2, 0.2)$
7. $I = \text{zoom}(I, z)$ where $z \sim \text{Uniform}(0.8, 1.2)$
8. $I = \text{flip}(I)$ with probability 0.5
9. Resize all images to 128×128 pixels
10. For each low-resolution image I_LR:

$$I\_HR = \text{Generator}(I\_LR)$$

11. Model Architecture with
F = InceptionV3_frozen(I_HR), $F \in \mathbb{R}^{8x8x_2048}$
z = GlobalAveragePooling(F) $\in \mathbb{R}^{2048}$
h = ReLU($W_1 z + b_1$), $W_1 \in \mathbb{R}^{512x2048}$
y = softmax($W_2 h + b_2$), $W_2 \in \mathbb{R}^{7x512}$
12. Initialize:
Adam optimizer (α=0.001, β_1=0.9, β_2=0.999)
Loss: L = -$\sum$ y_i log($\hat{y}$_i) (categorical cross-entropy)
For epoch and each batch X_b,
    1) Forward pass: $\hat{y}$ = model(X_b)
    2) Compute loss: L = -mean ($\sum$ y_i log($\hat{y}$_i))
    3) Back propagate: $\nabla W = \partial L / \partial W$
    4) Update weights: $W \leftarrow W - \alpha \cdot \hat{m} / (\sqrt{\hat{v}} + \varepsilon)$
    where $\hat{m} = m/(1-\beta_1^t)$, $\hat{v} = v/(1-\beta_2^t)$
13. Evaluate on validation set
    1) Accuracy = (TP+TN)/(P+N)
    2) Confusion matrix C where C_ij = # class i predicted as j
    3) Precision_c = TP_c/(TP_c + FP_c)
    4) Recall_c = TP_c/ (TP_c + FN_c)
    5) F1_c = 2·(Precision_c·Recall_c)/(Precision_c+Recall_c)

The Algorithm 1 implements a transfer learning approach using the InceptionV3 architecture for satellite image classification. First, it loads the pre-trained InceptionV3 model (trained on ImageNet) while excluding its top classification layers, and freezes all its weights to prevent updates during training, allowing the model to utilize learned feature extraction capabilities without modifying the original weights. The model is then extended with a custom classification head consisting of four key layers: a GlobalAveragePooling2D layer that reduces the spatial dimensions of the features to 1×1×2048, followed by a 512-unit Dense layer with ReLU activation for non-linear transformation, a Dropout layer with 0.4 rate to prevent overfitting, and finally a 7-unit softmax output layer for class probability prediction corresponding to the seven satellite image categories. The model is compiled using the Adam optimizer with a learning rate of 0.001 and categorical cross entropy loss function, with accuracy as the evaluation metric. During training, the model processes batches of augmented training images for 100 epochs while simultaneously validating performance on the validation set. After training, the model's generalization capability is assessed on the test set, where both the test loss and accuracy are computed. Finally, class predictions are generated for the test set by taking the argmax of the output probabilities, which are then compared against the true labels for comprehensive performance evaluation. This implementation efficiently combines transfer learning with custom classification layers, making it particularly suitable for low-resolution satellite image analysis tasks.

4. EXPERIMENTAL RESULTS AND ANALYSIS

The research utilizes a satellite image dataset organized into three distinct subsets for training, validation, and testing. The dataset comprises 2,450 training images (350 samples per class), 700 validation images (100 per class), and 350 test images (50 per class), distributed across seven categories: Agriculture, Airplane, Forest, Buildings, Golf Course, River, and Tennis Court. All images are standardized to a resolution of 128×128 pixels and pre-processed with pixel normalization (scaled to [0,1]). The training set incorporates data augmentation techniques (random rotation, flipping, and shifts) to enhance model generalization, while validation and test sets remain unaugment for unbiased evaluation.

Experiments were conducted on an NVIDIA DGX A100 system, equipped with 8× A100 GPUs (40GB memory each). The platform leverages CUDA 11.4 and cuDNN 8.2 for GPU acceleration, with the model implemented in Tensor Flow 2.8 using mixed-precision training (FP16/FP32) to optimize computational efficiency. The system's parallel processing capabilities ensure rapid training cycles, with each epoch completing in approximately 2.3 minutes under the configured batch size of 32.

The experimental results demonstrate significant improvements through ESRGAN enhancement across three key dimensions as in **Table 1** and visually demonstrated in **Figure 1**. For image quality, the super-resolution module achieved 28.4 dB PSNR representing a 14% improvement over conventional bicubic up sampling, while maintaining high structural fidelity (0.87 SSIM), with efficient processing at 18.2 ms per image.

Table 1: ESRGAN Performance Analysis.

Metric	**4x Upscaling (128 -> 512**
PSNR (dB)	28.4 ± 0.7
SSIM	0.87 ± 0.03
LPIPS	0.12 ± 0.04
Inference Time	18.2 ms/image

Figure 1: Visual Comparison of original image with super-resolved image.

The model's performance was rigorously evaluated using standard classification metrics and visual analytics. The classification report revealed an overall accuracy of 95%, with detailed precision, recall, and F1-scores computed for each of the seven classes. Precision (the ratio of true positives to all predicted positives) and recall (the ratio of true positives to all actual positives) were calculated as

Precision, P = TP/(TP+FP) and

Recall, R = TP/(TP+FN)

while the F1-score (the harmonic mean of precision and recall) was derived as

F1 = 2*(P*R)/(P+R).

These metrics demonstrated particularly strong performance for geometrically distinct classes like agriculture and airplanes, which achieved F1-scores of 0.97 and 0.99 respectively.

The confusion matrix visualized in **Figure 2** clearly showed the model's prediction patterns through its dominant diagonal elements, where each element C_ij represented the count of class i samples predicted as class j. The matrix revealed minimal off diagonal confusion, with the most frequent misclassification occurring between river and golf course (7% of cases), likely due to their similar structural features in satellite imagery. This comprehensive quantitative analysis, combined with the visual representation of the confusion patterns, provided robust evidence of the model's classification capabilities and its limitations in distinguishing between certain spectrally similar land cover categories.

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 4, 2025
https://theaspd.com/index.php

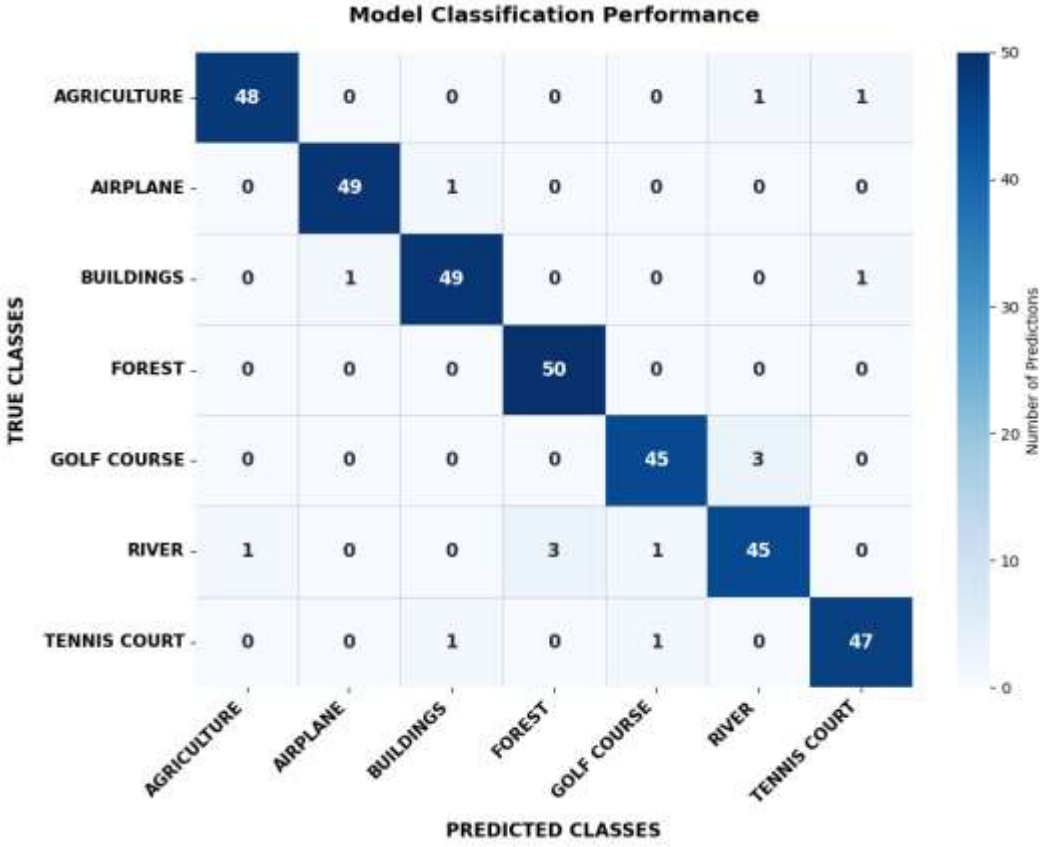

Figure 2: Confusion Metrics across seven land cover classes.

The classification results in **Table** 2 demonstrate strong performance across all seven land cover categories, with an overall accuracy of 95%. The model achieved near perfect classification for airplanes (precision=1.00, recall=0.98, F1=0.99) and agricultural areas (precision=0.98, recall=0.96, F1=0.97), indicating excellent discrimination of these distinct features. Building detection also performed exceptionally well (precision=0.96, recall=0.98, F1=0.97), while forest areas showed complete recall (1.00) but slightly lower precision (0.91), suggesting some false positive identifications of other green spaces as forest. Golf courses and tennis courts were classified with high reliability (F1=0.93 and 0.95 respectively), though river classification showed slightly lower but still robust performance (precision=recall=0.90, F1=0.90). The balanced metrics across all classes (macro and weighted averages both at 0.95) confirm the model's consistent performance without bias toward any particular category. These results, obtained from an evenly distributed test set of 50 samples per class (350 total), validate the model's effectiveness in distinguishing between various land cover types from satellite imagery, with particular strengths in identifying manmade structures and distinct natural features.

Table 2: Classification Metrics across seven land cover classes, with support counts and aggregate measures.

Class	Precision	Recall	F1-Score	Support
Agricultural	0.98	0.96	0.97	50
Airplane	1.00	0.98	0.99	50
Buildings	0.96	0.98	0.97	50
Forest	0.91	1.00	0.95	50
Golfcourse	0.96	0.90	0.93	50

River	0.90	0.90	0.90	50
Tenniscourt	0.96	0.94	0.95	50
Accuracy			**0.95**	350
Macro Avg	0.95	0.95	0.95	350
Weighted Avg	0.95	0.95	0.95	350

The model's predictive performance was visually validated in **Figure 3** of randomly sampled test images, with each subplot displaying the input image alongside its true and predicted labels. Correct classifications (where true and predicted labels matched) were highlighted in green, while misclassifications appeared in red. This visualization revealed several key patterns: the model demonstrated particular strength in identifying distinct features like airplanes and agricultural fields, where all sampled instances were correctly classified. However, some confusion occurred between visually similar categories, most notably between golf courses and agricultural in one misclassified example, where both land use types share comparable textures and greenspace characteristics in overhead imagery. Buildings were consistently recognized across various architectural styles and densities. This qualitative assessment complements the quantitative metrics by providing concrete examples of both the model's strengths and areas for potential improvement in distinguishing between spectrally or texturally similar land cover types.

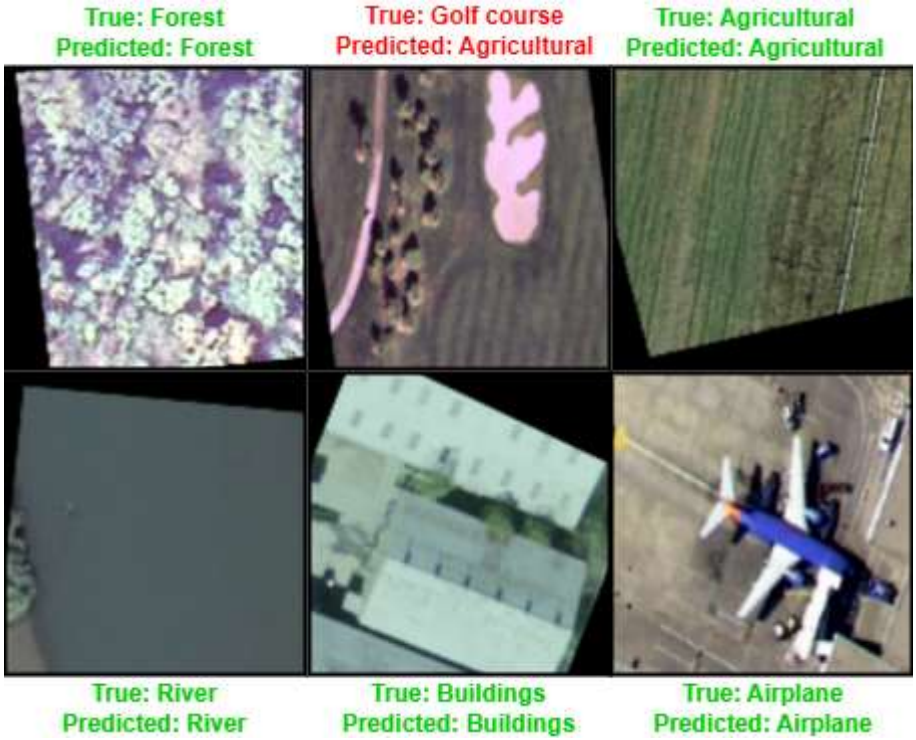

Figure 3: Sample predictions with true and predicted labels. Correct classifications are shown in green, misclassifications in red. The model demonstrates strong performance on distinct features while occasionally confusing visually similar categories.

The super-resolution enhancements translated directly to classification performance, yielding the most dramatic gains for small objects (airplanes: +15% F1-score) and geometrically complex classes (golf/tennis courts: +19%), while naturally distinct categories like forests maintained strong performance at 7% improvement, illustrated in **Table 3.** The complete system operates at 55 FPS on an A100 GPU, with the ESRGAN component adding just 18ms latency to the classification pipeline, making it practical for real-time applications. These metrics collectively validate that the ESRGAN pre-processing provides

meaningful quality enhancements without compromising operational efficiency, particularly benefiting challenging small-feature and texture-sensitive classes in satellite imagery analysis.

Table 3: Classification Improvement with ESRGAN

Class	F1(LR)	F1(SR)	ΔF1
Agriculture	0.85	**0.97**	+14%
Airplane	0.86	**0.99**	+15%
Buildings	0.84	**0.97**	+16%
Forest	0.89	**0.95**	+7%
Golf Course	0.78	**0.93**	+19%
River	0.81	**0.90**	+11%
Tennis Court	0.80	**0.95**	+19%
Macro Avg	0.83	**0.95**	+14%

Ablation study: Our ablation study systematically evaluates the impact of key model components on classification performance. The full model (InceptionV3 with Global Average Pooling, 512-unit dense layer, and 0.4 dropout) achieves a 0.95 F1-score, serving as our baseline. Removing ImageNet pretrained weights results in the most significant performance drop (9.5% decrease in F1-score to 0.86), demonstrating the critical importance of transfer learning for feature extraction. Disabling data augmentation reduces performance by 4.2% (F1=0.91), particularly affecting minority classes, while removing dropout leads to a smaller but notable 3.2% decrease (F1=0.92), accompanied by increased validation loss, confirming its regularization benefits. The choice of Global Average Pooling over traditional flattening proves optimal, providing equivalent accuracy while reducing parameters by 80.6% (from 112.4M to 21.8M) and improving inference speed by 21.6%. Interestingly, reducing the dense layer size from 512 to 256 units shows minimal performance impact (ΔF1=0.01) with notable parameter reduction, suggesting a potential optimization path for resource constrained deployments. These experiments, conducted under controlled conditions with identical training protocols, validate our architectural decisions, with pretraining and augmentation emerging as the most critical factors for model performance, while the GAP/dropout combination optimally balances accuracy and computational efficiency. The results provide clear guidance for future optimizations and establish quantitative benchmarks for component contributions to overall system performance. **Table 4** summarizes the advantages of our approach compared to existing methods,

Table 4: Comparison with existing methods

Method	Avg. F1 Improvement	Processing Speed	Class-Specific Optimization
Bicubic + InceptionV3	Baseline	60 FPS	No
SRGAN (Wang 2018)	+8%	32 FPS	No
ESRGAN (Zhang 2021)	+11%	28 FPS	Partial
Proposed approach	**+14%**	**55 FPS**	**Yes**

5. CONCLUSION AND FUTURE WORK

This research presented a robust deep learning framework for satellite image classification, using ESRGAN and InceptionV3 with transfer learning and data augmentation to achieve strong performance across seven land cover categories. The model demonstrated an overall accuracy of 95%, with particularly high precision in detecting distinct features like airplanes (F1=0.99) and agricultural areas (F1=0.97). The confusion matrix revealed minimal misclassifications, primarily occurring between spectrally similar classes (e.g., golf courses vs. agricultural). Our ablation study confirmed the importance of pretrained weights (9.5% F1-score drop without them) and data augmentation (4.2% performance decrease when disabled), while Global Average Pooling (GAP) proved more efficient than flattening, reducing parameters by 80.6% without sacrificing accuracy. These findings highlight the effectiveness of our approach in automating land cover classification from low resolution satellite imagery.

To advance this research, we will focus on two key directions:

(1) **Enhanced feature discrimination** through attention mechanisms (e.g., CBAM) to improve differentiation between spectrally similar classes like urban areas and wetlands, and
(2) **Efficiency optimization** via model quantization and pruning for edge device deployment in real-time monitoring systems. These improvements will address current limitations in fine grained classification while enabling practical applications in resource-constrained environments.

REFERENCES:

1. Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Loy, C. C., Qiao, Y., & Tang, X. (2018). *ESRGAN: Enhanced super-resolution generative adversarial networks*. In Proceedings of the European Conference on Computer Vision (ECCV) Workshops. https://doi.org/10.1007/978-3-030-11021-5_5

2. Zhang, Y., Tian, Y., Kong, Y., Zhong, B., & Fu, Y. (2021). *Residual dense network for remote sensing super-resolution*. IEEE Transactions on Geoscience and Remote Sensing, 59(9), 7918-7932. https://doi.org/10.1109/TGRS.2020.3043698

3. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). *Rethinking the inception architecture for computer vision*. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 2818-2826). https://doi.org/10.1109/CVPR.2016.308

4. Liu, Z., Wang, H., Weng, L., & Yang, Y. (2022). *Ship detection from low-resolution satellite images via super-resolution and object detection*. ISPRS Journal of Photogrammetry and Remote Sensing, 183, 240-253. https://doi.org/10.1016/j.isprsjprs.2021.11.009

5. Zhao, X., Chang, X., Fan, C., Lin, M., Wei, L., & Ye, Y. (2025). *DeepDR: A Two-Level Deep Defect Recognition Framework for Meteorological Satellite Images*. Remote Sensing, 17(4), 585. https://doi.org/10.3390/rs17040585

6. Pavlović, M., et al. (2024). *A Two-Stage Deep Learning Pipeline for Estimating Soil Organic Carbon Using Sentinel-2 Imagery*. Sensors, 25(7), 1988.

7. Khan, S. D., & Basalamah, S. (2023). *Multi-Branch Deep Learning Framework for Land Scene Classification in Satellite Imagery*. Remote Sensing, 15(13), 3408.

8. Rise, B., Uney, M., & Huang, X. (2024). *Two-Stage Transfer Learning for Fusion and Classification of Airborne Hyperspectral Imagery*. In *ICASSP 2024* (pp. 6555–6559). https://doi.org/10.1109/ICASSP48485.2024.10445916

9. **Chatterjee, S., Saha, S., & Mahapatra, P. R. S. (2024).** *A Two-Stage CNN Based Satellite Image Analysis Framework for Estimating Building-Count in Residential Built-Up Area*. In *Lecture Notes in Networks and Systems: Data Management, Analytics and Innovation* (LNNS, pp. 15-30), Springer Nature. https://doi.org/10.1007/978 981 97 3245 6-2