Remote Sensing Applications: Society and Environment 43 (2026) 102201

Contents lists available at ScienceDirect

Remote Sensing Applications: Society and Environment

journal homepage: www.elsevier.com/locate/rsase

A knowledge distillation deep channel attention super-resolution network for long-term NDVI reconstruction using Landsat 8 and Sentinel-2 imagery

Allen Patnaik [a,*], Bhavesh Joshi [a], Dipima Sarma [b], M.K. Bhuyan [a], Arup Kumar Sarma [c], Karl F. MacDorman [d]

[a] Department of Electronics and Electrical Engineering, IIT Guwahati, Guwahati, 781039, Assam, India
[b] Centre for Disaster Management and Research, IIT Guwahati, Guwahati, 781039, Assam, India
[c] Department of Civil Engineering, IIT Guwahati, Guwahati, 781039, Assam, India
[d] Luddy School of Informatics, Computing and Engineering, Indiana University, Indianapolis, IN, 46202, USA

ARTICLE INFO

Dataset link: https://github.com/allenptnk/KD-DCASRN

Keywords:
Channel attention
Deep learning
Knowledge distillation
NDVI super-resolution
Landsat
Sentinel-2
Vegetation monitoring

ABSTRACT

The normalized difference vegetation index (NDVI) is vital for monitoring vegetation health. However, NDVI imagery imposes trade-offs in spatial resolution, temporal coverage, and historical depth. Sentinel-2 MSI has provided 10-m imagery since 2015, while the Landsat archive extends to 1982, though its 30-m resolution limits fine-scale vegetation analysis. We propose a knowledge distillation strategy to enhance Landsat 8 imagery to match Sentinel-2 resolution. Our novel deep channel attention super-resolution network (DCASRN) serves as a lightweight student model guided by a deeper teacher, the widely used enhanced deep super-resolution network (EDSR). We experimentally validated the student model's accuracy across diverse land-cover types, including dense forests, wetlands and water bodies, agricultural fields, urban settlements, sparse vegetation, and open land, using study regions in Assam (Palashbari and Majuli) and Rajasthan (Ajmer), India. KD-DCASRN outperformed all baseline methods across the three study regions, achieving an average PSNR gain of approximately 7.97 dB over Bicubic, 3.92 dB over ESRT, 1.70 dB over SRCNN, 0.48 dB over EDSR, and 0.35 dB over the non-distilled DCASRN model. Correspondingly, KD-DCASRN reduced RMSE by 56.6%, 29.8%, 15.0%, 9.5%, and 8.1%, respectively, demonstrating its effectiveness for high-resolution NDVI reconstruction. Qualitative assessments show more structurally faithful reconstructions. The model's strong cross-biome performance demonstrates its practical utility in vegetation monitoring.

1. Introduction

The limited spatial resolution of long-term satellite datasets like Landsat's normalized difference vegetation index (NDVI) challenges fine-scale vegetation monitoring and precision agriculture applications. Although Landsat archives offer consistent observations since 1982, their 30-m resolution does not adequately capture detailed patterns crucial for monitoring small agricultural plots, urban greenery, and diverse land covers. In contrast, high spatial resolution sensors, such as Sentinel-2, launched in 2015, provide 10-m NDVI data, enabling detailed mapping and accurate vegetation assessment. However, Sentinel-2 lacks the historical

* Corresponding author.
E-mail address: allen.patnaik@iitg.ac.in (A. Patnaik).

https://doi.org/10.1016/j.rsase.2026.102201
Received 22 June 2025; Received in revised form 22 July 2026; Accepted 6 August 2026
Available online 12 August 2026

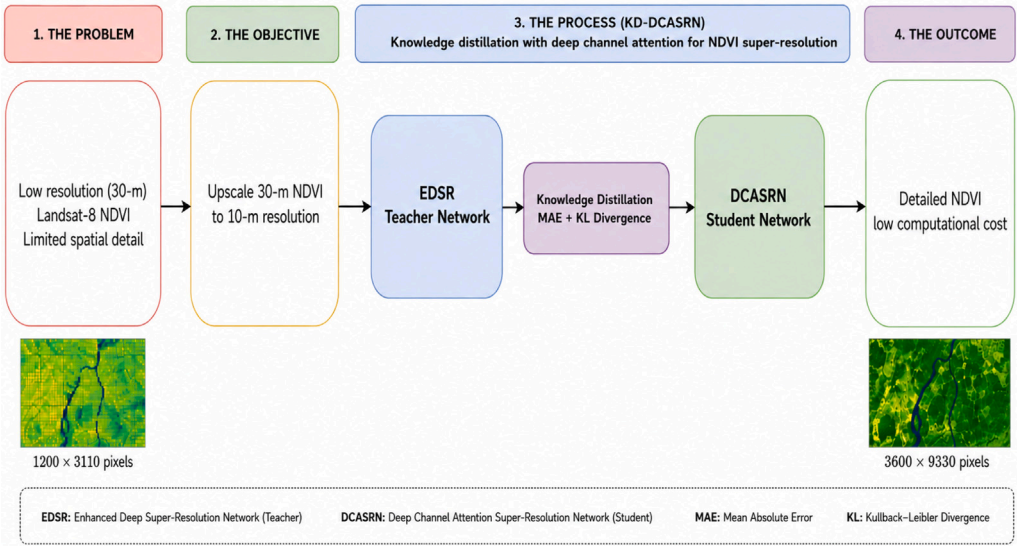


Fig. 1. Conceptual overview of the proposed framework.

depth of Landsat records. Their distinct orbital paths and revisit cycles offer complementary data acquisitions, which if integrated, could enhance temporal coverage.

These trade-offs in spatial resolution, temporal coverage, and historical depth pose a critical challenge. Researchers and practitioners require datasets that combine fine spatial detail with long-term temporal consistency for tasks like crop yield prediction, drought monitoring, and land degradation assessment (Al-Shammari et al., 2021; Henn and Peduzzi, 2024; Karmakar et al., 2024; Segarra, 2024; Zhang et al., 2020). Thus, to support data-driven decision making in agriculture and environmental management, super-resolving low-resolution imagery is essential (Crivellari et al., 2023; Rohith and Kumar, 2021).

Super-resolution comprises methods for reconstructing high-resolution images from low-resolution counterparts. These methods can be divided into classical and deep learning approaches. Classical methods for low-resolution satellite imagery mainly rely on resampling techniques, such as nearest-neighbor, bilinear, and bicubic interpolation (Keys, 1981; Matthews and Prate, 2024). Classical methods, though computationally simple and prevalent, are limited in reconstructing missing spatial details. They tend to smooth out fine structures, leading to blurred outputs and loss of critical texture. Several statistical and regression-based approaches have been proposed to improve on naive resampling. Methods such as regression kriging, geographically weighted regression, and Bayesian data fusion integrate spatial dependencies during upscaling (Wu et al., 2024). However, these classical methods lack robustness across varying land cover types and degrade when applied to complex scenes, such as mixed vegetation, urban–rural interfaces, and fragmented agricultural plots. Recent evaluations confirm that interpolation and regression-based techniques often show inconsistent performance, particularly in environments with high spatial variability.

Deep learning-based super-resolution models have emerged as effective tools for enhancing satellite imagery, with advantages over traditional methods (Brandt et al., 2025; Goodarzi and Sahebi, 2025; Palan et al., 2025; Sun et al., 2024a; Y. Zhao et al., 2023). Models based on convolutional neural networks (CNNs; Han et al., 2024; Pouliot et al., 2018; Liang et al., 2021; Xiong et al., 2023), generative adversarial networks (GANs; Wang et al., 2018; Pham and Bui, 2021; Wang et al., 2023, 2024; Q. Zhao et al., 2023), and multiscale fusion (Liu et al., 2025; Patnaik et al., 2024; Sun et al., 2024b) have delivered promising results in recovering fine spatial details in Earth observation tasks. Despite these advances, current NDVI super-resolution methods struggle to balance computational efficiency and reconstruction accuracy across heterogeneous landscapes, particularly for applications involving historical satellite datasets across climatically and ecologically diverse regions.

To improve performance, researchers have applied knowledge distillation (KD) to existing models, opening a new line of research in NDVI monitoring (Cui et al., 2025; Esmaeilzehi et al., 2025; Gong et al., 2023; Liu et al., 2024). Knowledge distillation transfers knowledge from a trained teacher model to a specialized student model, enabling the student to retain much of the teacher's representational capacity while reducing computational cost. Despite these advances, NDVI super-resolution remains constrained by limited generalizability, data scarcity, temporal inconsistency, and reduced accuracy for sparse or mixed vegetation. To address these unresolved challenges, we propose a knowledge-distillation framework built around deep channel attention, enabling a lightweight student model to preserve spatial detail while improving computational efficiency and robustness across diverse Landsat NDVI environments.

Specifically, the proposed framework uses a novel deep channel attention super-resolution network (DCASRN) as the student model to upscale Landsat 8 OLI/TIRS 30-m NDVI data to Sentinel-2 MSI 10-m resolution. The network's channel attention (CA) module adaptively emphasizes informative features, enabling DCASRN's compact architecture to outperform the highly ranked enhanced deep super-resolution network (EDSR; Lim et al., 2017) on satellite imagery (Khankeshizadeh et al., 2024; Tahermanesh

et al., 2025). Our network achieves this with a fraction of the parameters. This reduction in computational overhead makes DCASRN a more scalable, energy-efficient approach to enhancing spatial detail across the large geographical and historical datasets required for vegetation monitoring.

We further enhanced our network's performance by distilling knowledge from EDSR, recognizing that the two models have complementary strengths: EDSR's deep structure of 32 residual blocks captures rich hierarchical representations, while channel attention enables DCASRN to learn contextual features efficiently. EDSR serves as the teacher to guide DCASRN as the student, directing the student's attentive learning mechanism with the teacher's rich feature knowledge. Through this teacher–student transfer, DCASRN with knowledge distillation (KD-DCASRN) learns high-frequency spatial details more effectively and produces outputs closer to the reference images than either EDSR or DCASRN.

Fig. 1 presents the conceptual overview of the KD-DCASRN framework for NDVI super-resolution. It illustrates the overall problem-to-solution workflow, where low-resolution Landsat 8 NDVI imagery is enhanced to high-resolution NDVI using a knowledge distillation architecture. The proposed approach aims to improve spatial detail reconstruction while maintaining computational efficiency, thereby supporting long-term vegetation monitoring applications.

To validate our model, we used paired Landsat 8–Sentinel-2 samples from diverse regions of India: Palashbari (March 2017–April 2021) and Majuli (March 2022–January 2024) in Assam and Ajmer (April 2017–April 2021) in Rajasthan. In quantitative evaluations against widely used methods, including bicubic interpolation, super-resolution CNN (SRCNN; Dong et al., 2014), EDSR, and our own DCASRN without knowledge distillation, our model emerged as the best performer. Specifically, KD-DCASRN achieved the highest average scores across four metrics: peak signal-to-noise ratio (PSNR), structural similarity index measure (SSIM), root mean square error (RMSE), and mean absolute error (MAE). In qualitative evaluations, KD-DCASRN showed superior visual reconstruction of NDVI maps, preserving fine spatial details and structural nuances across diverse landscapes, such as narrow rivers, field boundaries, and mixed vegetation zones. These results, attained without compromising spatial fidelity or computational efficiency, establish KD-DCASRN as a robust practical solution to enhance low-resolution NDVI data in large-scale vegetation monitoring.

The main contributions of the study are as follows:

1. We present a novel knowledge distillation strategy to enhance the spatial resolution of NDVI imagery. A deeper teacher network (EDSR) guides a lightweight student network (DCASRN) to reconstruct high-fidelity NDVI data from low-resolution sources.
2. Our novel student network is computationally efficient. Its channel attention mechanism effectively captures spatial details without the parameter overhead of larger models, making it suitable for large-scale historical analysis.
3. Through quantitative evaluation using PSNR, SSIM, RMSE, and MAE, KD-DCASRN outperforms traditional upsampling methods (e.g., bicubic interpolation), standard deep learning models (SRCNN, ESRT), its deeper teacher (EDSR), and DCASRN without knowledge distillation. Qualitative results further confirm that the model generates reconstructions more structurally faithful to the reference image.
4. We validate the model's effectiveness and robustness using a curated dataset from Palashbari and Majuli in Assam and Ajmer in Rajasthan, India, representing humid floodplain and semi-arid environments. Together, these sites cover dense forests, wetlands and water bodies, agricultural land, urban and mixed-use areas, sparse vegetation, and open land. Its consistent performance across these biomes demonstrates its practical utility for vegetation monitoring.

2. Methodology

The NDVI is a widely used metric in remote sensing to assess vegetation health and monitor land cover change. It quantifies vegetation by measuring the difference between near-infrared (NIR) light, which vegetation reflects, and red light, which it absorbs.

NDVI is calculated from reflectance values in the NIR and red portions of the electromagnetic spectrum using

$$\mathrm{NDVI} = \frac{\rho_{\mathrm{NIR}} - \rho_{\mathrm{red}}}{\rho_{\mathrm{NIR}} + \rho_{\mathrm{red}}}$$

where ρ_{NIR} is the spectral reflectance in the near-infrared band, and ρ_{red} is the spectral reflectance in the red band.

We aim to enhance the spatial resolution of NDVI data from Landsat 8 OLI/TIRS (30 m) to match the finer detail of Sentinel-2 MSI imagery (10 m), used as reference high-resolution for training and evaluation. To this end, we constructed a dataset using temporally aligned satellite images from Palashbari (March 2017–April 2021) and Majuli (March 2022–January 2024) in Assam, and Ajmer (April 2017–April 2021) in Rajasthan. We formulated the task as 3 × super-resolution, enabling the model to reconstruct fine-grained spatial details lost in low-resolution inputs.

Let $X \in \mathbb{R}^{h \times w}$ denote a low-resolution (LR) NDVI image and $Y \in \mathbb{R}^{kh \times kw}$ the corresponding high-resolution (HR) image, where k is the scaling factor and $h \times w$ are the image dimensions.

Before training to map the relation between LR and HR images, data preprocessing is performed, as shown in Fig. 2. The preprocessing step ensures proper data alignment and formatting before the data are fed into the super-resolution model.

Invalid or missing pixel values in X and Y, typically encoded as extreme constants, are first replaced with the mean of the valid pixels in each image. This ensures numerical stability and consistent model learning.

The extracted LR–HR patch pairs are formatted as paired tensors, with LR inputs of shape $(N, H, W, 1)$ and corresponding HR targets of shape $(N, kH, kW, 1)$, where N is the number of patch pairs, H and W are the LR patch dimensions, $k = 3$ is the spatial-resolution scaling factor, and 1 represents a single NDVI band.

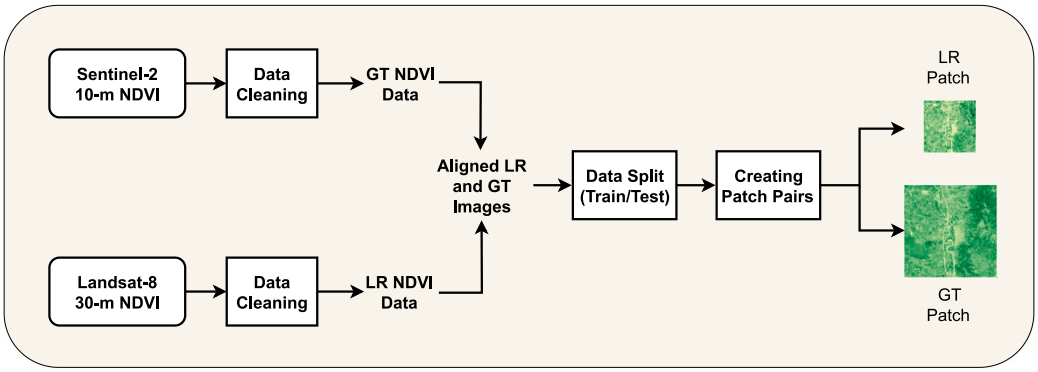
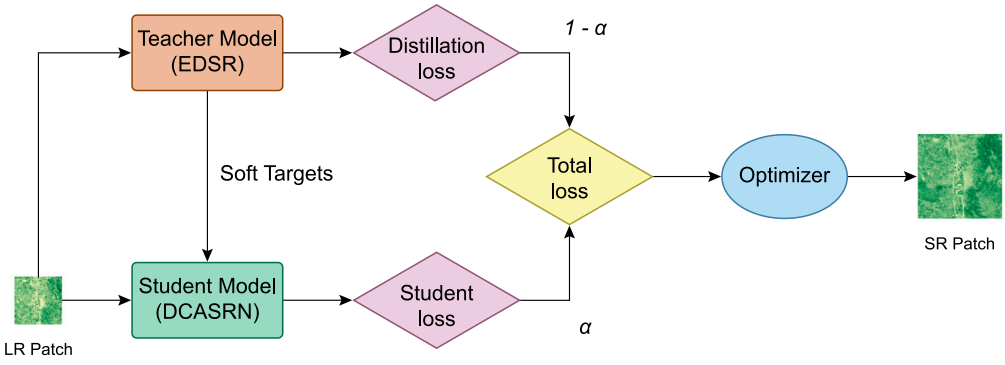
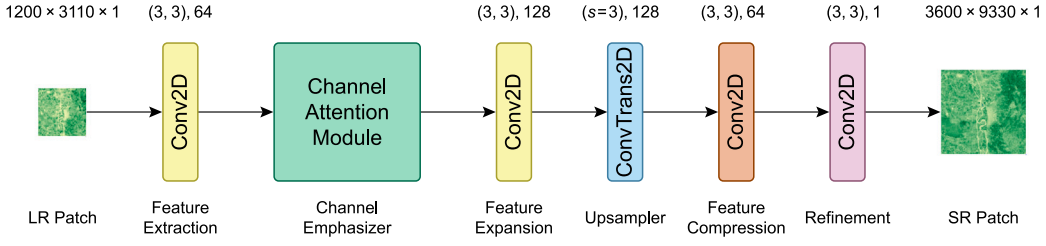


Fig. 2. Data preprocessing.



(a) Knowledge distillation (KD-DCASRN) approach



(b) Student model (DCASRN) architecture

Fig. 3. Knowledge distillation deep channel attention super-resolution network.

We propose a knowledge distillation framework to train a deep channel attention super-resolution network (KD-DCASRN), as shown in Fig. 3(a). Our network uses a student model (DCASRN) to extract knowledge from a teacher model. We selected EDSR (Lim et al., 2017) with 32 residual blocks as the teacher because of its high-fidelity output.

The student was trained using a combined objective consisting of a mean absolute error (MAE) reconstruction term and a Kullback–Leibler (KL) divergence distillation term. Because NDVI reconstruction is a continuous regression task, KL divergence was not applied directly to the raw NDVI values. Instead, before computing the KL-divergence term, the teacher and student outputs were converted into temperature-softened response distributions using a temperature-scaled softmax function. The KL-divergence term therefore acted only as a distillation loss, encouraging the compact student network to reproduce the teacher model's relative spatial response pattern, while the MAE term preserved sensitivity to pixel-level NDVI reconstruction errors. The combined objective was minimized using the Adam optimizer.

The novel student model, shown in Fig. 3(b), is designed to learn a mapping $\mathcal{F}_\theta : I_{LR} \rightarrow \hat{I}_{SR}$, where $I_{LR} \in \mathbb{R}^{H\times W\times 1}$ denotes the input low-resolution NDVI image and $\hat{I}_{SR} \in \mathbb{R}^{kH\times kW\times 1}$ denotes the predicted high-resolution NDVI output. Unless otherwise stated, all 3×3 convolutional layers use same padding to preserve spatial dimensions, and the transposed convolution uses stride $k = 3$ with padding and output settings chosen to map feature maps from $H \times W$ to $kH \times kW$. The architecture comprises the following components:

Initial Feature Extraction: A 3×3 convolution layer with ReLU activation extracts spatial features from the input:

$$F_1 = \text{ReLU}(\text{Conv}_{3\times3}(I_{LR})) \tag{1}$$

This operation increases the number of input channels from 1 to 64 by generating feature maps that capture initial details.

Channel emphasizer block: This block emphasizes important feature channels. It uses global average pooling (GAP) followed by two dense layers to compute a channelwise attention map used to compress channelwise features into vector form:

$$GAP_c(F_1) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} F_1(i, j, c), \tag{2}$$

$$F_{ca} = F_1 \cdot \sigma \left(W_2 \cdot \text{ReLU}(W_1 \cdot \text{GAP}(F_1))\right) \tag{3}$$

where W_1 and W_2 are dense layer weights, σ is a sigmoid activation function, and $\cdot$ denotes channelwise multiplication with the attention weights broadcast across spatial dimensions.

Feature Expansion: A 3×3 convolutional layer expands the feature depth of attention-weighted features, increasing channels from 64 to 128. This enables the network to learn more complex patterns and finer details:

$$F_2 = \text{ReLU}(\text{Conv}_{3\times3}(F_{ca})) \tag{4}$$

Upsampling Module: A transposed convolution layer with a stride of 3 increases spatial resolution by a scale factor of 3. It performs learned upsampling to reconstruct high-resolution details from low-resolution inputs:

$$F_{\text{up}} = \text{ReLU}\left(\text{ConvTranspose}_{3\times3,\, s=3}(F_2)\right) \tag{5}$$

Feature Compression: After upsampling, a 3×3 convolutional layer compresses the feature maps, reducing the channels from 128 to 64 while refining spatial structures for better reconstruction:

$$F_{\text{comp}} = \text{ReLU}(\text{Conv}_{3\times3}(F_{\text{up}})) \tag{6}$$

Refinement: A final 3×3 convolutional layer refines the compressed features and reduces the number of channels from 64 to 1, producing the reconstructed high-resolution output with enhanced spatial details:

$$\hat{I}_{SR} = \text{Conv}_{3\times3}(F_{\text{comp}}) \tag{7}$$

The student model is trained using a weighted combination of supervised reconstruction loss and knowledge distillation loss:

$$\mathcal{L}_{\text{total}} = \alpha\mathcal{L}_{\text{student}} + (1 - \alpha)\mathcal{L}_{\text{KD}}, \tag{8}$$

where α controls the contribution of the supervised reconstruction loss and $(1 - \alpha)$ controls the contribution of the knowledge distillation loss. The supervised reconstruction loss is defined as

$$\mathcal{L}_{\text{student}} = \frac{1}{NP} \sum_{i=1}^{N} \sum_{p=1}^{P} \left| \hat{I}_{\text{SR}}^{(i)}(p) - I_{\text{HR}}^{(i)}(p) \right|, \tag{9}$$

where N is the number of training patches, P is the number of pixels in each high-resolution patch, $\hat{I}_{\text{SR}}^{(i)}$ denotes the student prediction, and $I_{\text{HR}}^{(i)}$ represents the corresponding Sentinel-2 reference NDVI image.

To perform knowledge distillation, the teacher and student outputs are first vectorized across their spatial dimensions and converted into softened probability distributions using a temperature-scaled softmax function. Specifically, the teacher and student outputs $T^{(i)}$ and $\hat{I}_{\text{SR}}^{(i)}$, each having dimensions $kH \times kW \times 1$, are flattened into vectors containing $P = kHkW$ spatial locations. The softened teacher and student distributions are defined as:

$$P_T^{(i)} = \text{softmax}\left(\frac{\text{vec}\left(T^{(i)}\right)}{\tau} \right), \tag{10}$$

$$P_S^{(i)} = \text{softmax}\left(\frac{\text{vec}\left(\hat{I}_{\text{SR}}^{(i)}\right)}{\tau} \right), \tag{11}$$

where $\text{vec}(\cdot)$ denotes vectorization across the spatial dimensions, $P = kHkW$ is the number of spatial locations, and τ is the temperature parameter.

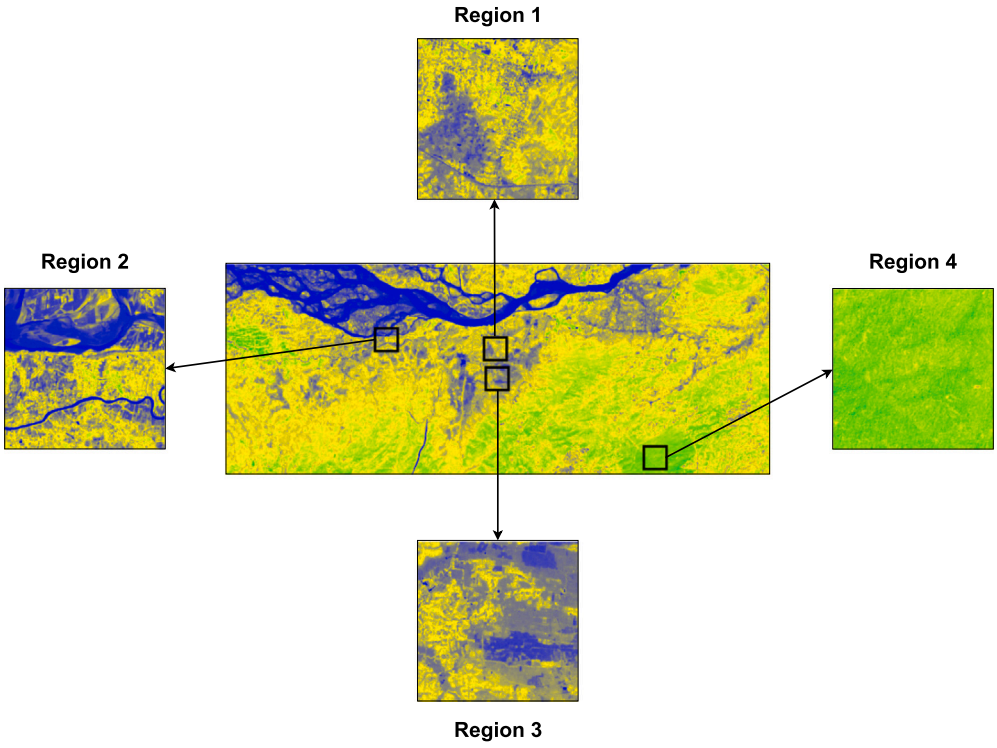


Fig. 4. Palashbari area.

The distillation loss is computed using the Kullback–Leibler (KL) divergence between the softened teacher and student spatial distributions:

$$\mathcal{L}_{\mathrm{KD}} = \frac{\tau^2}{N} \sum_{i=1}^{N} \sum_{p=1}^{P} P_T^{(i)}(p) \log \frac{P_T^{(i)}(p)}{P_S^{(i)}(p)}, \tag{12}$$

where $P_T^{(i)}(p)$ and $P_S^{(i)}(p)$ denote the teacher and student probabilities at spatial location p, respectively. The temperature-scaled softmax therefore converts the teacher and student outputs into softened spatial response distributions, enabling the student network to learn the relative spatial response pattern encoded by the teacher predictions.

3. Experiment

To evaluate the efficacy and scalability of our super-resolution NDVI enhancement model, we selected contrasting regions in India: Palashbari and Majuli in Assam and Ajmer in Rajasthan. These areas represent a spectrum of vegetation densities and land cover types, enabling a comprehensive assessment of our model's performance across varying ecological conditions. We compared our model, KD-DCASRN, with bicubic interpolation, SRCNN (Dong et al., 2014), ESRT (Lu et al., 2022), EDSR (Lim et al., 2017), and DCASRN without knowledge distillation.

3.1. Study area

Palashbari is highly vulnerable to erosion and has undergone significant riverbank protection interventions. Four regions are selected within this area, given in Figs. 4 and 5.

Region 1: Bijoynagar — A rapidly urbanizing zone with expanding infrastructure and reduced green cover. NDVI values often fall below 0.3, reflecting sparse vegetation. This region represents urban environments, testing the model's ability to reconstruct NDVI with limited vegetation.

Region 2: Hatipara — A mix of vegetated land and water bodies, including wetlands and small lakes. NDVI values range from 0.05 to 0.5, indicating moderate vegetation interspersed with aquatic features. The mixed land–water interface challenges NDVI reconstruction, providing a suitable case for evaluating model performance.

Region 3: Sikrhari — A rural zone with sparse vegetation and open land. NDVI values range from 0.2 to 0.4, reflecting limited vegetative cover. This region tests the model's ability to reconstruct NDVI in areas with minimal vegetation.

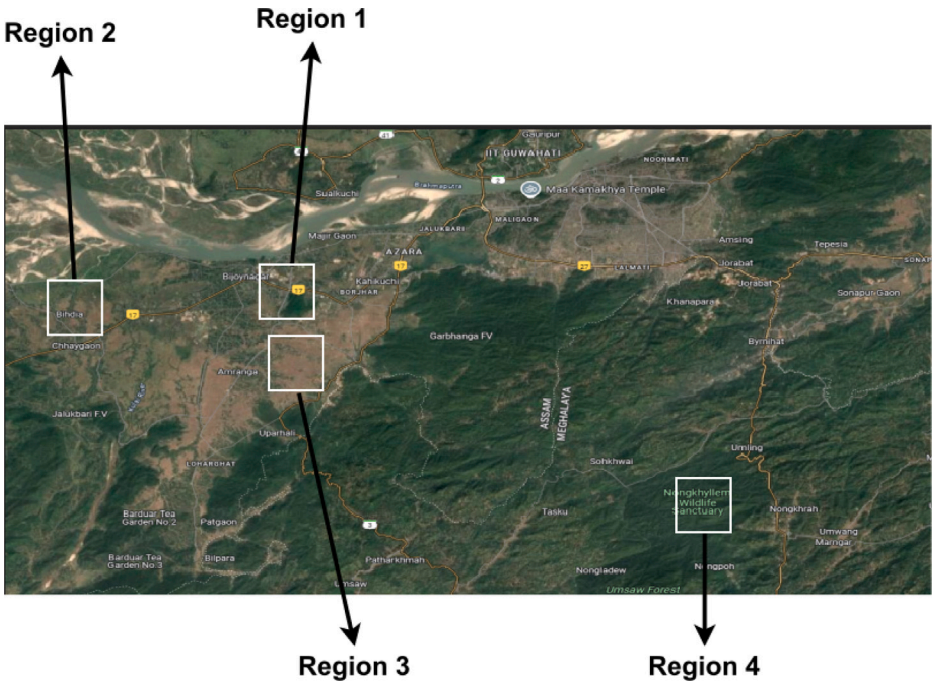


Fig. 5. Palashbari image area in map.

Region 4: Nongkhyllem Wildlife Sanctuary — A protected area with subtropical forests and rich biodiversity. NDVI values typically range from 0.6 to 0.9, indicating dense, healthy vegetation. As a benchmark for forested regions, this area evaluates the model's performance in reconstructing high NDVI values.

Majuli, the world's largest river island, lies on the Brahmaputra River and is known for its unique ecological and cultural landscape. Two regions were selected within this area, shown in Figs. 6 and 7.

Region 1: Pichola Forestry Area — A densely forested zone comprising a mix of evergreen and deciduous trees. NDVI values typically range from 0.6 to 0.9, indicating robust vegetative health. This region assesses the model's effectiveness in reconstructing NDVI in dense forest ecosystems.

Region 2: Bihpuria — A heterogeneous landscape with water bodies, urban settlements, and sparse vegetation. NDVI values range from 0.2 to 0.5, reflecting mixed land cover. This region tests the model's adaptability to complex environmental conditions.

To further evaluate the generalization capability of the proposed framework across different environmental conditions, an additional study area was selected in Ajmer, Rajasthan, India, as shown in Figs. 8 and 9. Unlike the floodplain and agricultural landscapes of Assam, Ajmer represents a semi-arid environment characterized by heterogeneous land-cover classes, including urban settlements, sparse vegetation, open land, and water bodies. The inclusion of this region enables assessment of the model's performance across contrasting climatic and land-cover conditions.

Region 1: Adarsh Nagar Area — A mixed land-use region containing residential areas, vegetation patches, open land, and nearby water bodies. NDVI values vary from low values over built-up and water regions to moderate values over vegetated areas. This region assesses the model's ability to reconstruct NDVI accurately across heterogeneous landscapes.

By selecting these varied regions, we aim to evaluate our super-resolution model's performance across different biomes, ensuring its robustness and applicability to real-world scenarios.

3.2. Training settings

Low-resolution NDVI data (1200 × 3110 pixels) were derived from Landsat 8 OLI/TIRS Level-2 surface reflectance imagery at 30-m spatial resolution, while high-resolution NDVI reference data (3600 × 9330 pixels) were generated from Sentinel-2 MSI imagery at 10-m resolution. Data were obtained from USGS EarthExplorer for Landsat 8 imagery and the Copernicus Open Access Hub for Sentinel-2 imagery. The NDVI maps were spatially aligned to ensure pixel-level correspondence for training the super-resolution model. Dataset characteristics and preprocessing are summarized in Table 1, while the training configuration and hyperparameters are presented in Table 2.

The baseline and teacher models were trained on paired low- and high-resolution NDVI data using MSE loss, while the proposed student model was trained using the combined reconstruction and distillation objective described above. Training ran for 30 epochs with a batch size of 16. A default learning rate of 0.001 was used. Training was conducted using the Adam optimizer using the TensorFlow framework on an NVIDIA GeForce RTX 3080 GPU.

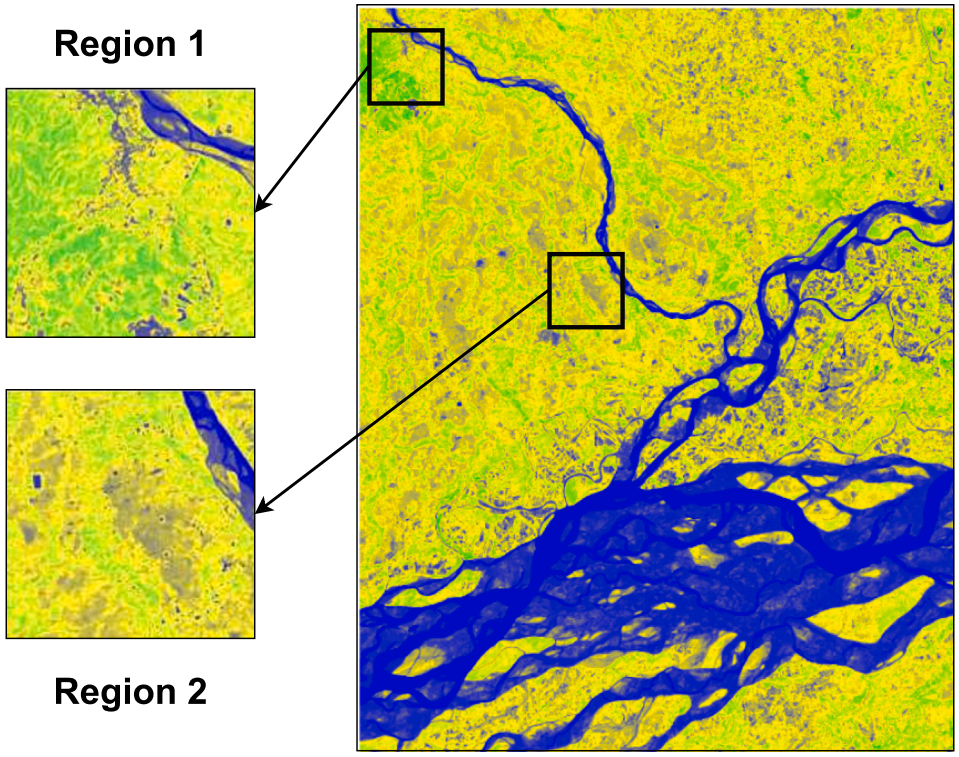


Fig. 6. Majuli study area.

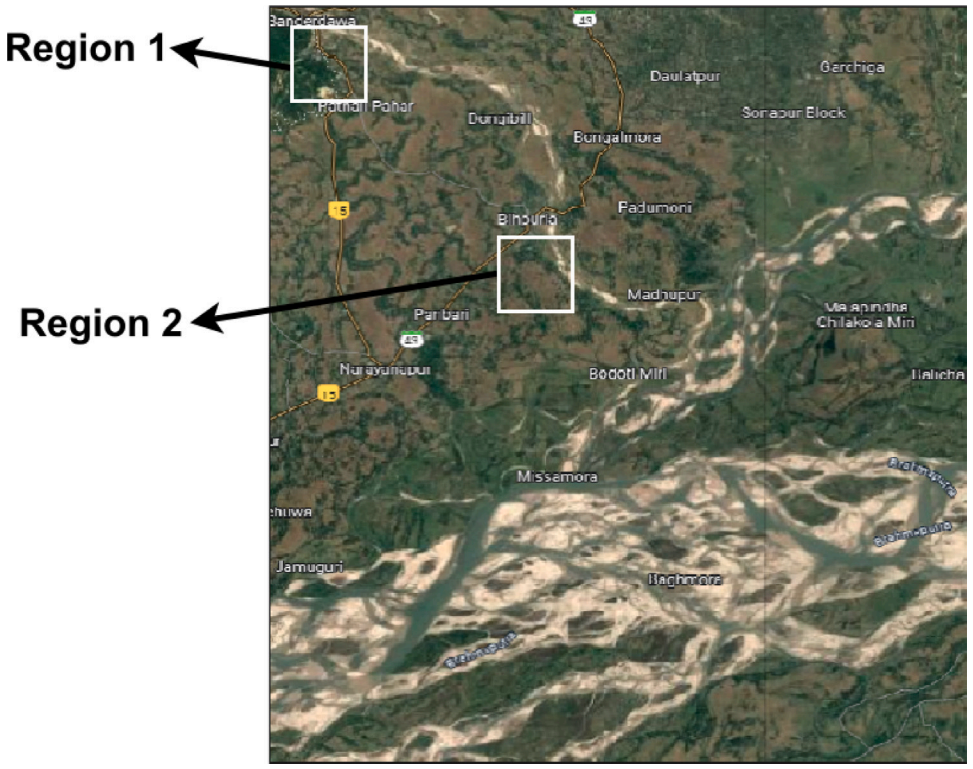


Fig. 7. Majuli image area in map.

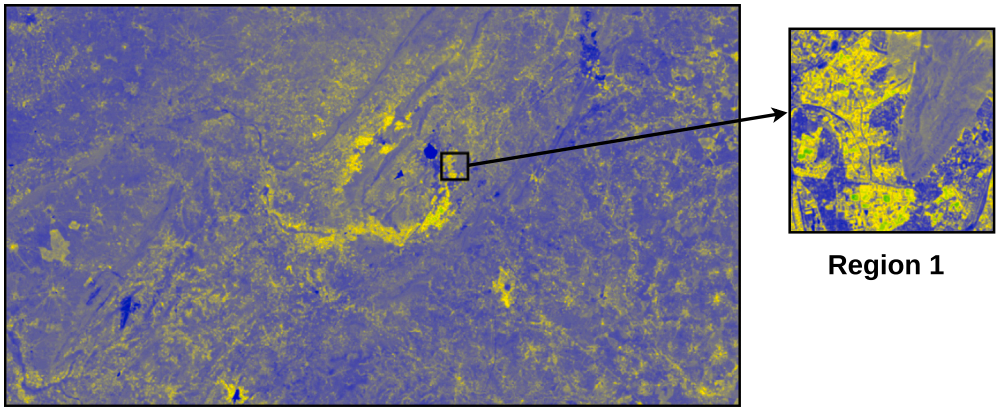


Fig. 8. Ajmer study area.

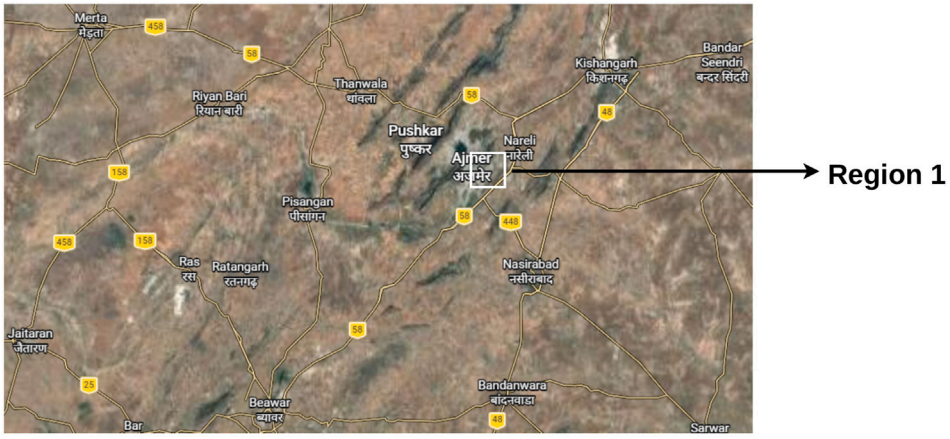


Fig. 9. Ajmer image area in map.

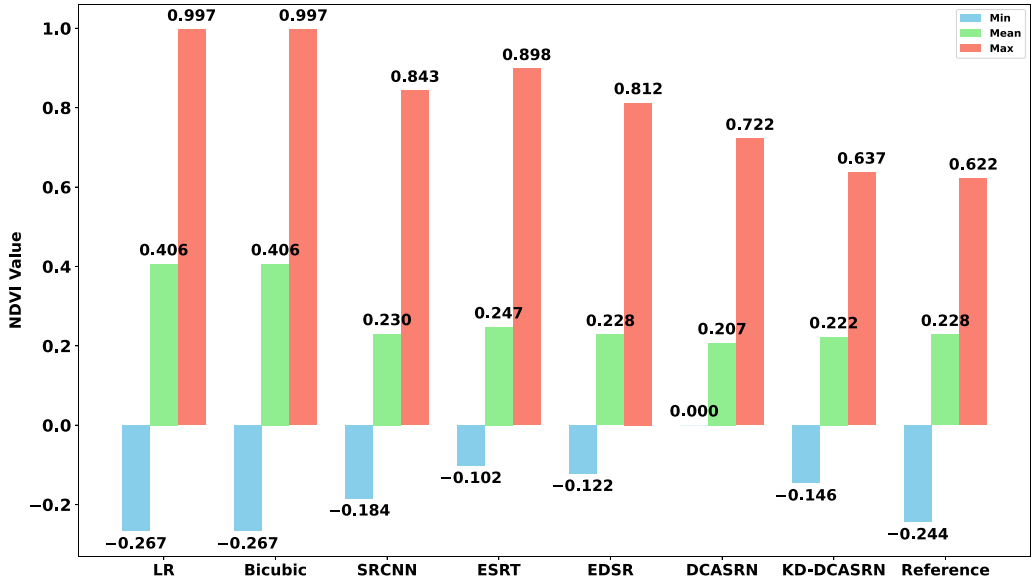


Fig. 10. NDVI value comparison for Palashbari, March 2017.

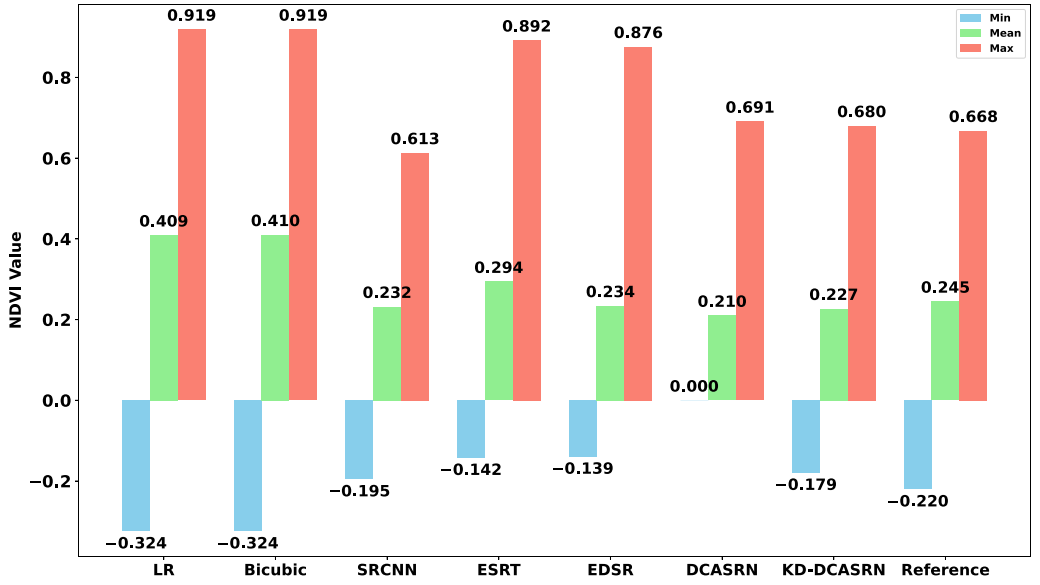


Fig. 11. NDVI value comparison for Majuli, March 2022.

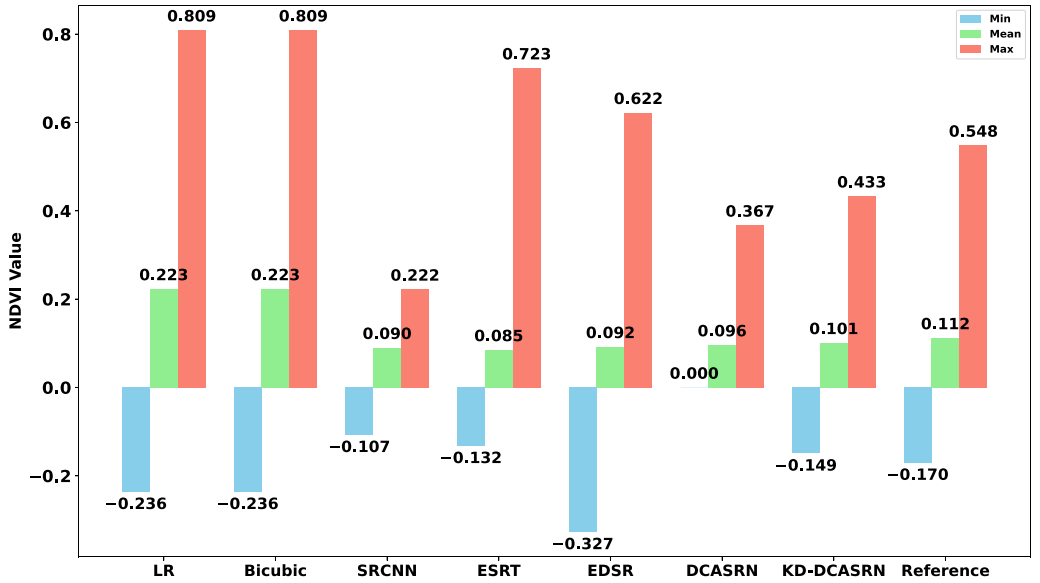


Fig. 12. NDVI value comparison for Ajmer, April 2021.

3.3. Evaluation metrics

The experimental results were evaluated using the following metrics:

3.3.1. Peak signal to noise ratio

$$\mathrm{PSNR}_i = 10\log_{10}\left(\frac{\mathrm{MAX}_I^2}{\mathrm{MSE}_i}\right) \tag{13}$$

where MSE_i is the mean squared error for image i, and MAX_I is the maximum possible pixel value of the image.

Table 1
Dataset summary and preprocessing details.

Parameter	Description
Study region	Palashbari and Majuli, Assam; Ajmer, Rajasthan
Sensor pair	Landsat 8 OLI/TIRS (30 m) and Sentinel-2 MSI (10 m)
Date range	2017–2024
Training samples	1728 paired patches
LR patch size	128 × 128 × 1
HR patch size	384 × 384 × 1
Preprocessing	Pixel alignment, no-data correction, and patch extraction

Table 2
Hyperparameters and training settings used for NDVI super-resolution.

Parameter	Value
Low-resolution input data	Landsat 8 OLI/TIRS NDVI (30 m)
High-resolution target data	Sentinel-2 NDVI (10 m)
Low-resolution image size	1200 × 3110 pixels
High-resolution image size	3600 × 9330 pixels
Scale factor	3×
Optimizer	Adam
Learning rate	0.001
Batch size	16
Epochs	30
Student loss function	Mean absolute error (MAE)
Distillation loss function	Kullback–Leibler (KL) divergence
Student loss weight (α)	0.1
Temperature parameter (τ)	10
Hardware	NVIDIA GPU-enabled system

3.3.2. Structural similarity index measure

$$\mathrm{SSIM}(s,h) = \frac{(2\mu_s\mu_h + c_1)(2\sigma_{sh} + c_2)}{(\mu_s^2 + \mu_h^2 + c_1)(\sigma_s^2 + \sigma_h^2 + c_2)} \tag{14}$$

where μ_s and μ_h denote the mean values of the super-resolved image s and the reference high-resolution image h, σ_s^2 and σ_h^2 are their variances, and σ_{sh} is the covariance between them. The constants $c_1 = (k_1 L)^2$ and $c_2 = (k_2 L)^2$ are used to stabilize division when the denominator is close to zero, where L is the dynamic range of pixel values, and k_1 and k_2 are small scalar constants (typically 0.01 and 0.03).

3.3.3. Root mean square error and mean absolute error

RMSE calculates the square root of the average squared difference between corresponding pixels in the predicted and reference images:

$$\mathrm{RMSE} = \sqrt{\frac{1}{P}\sum_{p=1}^{P}\left(I_{\mathrm{ref}}(p) - I_{\mathrm{pred}}(p)\right)^2} \tag{15}$$

and MAE computes the average absolute difference:

$$\mathrm{MAE} = \frac{1}{P}\sum_{p=1}^{P}\left|I_{\mathrm{ref}}(p) - I_{\mathrm{pred}}(p)\right| \tag{16}$$

where $I_{\mathrm{ref}}(p)$ is the pixel value at position p in the reference image, $I_{\mathrm{pred}}(p)$ is the pixel value at position p in the predicted (upscaled) image, and P is the total number of pixels.

3.4. Results

After training, we evaluated super-resolution model performance on unseen test images by calculating PSNR, SSIM, RMSE, and MAE between reconstructed and reference images, and compared the results with bicubic resampling. Higher PSNR and SSIM values and lower RMSE and MAE indicate better performance. For long-term observation, we explored minimum, maximum, and mean NDVI values to assess closeness to reference for Palashbari Fig. 10, Majuli Fig. 11, and Ajmer Fig. 12. Over different time periods, our model reconstructed images with NDVI value spans closest to the reference image.

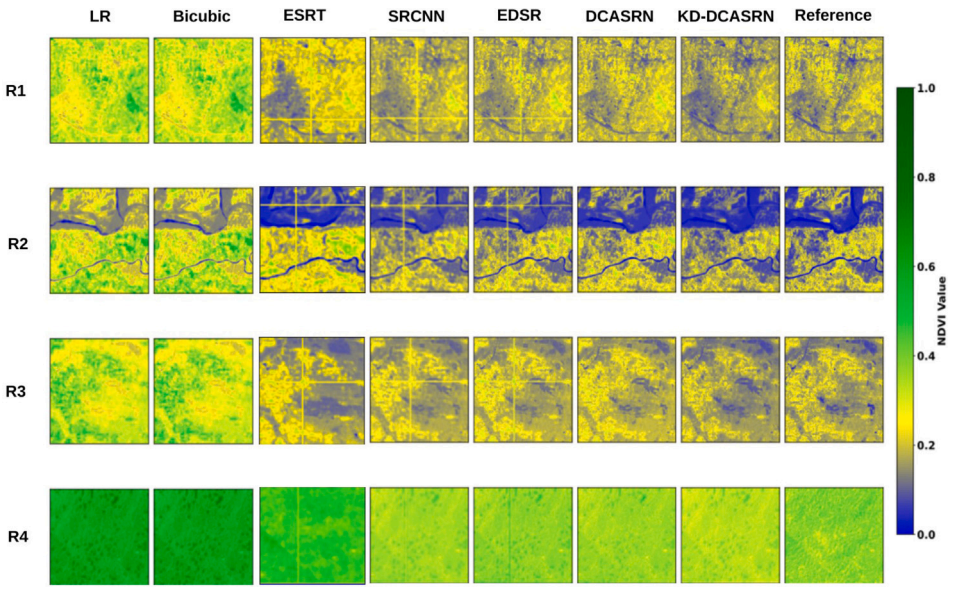


Fig. 13. Qualitative comparison for Palashbari, March 2017.

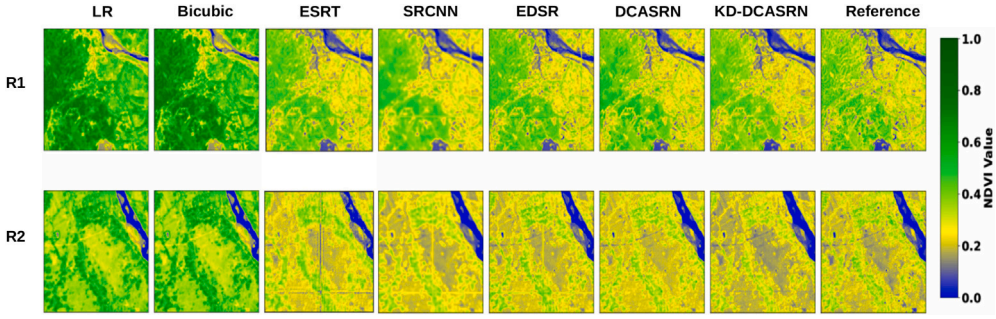


Fig. 14. Qualitative comparison for Majuli, March 2022.

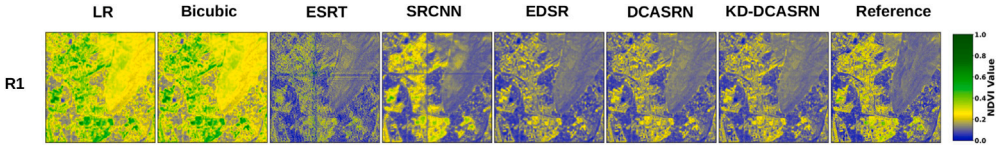


Fig. 15. Qualitative comparison for Ajmer, April 2021.

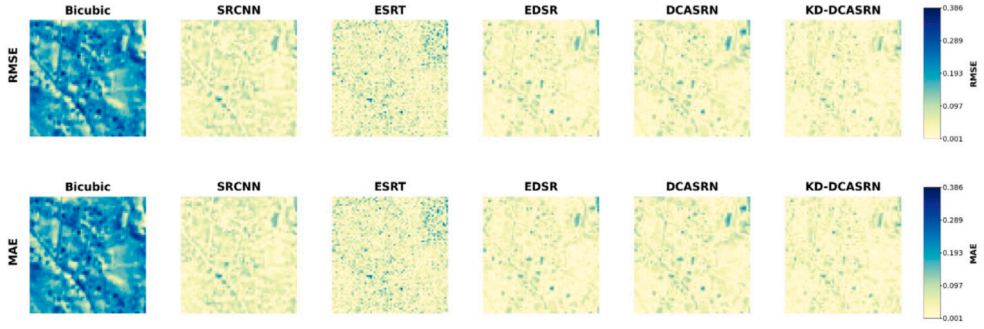


Fig. 16. Error map between reference image and super-resolution model for Palashbari.

Table 3
Quantitative comparison of different models for Palashbari.

Method	PSNR			SSIM			RMSE			MAE			Trainable params
	Mar. 17	Mar. 19	Apr. 21	Mar. 17	Mar. 19	Apr. 21	Mar. 17	Mar. 19	Apr. 21	Mar. 17	Mar. 19	Apr. 21	
Bicubic	12.92	13.29	12.68	0.4912	0.5058	0.4865	0.2310	0.2240	0.2011	0.2085	0.2069	0.1861	0
SRCNN	19.50	21.01	19.14	0.6031	0.5916	0.5970	0.0998	0.0931	0.1012	0.0691	0.0585	0.0688	103 731
ESRT	19.03	19.11	18.92	0.6011	0.5898	0.5723	0.1091	0.1128	0.0998	0.0922	0.0877	0.0984	659 849
EDSR	21.19	21.17	20.25	0.6461	0.6658	0.6102	0.0925	0.0579	0.0921	0.0592	0.0920	0.0570	684 961
DCASRN	21.25	21.35	20.31	0.6502	0.6690	0.6108	0.0897	0.0577	0.0911	0.0538	0.0891	0.0501	297 545
KD-DCASRN	**21.35**	**21.58**	**20.46**	**0.6690**	**0.7172**	**0.6121**	**0.0786**	**0.0886**	**0.0820**	**0.0522**	**0.0522**	**0.0489**	297 545

Table 4
Quantitative comparison of different models for Majuli.

Method	PSNR			SSIM			RMSE			MAE			Trainable params
	Mar. 22	Mar. 23	Jan. 24	Mar. 22	Mar. 23	Jan. 24	Mar. 22	Mar. 23	Jan. 24	Mar. 22	Mar. 23	Jan. 24	
Bicubic	14.13	13.06	13.30	0.5758	0.5412	0.4972	0.1626	0.1620	0.1965	0.1465	0.1438	0.1793	0
SRCNN	16.56	18.10	15.95	0.5829	0.6010	0.5230	0.1234	0.1049	0.1310	0.0859	0.0638	0.0920	103 731
ESRT	16.52	16.44	15.40	0.5932	0.5897	0.5248	0.1344	0.1204	0.1341	0.0922	0.0731	0.0937	659 849
EDSR	19.27	19.53	16.05	0.6190	0.6115	0.5258	0.1230	0.0982	0.1257	0.0845	0.0612	0.0853	684 961
DCASRN	19.41	19.61	16.12	0.6270	0.6112	0.5261	0.1180	0.0991	0.1282	0.0821	0.0611	0.0789	297 545
KD-DCASRN	**19.56**	**19.90**	**17.70**	**0.6358**	**0.6256**	**0.5924**	**0.0985**	**0.0973**	**0.1135**	**0.0781**	**0.0598**	**0.0738**	297 545

Table 5
Quantitative comparison of different models for Ajmer.

Method	PSNR			SSIM			RMSE			MAE			Trainable params
	Apr. 17	Mar. 19	Apr. 21	Apr. 17	Mar. 19	Apr. 21	Apr. 17	Mar. 19	Apr. 21	Apr. 17	Mar. 19	Apr. 21	
Bicubic	15.50	15.70	15.90	0.5116	0.4909	0.6226	0.1268	0.1365	0.1152	0.1185	0.1291	0.1110	0
SRCNN	24.41	25.10	23.15	0.5938	0.6532	0.5638	0.0455	0.0463	0.0500	0.0269	0.0293	0.0300	103 731
ESRT	19.04	19.68	18.79	0.5521	0.5976	0.5654	0.0844	0.0853	0.0827	0.0602	0.0611	0.0590	659 849
EDSR	24.95	25.71	25.73	0.6273	0.6727	0.6956	0.0427	0.0431	0.0372	0.0256	0.0281	0.0223	684 961
DCASRN	25.44	25.75	25.80	0.6567	0.7016	0.7270	0.0404	0.0429	0.0369	0.0245	0.0296	0.0213	297 545
KD-DCASRN	**25.45**	**25.97**	**26.23**	**0.6636**	**0.7017**	**0.7441**	**0.0403**	**0.0419**	**0.0351**	**0.0238**	**0.0272**	**0.0195**	297 545

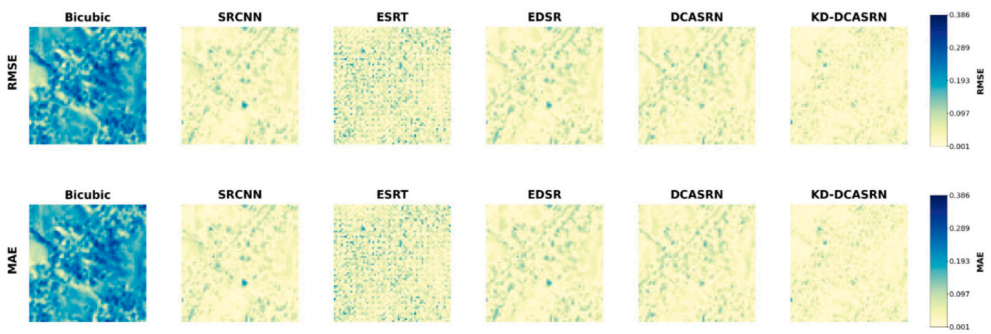


Fig. 17. Error map between reference image and super-resolution model for Majuli.

3.4.1. Quantitative results

Tables 3–5 present the quantitative results for the Palashbari and Majuli regions in Assam and the Ajmer region in Rajasthan, respectively, across multiple time points, enabling evaluation under diverse geographic and land-cover conditions.

Performance of KD-DCASRN: Among all tested models, our proposed knowledge-distilled model, KD-DCASRN, which uses DCASRN as the student network and EDSR as the teacher, consistently outperformed other models across all metrics and all three regions. This highlights the effectiveness of transferring rich feature representations from the deeper teacher model to the shallower student model, enabling the student to generalize better while remaining computationally efficient.

In the Palashbari region, KD-DCASRN achieved the highest PSNR values of 21.35 dB, 21.58 dB, and 20.46 dB across the three dates, significantly surpassing all other models, especially bicubic interpolation (12.92–13.29 dB). For SSIM, which captures structural fidelity, KD-DCASRN scored 0.6690, 0.7172, and 0.6121—again, the highest among all models. These improvements indicate superior visual quality and structural reconstruction.

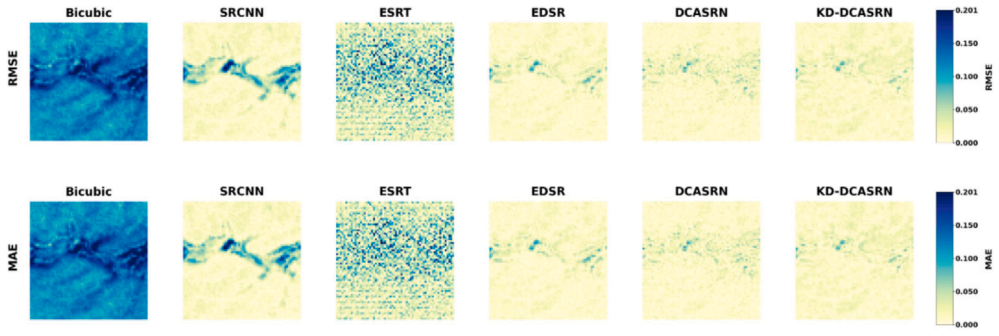


Fig. 18. Error map between reference image and super-resolution model for Ajmer.

Moreover, KD-DCASRN achieved the lowest RMSE values (0.0786, 0.0886, 0.0820) and MAE values (0.0522, 0.0522, 0.0489), reflecting substantial error reduction compared with other methods. While individual models like EDSR and DCASRN performed reasonably well, their performance lagged slightly behind KD-DCASRN. This confirms that the knowledge distillation mechanism enhanced the spatial and spectral information in the NDVI images.

Similar trends were observed in the Majuli region. KD-DCASRN again delivered the best PSNR (19.56 dB, 19.90 dB, 17.70 dB) and SSIM scores (0.6358, 0.6256, 0.5924). The RMSE and MAE values also remained the lowest among all models, showing that the improvements generalize well across regions with different topography and vegetation cover.

The inter-regional differences in Tables 3–5 further indicate that reconstruction accuracy was influenced by landscape characteristics. Across the three evaluated dates, KD-DCASRN achieved a higher mean PSNR in Ajmer (25.88 dB) than in Palashbari (21.13 dB) or Majuli (19.05 dB), with the same pattern reflected in lower RMSE and MAE values. This difference may reflect the distinct spatial structure of the Ajmer scene, where semi-arid conditions, open land, sparse vegetation, built-up areas, and water bodies likely produce smoother NDVI transitions than the floodplain, riverine, forested, and mixed agricultural landscapes represented in Palashbari and Majuli. In the Assam sites, dense vegetation, fragmented land-cover boundaries, narrow riverine features, and mixed land–water interfaces create more high-frequency spatial variation, making the 3× reconstruction task more challenging. Thus, the regional variation in PSNR, RMSE, and MAE suggests that model performance depends not only on network architecture and knowledge distillation but also on the spatial complexity of the target landscape.

The Majuli results also show date-specific variation. KD-DCASRN achieved lower PSNR and SSIM in Jan. 24 (17.70 dB and 0.5924) than in Mar. 22 (19.56 dB and 0.6358) and Mar. 23 (19.90 dB and 0.6256), while RMSE was also higher in Jan. 24 than in the two March cases. Because similar reductions are observed for the other deep learning models, this pattern likely reflects a more challenging seasonal or scene-specific reconstruction condition rather than a limitation of KD-DCASRN alone. In Majuli, seasonal changes in vegetation density, exposed soil, moisture conditions, riverine boundaries, and mixed land–water interfaces may alter NDVI texture and spectral distributions, making 3× reconstruction more difficult. These results suggest that future work should explicitly evaluate seasonal and phenological generalization using continuous time-series data rather than isolated acquisition dates.

The superior performance of KD-DCASRN can be attributed to synergy between the student and teacher models. While DCASRN is efficient in learning contextual features using channel attention, the teacher EDSR provides deep hierarchical representations. Through distillation, KD-DCASRN inherits DCASRN's channel attention and lightweight design and EDSR's rich feature knowledge. This leads to enhanced learning of spatial textures and finer NDVI details, surpassing the performance of SRCNN, ESRT, EDSR, and DCASRN without knowledge distillation.

In summary, our KD-DCASRN model surpasses traditional interpolation methods and outperforms highly ranked deep learning models like EDSR or ESRT.

3.4.2. Qualitative results

To visually assess NDVI, we used a custom colormap where blue shades indicate low values, yellow indicates intermediate values, and green indicates high values. This color progression clearly highlights variations in vegetation density and land cover.

As illustrated in Figs. 13–15, we selected representative subregions across varying vegetation biomes for Palashbari (Regions 1–4), Majuli (Regions 1 and 2), and Ajmer (Region 1). These regions include forested zones, water bodies, agricultural land, mixed-use areas, sparse vegetation, and open land, offering diverse terrain to evaluate qualitative performance.

Visual assessment reveals that our proposed KD-DCASRN model consistently produces NDVI maps that are visually closer to the reference image across all regions. The KD-DCASRN outputs capture subtle color gradients and fine vegetation structures — such as narrow river outlines, field boundaries, and mixed vegetation patches — that are either smoothed out or lost in other methods like bicubic interpolation, SRCNN, EDSR, and even nondistilled DCASRN.

Notably, in high-variation regions such as Palashbari–Region 3 and Majuli–Region 2, KD-DCASRN preserves the balance between yellow and green areas more effectively, closely matching the distribution observed for the reference image. This suggests that KD-DCASRN is better at retaining moderate vegetation zones and edge details between land-cover types.

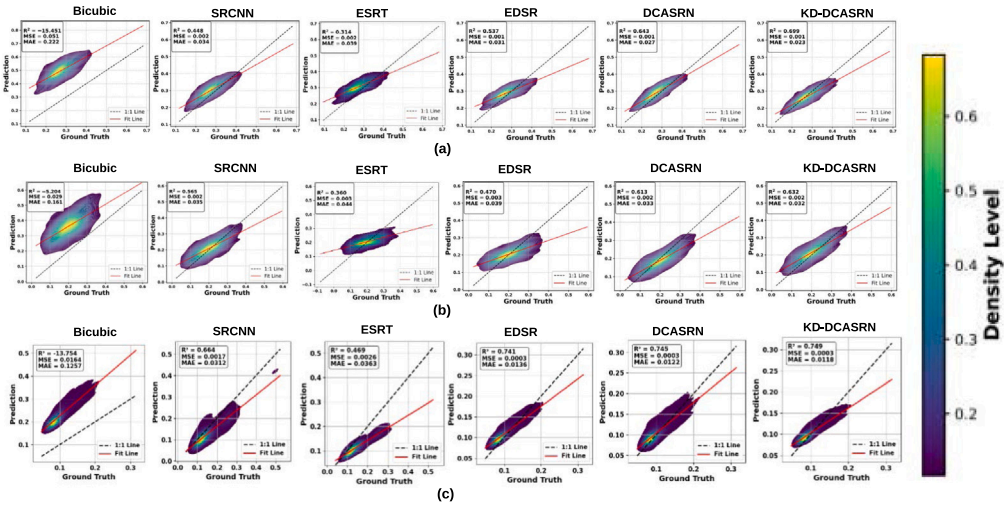


Fig. 19. (a) Density scatterplot for Palashbari, (b) Density scatterplot for Majuli, (c) Density scatterplot for Ajmer.

Table 6
Architectural and computational comparison of different super-resolution models.

Model	Conv. layers	Attention	Architecture / training strategy	Trainable params	GFLOPs[a]	Inference[b] (ms)
Bicubic	–	–	Interpolation	0	–	–
SRCNN	5	None	CNN-based SR	103731	3.394	15.389
ESRT	9	Efficient multi-head attention (EMHA)	CNN + transformer	659849	32.640	69.502
EDSR	68	None	Residual learning	684961	22.515	39.132
DCASRN	5	Channel attention (CA)	CNN + channel attention	297545	29.213	11.501
KD-DCASRN	5	Channel attention (CA)	CNN + CA + knowledge distillation	297545	29.213	11.508

[a] Computed for an input size of $128 \times 128 \times 1$. KD-DCASRN uses the same inference architecture as DCASRN and therefore has the same parameter count and computational complexity. Its performance improvement comes from knowledge distillation, using EDSR as the teacher model.

[b] Average forward-pass time for a single $128 \times 128 \times 1$ input patch, measured over 100 runs on the same hardware platform.

In contrast, outputs from low-resolution inputs and traditional upsampling methods over-smoothed colors, reduced detail, and overestimated vegetation (e.g., large dark-green patches). Even deep learning models like EDSR and DCASRN, used without knowledge distillation, show slight loss in intermediate vegetation gradients.

We used long-term, different time period satellite images to evaluate our model's capability for crucial tasks like vegetation monitoring. These qualitative results affirm the strength of our knowledge distillation approach, where the KD-DCASRN model benefits from the rich feature learning of the teacher (EDSR) while maintaining the student's attention mechanism and reduced parameter count (DCASRN). The outcome is sharper, structurally faithful NDVI reconstructions suitable for vegetation monitoring.

Bicubic resampling, being a simple interpolation method, does not enhance spatial details and results in uniform NDVI patches, while our model restores structural variations and finer textures. This qualitative improvement aligns with our quantitative results, highlighting the superior capability of our super-resolution model in reconstructing more detailed, realistic representations of land surface regions.

3.5. Ablation study

To validate the effectiveness of KD-DCASRN, we conducted an ablation study using density scatterplots and spatial error maps for Palashbari, Majuli and Ajmer. These tools enable detailed analysis of prediction quality and spatial consistency, highlighting performance differences among models. In addition, Table 6 summarizes the architectural characteristics and computational complexity of the evaluated methods, including network design, attention mechanisms, training strategy, trainable parameters, computational cost (GFLOPs), and inference time (ms). This comparison provides further context for interpreting the trade-off between reconstruction performance and model efficiency.

Palashbari region: The density scatterplots in Fig. 19(a) show clear performance differences. Bicubic interpolation performed poorly ($R^2 = -15.451$, MAE = 0.222) due to scale distortion, while SRCNN and ESRT provided moderate improvements, with R^2

values of 0.448 and 0.374, respectively. EDSR further improved reconstruction accuracy (R^2 = 0.537, MAE = 0.031). DCASRN further reduced errors (MSE = 0.001, MAE = 0.027). KD-DCASRN yielded the best overall performance, with R^2 = 0.699, MSE = 0.001, and MAE = 0.023. The error maps reinforce this result, showing that KD-DCASRN minimized reconstruction errors in dense, sparse, and mixed vegetation zones.

Majuli region: Similar results were obtained for the Majuli region, as shown in Fig. 19(b). Bicubic interpolation yielded the weakest results, with R^2 = −5.204, MSE = 0.029, and MAE = 0.161, showing a wide prediction spread and poor correlation with the reference image. ESRT also yielded poor results, with R^2 = 0.360, MSE = 0.003, and MAE = 0.044. SRCNN and EDSR showed improvement (R^2 = 0.565 and 0.470, respectively) but still struggled to align with the 1:1 reference line. DCASRN provided more accurate predictions, with R^2 = 0.613, while KD-DCASRN achieved the best correlation, with R^2 = 0.632, MSE = 0.002, and MAE = 0.032.

Ajmer region: The density scatterplots in Fig. 19(c) demonstrate consistent performance trends for the Ajmer region. Bicubic interpolation exhibited the weakest agreement with the reference NDVI data (R^2 = 0.177, MAE = 0.1197), reflecting its inability to recover fine-scale spatial variations. SRCNN and ESRT improved reconstruction quality, achieving R^2 values of 0.654 and 0.496, respectively. EDSR and DCASRN further enhanced prediction accuracy, with R^2 values of 0.741 and 0.745, respectively, and lower reconstruction errors. KD-DCASRN achieved the best overall performance, yielding the highest coefficient of determination (R^2 = 0.749) and the lowest MAE (0.0140), while maintaining a low MSE of 0.0003. These results indicate that the framework effectively reconstructed high-resolution NDVI patterns in the heterogeneous semi-arid landscape of Ajmer, including urban areas, sparse vegetation, open land, and water bodies.

For the spatial error maps, plotted using a yellow–green–blue gradient, yellow indicates lower local error magnitudes, while blue indicates higher errors. As shown in Figs. 16–18, KD-DCASRN reconstructions show fewer blue regions than those produced by SRCNN, ESRT, EDSR, and DCASRN. In contrast, bicubic interpolation shows the most blue regions.

This ablation study confirms the benefit of using a knowledge distillation strategy. Although DCASRN already performs well because of its channel attention mechanism, guidance from the teacher model, EDSR, enables KD-DCASRN to learn high-frequency spatial details more effectively and align more closely with the reference image. The consistent improvements in pixel-level scatterplots and spatial error distributions support the robustness and adaptability of KD-DCASRN across diverse remote sensing scenarios.

3.6. Uncertainty associated with cross-sensor reference data

Landsat and Sentinel-2 acquisitions used in this study were temporally matched within a window of ∼5 days to minimize phenological discrepancies between the two sensors. Despite this matching, residual differences may arise from the distinct spectral response functions of the Operational Land Imager (OLI) and MultiSpectral Instrument (MSI), as well as from differences in the atmospheric correction algorithms applied to each product. These factors can introduce small, systematic biases in NDVI values that are independent of the super-resolution model itself. Consequently, the PSNR, SSIM, RMSE, and MAE values reported here should be interpreted as a measure of agreement with the Sentinel-2 reference product rather than as an absolute measure of accuracy against a truly error-free ground reference. Future work could explore harmonization techniques (e.g., cross-sensor bias correction) to further isolate model performance from reference-data uncertainty.

4. Conclusion

In this work, we proposed a super-resolution approach that combines Landsat 8 and Sentinel-2 imagery to enhance NDVI reconstruction by leveraging their complementary strengths. The framework uses a knowledge distillation strategy in which a compact student network with channel attention (DCASRN) is guided by a deeper teacher network (EDSR), enabling the student to learn refined structural and contextual representations.

By combining Landsat's historically deep but spatially coarse 30 m data with Sentinel-2's spatially fine 10 m data, the framework reconstructs high-resolution NDVI images with improved spatial detail. This integration addresses a key challenge in remote sensing: enhancing the spatial resolution of long-term vegetation records without sacrificing computational efficiency.

Quantitative results show that KD-DCASRN outperforms bicubic interpolation, SRCNN, nondistilled DCASRN, and its teacher model, EDSR, across PSNR, SSIM, RMSE, and MAE metrics over multiple test cases. Qualitative evaluations across diverse land-cover conditions in the Palashbari, Majuli, and Ajmer regions show that KD-DCASRN reconstructs coherent NDVI patterns that more closely match the reference images, preserving subtle structures such as rivers, mixed vegetation zones, and fragmented landscapes.

While the proposed KD-DCASRN framework demonstrated consistent and promising performance across the three study sites evaluated in this work — Palashbari and Majuli in Assam (humid subtropical) and Ajmer, Rajasthan (semi-arid) — these results do not yet constitute global validation. Although the selected study areas represent geographically distinct environments, they cover only a limited range of climatic conditions and land-cover types. Future studies should evaluate the proposed framework across a broader range of eco-regions, including arid deserts, boreal forests, alpine ecosystems, and tropical rainforests, to further assess its robustness and generalizability under diverse environmental conditions.

CRediT authorship contribution statement

Allen Patnaik: Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Bhavesh Joshi:** Writing – original draft, Methodology, Conceptualization. **Dipima Sarma:** Validation, Resources, Investigation, Data curation. **M.K. Bhuyan:** Validation, Supervision, Investigation. **Arup Kumar Sarma:** Supervision, Resources, Project administration, Investigation, Data curation.

Ethical statement for solid state ionics

Authors here by assure that the following is fulfilled:

(1) This material is the authors' own original work, which has not been previously published elsewhere.

(2) The paper is not currently being considered for publication elsewhere.

(3) The paper reflects the authors' own research and analysis in a truthful and complete manner.

(4) The paper properly credits the meaningful contributions of co-authors and co-researchers.

(5) The results are appropriately placed in the context of prior and existing research.

(6) All sources used are properly disclosed (correct citation). Literally copying of text must be indicated as such by using quotation marks and giving proper reference.

(7) All authors have been personally and actively involved in substantial work leading to the paper, and will take public responsibility for its content.

Funding

The authors received no specific funding for this work.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Code and data availability

The Landsat 8 and Sentinel-2 imagery used in this study are publicly available from USGS EarthExplorer and the Copernicus Data Space Ecosystem, respectively. The processed NDVI datasets, sample training pairs, and source code for the proposed KD-DCASRN framework will be made publicly available at https://github.com/allenptnk/KD-DCASRN.

References

Al-Shammari, D., Whelan, B.M., Wang, C., Bramley, R.G.V., Fajardo, M., Bishop, T.F.A., 2021. Impact of spatial resolution on the quality of crop yield predictions for site-specific crop management. Agricult. Forest. Meterol. 310, 108622. http://dx.doi.org/10.1016/j.agrformet.2021.108622.

Brandt, M., Chave, J., Li, S., Fensholt, R., Ciais, P., Wigneron, J.-P., Gieseke, F., Saatchi, S., Tucker, C.J., Igel, C., 2025. High-resolution sensors and deep learning models for tree resource monitoring. Nat. Rev. Electr. Eng. 2 (1), 13–26. http://dx.doi.org/10.1038/s44287-024-00116-8.

Crivellari, A., Wei, H., Wei, C., Shi, Y., 2023. Super-resolution GANs for upscaling unplanned urban settlements from remote sensing satellite imagery–the case of Chinese urban village detection. Int. J. Digit. Earth 16 (1), 2623–2643. http://dx.doi.org/10.1080/17538947.2023.2230956.

Cui, J., Liu, J., Ni, Y., Sun, Y., Guo, M., 2025. MCKTNet: Multi-scale cross-modal knowledge transfer network for semantic segmentation of remote sensing images. IEEE Trans. Geosci. Remote Sens. 63, 4406015. http://dx.doi.org/10.1109/TGRS.2025.3547442.

Dong, C., Loy, C.C., He, K., Tang, X., 2014. Learning a deep convolutional network for image super-resolution. In: Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6–12, 2014, Proceedings, Part IV 13, vol. 8692, Springer, pp. 184–199. http://dx.doi.org/10.1007/978-3-319-10593-2_13.

Esmaeilzehi, A., Zaredar, H., Ahmad, M.O., 2025. DCSR: A deep continual learning-based scheme for image super resolution using knowledge distillation. Appl. Intell. 55 (7), 1–19. http://dx.doi.org/10.1007/s10489-025-06490-6.

Gong, M., Zhang, H., Xu, H., Tian, X., Ma, J., 2023. Multipatch progressive pansharpening with knowledge distillation. IEEE Trans. Geosci. Remote Sens. 61, 1–15. http://dx.doi.org/10.1109/TGRS.2023.3254053.

Goodarzi, R., Sahebi, M.R., 2025. Integrating Landsat-8 imagery and spectroscopic data for accurate soil lead content estimation using the LeadNet model based on transfer learning. Remote. Sens. Appl.: Soc. Environ. 38, 101546. http://dx.doi.org/10.1016/j.rsase.2025.101546.

Han, Y., Huang, J., Ling, F., Gao, X., Cai, W., Chi, H., 2024. RDNRnet: A reconstruction solution of NDVI based on SAR and optical images by residual-in-residual dense blocks. IEEE Trans. Geosci. Remote Sens. 62, 1–14. http://dx.doi.org/10.1109/TGRS.2024.3354255.

Henn, K.A., Peduzzi, A., 2024. Surface heat monitoring with high-resolution UAV thermal imaging: Assessing accuracy and applications in urban environments. Remote. Sens. (ISSN: 2072-4292) 16 (5), 930. http://dx.doi.org/10.3390/rs16050930.

Karmakar, P., Teng, S.W., Murshed, M., Pang, S., Li, Y., Lin, H., 2024. Crop monitoring by multimodal remote sensing: A review. Remote. Sens. Appl.: Soc. Environ. 33, 101093. http://dx.doi.org/10.1016/j.rsase.2023.101093.

Keys, R., 1981. Cubic convolution interpolation for digital image processing. IEEE Trans. Acoust. Speech Signal Process. 29 (6), 1153–1160. http://dx.doi.org/10.1109/TASSP.1981.1163711.

Khankeshizadeh, E., Mohammadzadeh, A., Mohsenifar, A., Moghimi, A., Pirasteh, S., Feng, S., Hu, K., Li, J., 2024. Building detection in VHR remote sensing images using a novel dual attention residual-based U-Net (DAttResU-Net): An application to generating building change maps. Remote. Sens. Appl.: Soc. Environ. 36, 101336. http://dx.doi.org/10.1016/j.rsase.2024.101336.

Liang, J., Cao, J., Sun, G., Zhang, K., Van Gool, L., Timofte, R., 2021. SwinIR: Image restoration using swin transformer. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 1833–1844.

Lim, B., Son, S., Kim, H., Nah, S., Lee, K.M., 2017. Enhanced deep residual networks for single image super-resolution. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. pp. 136–144. http://dx.doi.org/10.1109/CVPRW.2017.151.

Liu, Z., Wang, S., Li, Y., Gu, Y., Yu, Q., 2024. DSRKD: Joint despecking and super-resolution of SAR images via knowledge distillation. IEEE Trans. Geosci. Remote Sens. 62, 5218613. http://dx.doi.org/10.1109/TGRS.2024.3432193.

Liu, D., Zhong, L., Wu, H., Li, S., Li, Y., 2025. Remote sensing image super-resolution reconstruction by fusing multi-scale receptive fields and hybrid transformer. Sci. Rep. 15 (1), 2140. http://dx.doi.org/10.1038/s41598-025-86446-5.

Lu, Z., Li, J., Liu, H., Huang, C., Zhang, L., Zeng, T., 2022. Transformer for single image super-resolution. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 457–466.

Matthews, E., Prate, N., 2024. Nearest neighbor classification for classical image upsampling. arXiv preprint arXiv:2403.19611.

Palan, V.A., Thakur, S., Sumith, N., 2025. Leveraging super-resolution technology in drone imagery for advanced plant disease diagnosis and prognosis. IEEE Access 13, 66432–66444. http://dx.doi.org/10.1109/ACCESS.2025.3550257.

Patnaik, A., Bhuyan, M.K., MacDorman, K.F., 2024. A two-branch multi-scale residual attention network for single image super-resolution in remote sensing imagery. IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens. 17, 6003–6013. http://dx.doi.org/10.1109/JSTARS.2024.3371710.

Pham, V.-D., Bui, Q.-T., 2021. Spatial resolution enhancement method for Landsat imagery using a generative adversarial network. Remote. Sens. Lett. 12 (7), 654–665. http://dx.doi.org/10.1080/2150704X.2021.1918789.

Pouliot, D., Latifovic, R., Pasher, J., Duffe, J., 2018. Landsat super-resolution enhancement using convolution neural networks and Sentinel-2 for training. Remote. Sens. 10 (3), 394. http://dx.doi.org/10.3390/rs10030394.

Rohith, G., Kumar, L.S., 2021. Effectiveness of super-resolution technique on vegetation indices. IEEE Access 9, 97197–97227. http://dx.doi.org/10.1109/ACCESS.2021.3094283.

Segarra, J., 2024. Satellite imagery in precision agriculture. In: Digital Agriculture: A Solution for Sustainable Food and Nutritional Security. Springer, pp. 325–340. http://dx.doi.org/10.1007/978-3-031-43548-5_10.

Sun, M., Zhao, X., Yang, H., Rahmati, M., Chen, S., Shi, W., Zhao, S., 2024a. Enhancing long-term vegetation monitoring with high-performance super-resolution network. http://dx.doi.org/10.2139/ssrn.4762403, Available At SSRN 4762403.

Sun, M., Zhao, X., Zhao, J., Liu, N., Zhao, S., Guo, Y., Shi, W., Si, L., 2024b. A new spatial downscaling method for long-term AVHRR NDVI by multi-scale residual convolutional neural network. IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens. 17, 7068–7088. http://dx.doi.org/10.1109/JSTARS.2024.3373884.

Tahermanesh, S., Mohammadzadeh, A., Mohsenifar, A., Moghimi, A., 2025. SISCNet: A novel Siamese inception-based network with spatial and channel attention for flood detection in Sentinel-1 imagery. Remote. Sens. Appl.: Soc. Environ. 38, 101571. http://dx.doi.org/10.1016/j.rsase.2025.101571.

Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Qiao, Y., Loy, C.C., 2018. Esrgan: Enhanced super-resolution generative adversarial networks. In: Proceedings of the European Conference on Computer Vision (ECCV) Workshops.

Wang, C., Zhang, X., Yang, W., Wang, G., Zhao, Z., Liu, X., Lu, B., 2023. Landsat-8 to Sentinel-2 satellite imagery super-resolution-based multiscale dilated transformer generative adversarial networks. Remote. Sens. 15 (22), 5272. http://dx.doi.org/10.3390/rs15225272.

Wang, X., Zhong, B., Ao, K., Du, B., Hu, L., Cai, H., Qiao, Y., Wu, J., Yang, A., Wu, S., et al., 2024. Reconstruction of 30 m land cover in the Qilian Mountains from 1980 to 1990 based on super-resolution generative adversarial networks. Remote. Sens. 16 (22), 4252. http://dx.doi.org/10.3390/rs16224252.

Wu, X., Walker, J.P., Ye, N., 2024. Evaluation of the Bayesian downscaling algorithm for achieving higher resolution soil moisture data. IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens. 17, 5332–5344. http://dx.doi.org/10.1109/JSTARS.2024.3366886.

Xiong, C., Ma, H., Liang, S., He, T., Zhang, Y., Zhang, G., Xu, J., 2023. Improved global 250 m 8-day NDVI and EVI products from 2000–2021 using the LSTM model. Sci. Data 10 (1), 800. http://dx.doi.org/10.1038/s41597-023-02695-x.

Zhang, C., Marzougui, A., Sankaran, S., 2020. High-resolution satellite imagery applications in crop phenotyping: An overview. Comput. Electron. Agric. (ISSN: 0168-1699) 175, 105584. http://dx.doi.org/10.1016/j.compag.2020.105584.

Zhao, Y., Hou, P., Jiang, J., Zhao, J., Chen, Y., Zhai, J., 2023. High-spatial-resolution NDVI reconstruction with GA-ANN. Sensors 23 (4), 2040. http://dx.doi.org/10.3390/s23042040.

Zhao, Q., Ji, L., Su, Y., Zhao, Y., Shi, J., 2023. SRSF-GAN: A super-resolution-based spatial fusion with GAN for satellite images with different spatial and temporal resolutions. IEEE Trans. Geosci. Remote Sens. 61, 1–19. http://dx.doi.org/10.1109/TGRS.2023.3329115.