CC BY

SCUPNET: Classification of Temple Sculptures using deep learning approach

Pushpa T S [1], Vijayakumar Kadappa [2]Arjun Srinivas R[3], Shreesha B R[4], Rajeshwari Devi Doddapoojari Veerabhaskar[*5]

[1-4]Department of Computer Applications, B. M. S. College of Engineering, Bengaluru, Karnataka 560019, India
[5*]Department of Electronics and Communication Engineering, Global Academy of Technology, Bengaluru, Karnataka 560098, India
[1]pushpa.mca@bmsce.ac.in
[2]vijaykumark.mca@bmsce.ac.in
[3]arjun.mca23@bmsce.ac.in
[4]shreesha.mca23@bmsce.ac.in
[5]rajeshwaridevidv@gmail.com

Abstract:
Sculptures constitute a vital component of a nation's artistic and cultural heritage, reflecting intricate craftsmanship, symbolic meaning, and distinct regional styles across historical periods. Recent advances in Deep Learning (DL) have shown strong potential in digital heritage preservation, cultural analytics, and automated museum curation. However, temple sculpture classification remains challenging due to fragmented or eroded artifacts, missing inscriptions, and overlapping stylistic features across regions and dynasties. Manual classification is time-consuming, subjective, and reliant on domain expertise.
To address these challenges, this study proposes a Deep Learning-based approach for automated sculpture classification. A curated dataset of over 1,600 high-resolution images was compiled from the IHIRD repository and Google Images. Preprocessing techniques, including resizing, normalization, and data augmentation, were applied to enhance robustness and generalization.
To effectively capture hierarchical visual features, a customized deep neural network, termed SCULPture classification network (SCULPNET), was developed. Experimental results show that SCULPNET achieves an accuracy of 86.25% and an AUROC of 0.97 in classifying Bengal Terracotta, Dravida, Hoysala, and Vijayanagara sculptures. The model outperforms existing approaches such as ANN, MobileNetV2, and ResNet50. These findings highlight the potential of deep learning in supporting heritage preservation, digital archiving, and AI-assisted cultural research.

Keywords: Archaeology, Sculpture Classification; Deep Learning (DL); Artificial Intelligence (AI).

1 Introduction

In the digital era, the preservation of cultural heritage has become increasingly important, as digital technologies enable scholars and the public to study artifacts without the risks associated with physical degradation. Indian sculptures constitute a vital component of the nation's rich artistic and cultural legacy, reflecting intricate craftsmanship and distinct regional styles across various historical periods. Traditionally, sculptures in India are classified based on ruling dynasties (e.g., Maurya, Gupta, Chola), regions (e.g., Gandhara, Mathura), materials (stone, bronze, terracotta), and religious themes (Hindu, Buddhist, Jain). However, factors such as fragmented or eroded structures, missing inscriptions, and overlapping artistic influences across dynasties make accurate classification highly challenging. Furthermore, illegal excavations and inadequate digital documentation further complicate systematic preservation and categorization efforts.
Existing approaches lack intelligent automated tools capable of distinguishing traditional Indian architectural and sculptural styles such as Bengal Terracotta, Dravida, Hoysala, and Vijayanagara. At present, identification and classification largely depend on manual analysis by domain experts, a process that is time-consuming, subjective, and occasionally inconsistent. Moreover, non-experts, students, and tourists often face difficulties in recognizing and understanding different Indian architectural styles due to limited accessible tools and structured information. Machine learning techniques, such as Decision Trees, Support Vector Machines (SVM), and K-Nearest Neighbours (KNN), have been applied to image recognition tasks. However, these methods rely heavily on handcrafted features such as color histograms, texture descriptors, and edge detection techniques. Such feature engineering requires substantial domain expertise and often fails to generalize effectively to complex and diverse datasets, such as sculptures representing multiple Indian artistic traditions. Pre-trained AI models such as VGG16, ResNet50, and MobileNetV2, originally trained on large-scale datasets like ImageNet, have been utilized for transfer learning in related classification tasks. Nevertheless, these models are not specifically optimized for domain-specific datasets like Indian sculptures and may not adequately capture subtle stylistic distinctions among classes.
Deep learning techniques offer significant potential for digital conservation, cultural analytics, and intelligent museum curation of heritage artifacts. In particular, Convolutional Neural Networks (CNNs) have demonstrated remarkable performance in image classification tasks, making them highly suitable for analyzing complex visual structures such as Indian architectural and sculptural monuments.

To address these limitations, this study proposes a customized deep learning architecture, SCULPNET, designed to classify temple sculptures into four representative categories: Bengal Terracotta, Dravida, Hoysala, and Vijayanagara. Limited research has focused specifically on the automated classification of Indian temple sculptures. The proposed SCULPNET architecture encompasses five convolutional layers, three pooling layers, and three fully connected layers, using ReLU activation functions and a Softmax output layer. Simulation results indicate that the proposed model outperforms state-of-the-art architectures such as MobileNet and ResNet50 in terms of classification performance.

2 Related Work

The categorization and identification of uncommon Indian Hindu deities -Lord Ganesha, Hanuman, Krishna, Shiva, and Vishnu – is carried out in a study. The model was trained on a manually curated dataset of 500 images, comprising 100 images per deity. A CNN with a sequential architecture and Softmax activation was employed, and performance was evaluated using a SVM. The proposed approach achieved a classification accuracy of 94% [1]. Another work proposes a method for detecting sculptures in images using CNNs, addressing challenges related to varying materials and viewing perspectives. Regions of Interest (ROI) were annotated in images containing full or partial sculptures. The Single Shot MultiBox Detector (SSD) network was utilized for training, achieving an accuracy of approximately 82% on test images [2]. A study introduces two techniques for classifying Indian monuments based on architectural styles: Radon Barcodes and CNNs. Due to the absence of standard datasets for Indian architecture, a new dataset was created. While Radon Barcodes offered faster processing and lower memory consumption, CNNs achieved superior classification accuracy (82%) compared to 76% for Radon Barcodes [3].
An investigation for classifying eight architectural styles—Buddhist, Indo-Islamic, Rajput, Dravidian, Hindu, Sikh, British, and Modern—using machine learning is presented. Histogram of Oriented Gradients (HOG) was used for feature extraction, and models such as Logistic Regression, Random Forest, XGBoost, DenseNet, and InceptionV3 were evaluated. DenseNet achieved the highest accuracy of 86%, demonstrating the capability of deep learning models in capturing complex architectural features [4]. Another research proposes a specialized Generative Adversarial Network (GAN) to generate hyper-realistic and culturally representative images of Indian architectural monuments. Recognizing the limitations of conventional GANs in capturing intricate details and cultural nuances, a customized model was developed to learn architectural features, historical context, and symbolism, contributing to digital preservation efforts [5]. One study addresses the identification of objects depicted in Indian temple sculptures, which hold significant cultural and historical value. Image classification techniques were applied to digital archives to facilitate iconographic studies and enhance understanding of Indian temple architecture and heritage art [6].
A two-level hierarchical classification of Indian historical monuments based on construction period and architectural design has been introduced using machine learning (ML) and deep learning (DL) approaches. Various feature extraction techniques, including SURF, SIFT, FAST, ORB, Hu moments, and Zernike moments, were utilized. Classification was performed using ML algorithms such as ID3 and SVM alongside CNN models. The results indicate that CNN models outperform traditional ML approaches in classification tasks [7]. Another study focuses on the identification and categorization of Buddha statues, particularly those from Borobudur, Indonesia, using visual characteristics. The approach incorporates deep neural networks for feature encoding followed by classification using Support Vector Machines (SVM), forming a two-stage detection and classification method [8].
Research on Indian sculpture recognition using images captured from multiple angles and orientations applied SIFT feature extraction with classifiers such as KNN, SVM, and ANN. CNN models were also evaluated, and Contrast Limited Adaptive Histogram Equalization (CLAHE) was used for image enhancement. The highest accuracy achieved was 70.66% using CNN with CLAHE on a dataset containing 15 sculpture classes [9]. Another work enhances architectural style classification by integrating CNN with a channel-spatial attention mechanism. The process involves preprocessing to identify significant architectural regions, followed by deep feature extraction and attention-based refinement. Classification is performed using a Softmax classifier on architectural datasets [10]. A study exploring AI applications in tourism focused on temple architecture in Tamil Nadu. Several pre-trained CNN models, including ResNet50, VGG16, InceptionV3, MobileNetV2, and EfficientNetB1, were evaluated for temple recognition and classification. The categorized temples include Brihadeeswara, Meenakshi, Padmanabhaswamy, Five Rathas, and Shore Temple [11]. Another research addresses the identification of Indian deities using deep learning and transfer learning approaches. A custom dataset of 2,200 images across 11 deity classes was developed. CNN-based transfer learning models, including MobileNet, VGG16, and Xception, achieved accuracies of 98.76%, 96%, and 94%, respectively [12]. Another study investigates AI-based classification of Indian monuments using several CNN architectures, including ResNet50, InceptionResNetV2, EfficientNetB1, EfficientNetB3, and MobileNetV2. The models were trained on a dataset of over 4,000 images across 24 monument types, achieving training accuracies ranging from 13% to over 98% depending on the architecture used [13].
The existing work focuses on classification of deities, religions, monuments, or specific temples in one region. They do not address the classification of Indian temples spread across multiple regions. In our study, we consider temples located in Karnataka, West Bengal, and Andhra Pradesh [14].

3 Methodology

The approach adopted in this study encompasses multiple stages, beginning with data acquisition and preprocessing, followed by model selection, training, testing, and performance optimization.
Data collection:

The initial phase of the research involves acquiring a diverse dataset of 1,600 images of Indian sculptures, categorized into four distinct classes. The images were collected from online repositories, including the IHIRD website [14] and Google Images. The dataset was carefully curated to include sculptures spanning different historical periods (324 BC to the early 16th century AD), ensuring balanced representation across all four classes. The final dataset comprises a combination of original and augmented images for each class. One of the categories included in the study is Bengal Terracotta architecture (Fig. 1), which developed in West Bengal and Bangladesh between the 16th and 19th centuries. Due to the scarcity of stone in the region, terracotta (baked clay) was extensively used, resulting in intricate carvings and narrative relief panels. Temples such as the Jor Bangla Temple are characterized by distinctive curved roofs and decorative panels depicting mythological epics and scenes from daily life.

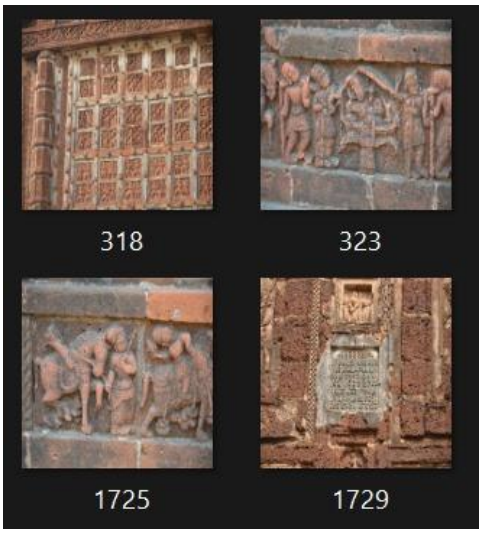


Figure 1 A sample of Bengal Terracotta architecture

Dravida architecture (Fig. 2) originated in Tamil Nadu during the early Chola period around the 7th century. It is characterized by pyramid-shaped vimanas (temple towers), monumental gopurams (gateway towers), and well-defined enclosed temple complexes. This architectural style emphasizes symmetry, vertical elevation, and elaborate sculptural detailing. Prominent examples include the Brihadeeswara Temple in Thanjavur and the Meenakshi Temple in Madurai.



Figure 2 A sample of Dravida architecture

Hoysala architecture (Fig. 3) flourished in Karnataka between the 11th and 14th centuries under the Hoysala dynasty. This style is distinguished by star-shaped temple platforms, the use of soapstone as the primary construction material, and highly intricate sculptural carvings. Notable examples include the Chennakesava Temple at Belur and the Hoysaleswara Temple at Halebidu.



Figure 3 A sample of Hoysala architecture

Vijayanagara architecture (Fig. 4) emerged in the regions of present-day Karnataka and Andhra Pradesh between the 14th and 17th centuries. This style represents a synthesis of Dravidian and Hoysala architectural traditions, characterized by imposing gopurams (gateway towers), expansive pillared halls (mandapas), and intricately designed stepped tanks (pushkarnis). Key sites include the Virupaksha Temple and the Vittala Temple at Hampi.

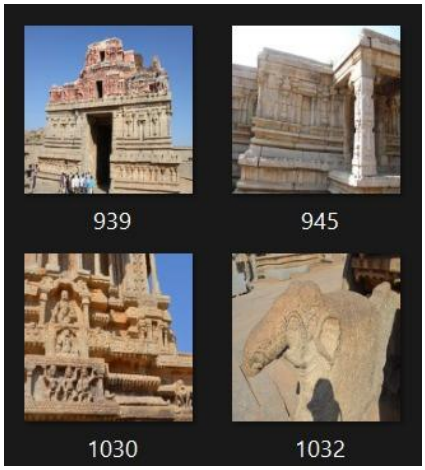


Figure 4 A sample of Vijayanagara architecture

Data Pre-processing

After collecting the dataset, all images are resized to a standard resolution to maintain uniformity in input dimensions. Data augmentation methods, such as rotation, flipping, zooming, and cropping, are employed to synthetically expand the dataset and improve the model's ability to generalize to unseen images. Additionally, background noise is minimized to emphasize the structural and architectural features of the sculptures.

Design and Training

A CNN architecture is designed specifically for Sculpture classification with convolution, pooling, and fully-connected layers. The dataset is split into 70% training, 15% validation, and 15% testing sets. The model is trained using the Adam optimizer and categorical cross-entropy loss function. Mini-batch gradient descent is applied for efficient weight updates. To further enhance generalization and prevent overfitting, regularization techniques such as early stopping are employed.

Testing

The performance metrics like accuracy, precision, recall, F1-score, AUROC are used to evaluate the models.

4 Proposed Sculpture-Convolutional Neural Network (SCULPNET)

The architecture of the proposed Convolutional Neural Network (CNN) is illustrated in Fig. 1. The network accepts an input image of size (224, 224, 3), representing RGB images. The first layer is a 2D convolutional layer comprising 96 filters of size (11 × 11), applied with a stride of (4 × 4) and ReLU activation. This configuration significantly reduces the spatial dimensions and produces a feature map of shape (54, 54, 96). A MaxPooling2D layer with a pool size of (3 × 3) and a stride of (2 × 2) follows, reducing the output to (26, 26, 96).

The second convolutional layer consists of 256 filters of size (5 × 5) with 'same' padding and ReLU activation, generating an output of shape (26, 26, 256). This is again followed by a max pooling layer with the same configuration, reducing the spatial dimensions to (12, 12, 256).

The third convolutional block contains three successive Conv2D layers. The first two layers employ 384 filters of size (3 × 3) with 'same' padding and ReLU activation, each producing feature maps of shape (12, 12, 384). The

third Conv2D layer uses 256 filters of size (3 × 3) with 'same' padding, resulting in an output of shape (12, 12, 256). A final max pooling layer with a pool size of (3 × 3) and a stride of (2 × 2) reduces the feature map to (5, 5, 256).

The resulting feature maps are flattened into a one-dimensional vector of size 6,400, which is then passed through two fully connected dense layers, each containing 4,096 neurons with ReLU activation, enabling high-level feature abstraction. The final output layer consists of four neurons, corresponding to the four sculpture classes, with a Softmax activation function for multi-class classification.

The network is trained using the Adam optimizer and categorical cross-entropy loss function, achieving strong classification performance. The detailed configuration of layer types, output shapes, and parameter values of the proposed deep network is presented in Table 1.

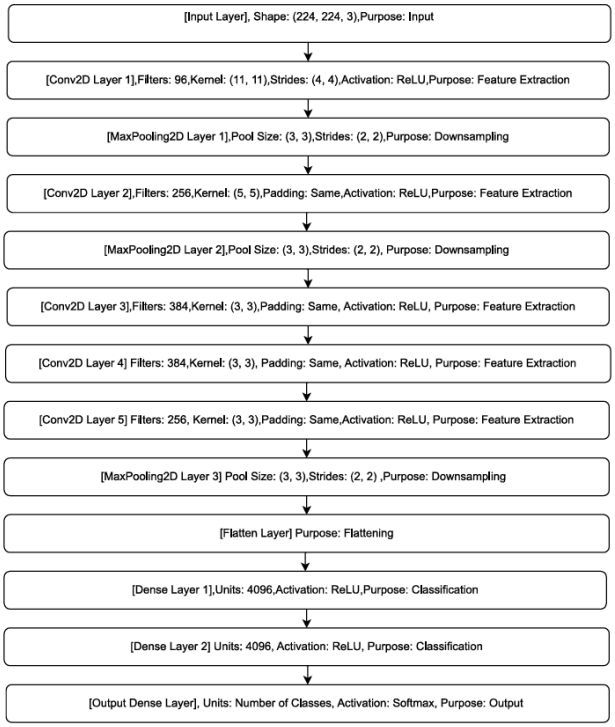


Figure 1 Architecture of SCULPNET

Table 1: Model layer type, output shape, and parameters

Layer Type	**Output Shape**	**Number of parameters**
Input Layer	(224, 224, 3)	0
Conv2D Layer 1	(54, 54, 96)	34,944
MaxPooling2D Layer 1	(26, 26, 96)	0
Conv2D Layer 2	(26, 26, 256)	6,14,656
MaxPooling2D Layer 2	(12, 12, 256)	0
Conv2D Layer 3	(12, 12, 384)	885,120
Conv2D Layer 4	(12, 12, 384)	1,327,488
Conv2D Layer 5	(12, 12, 256)	885,992
MaxPooling2D Layer 3	(5, 5, 256)	0
Flatten Layer	(6400,)	0
Dense Layer 1	(4096,)	26,218,496
Dense Layer 2	(4096,)	16,781,312
Output Layer (Dense)	(4,)	16,388

5 Experimental Results and Discussion

In this section, the proposed model is compared and evaluated against other deep learning models, namely ANN, MobileNet, and ResNet50. The performance of these models is assessed using evaluation metrics such as accuracy, precision, recall, F1-score, and AUROC. In addition, a simple graphical user interface (GUI) is presented for the classification of sculptures.

As described in Section 4, the study utilizes an Indian sculpture dataset comprising 1,600 images, with 400 images per class. The experiments are conducted over five iterations to ensure reliability and robustness of the results. In each iteration, the dataset is divided into 70% for training, 15% for validation, and 15% for testing.

Classification accuracy:

Table 2 presents the classification accuracy (in percentage) of four deep learning algorithms—SCULPNET, ANN, MobileNet, and ResNet50—across five iterations of training and testing. These accuracy values reflect the models' effectiveness in classifying images from the sculpture dataset. Accuracy is computed as the ratio of correctly classified images to the total number of test images under consideration.

The proposed SCULPNET achieves the highest average accuracy of 86.25%, whereas ResNet50 records the lowest performance among the compared models. SCULPNET outperforms ANN, MobileNet, and ResNet50 by margins of 17%, 8%, and 32%, respectively. Furthermore, SCULPNET consistently demonstrates superior accuracy across all five iterations compared to the other models.

Table 2. Classification Accuracy of Deep learning algorithms (%)

Iteration	SCULPNET	ANN	MobileNet	Resnet50
I	85.42	73.75	84.58	57.08
II	85	57.5	79.58	53.75
III	80.42	70.42	81.25	56.25
IV	89.58	74.58	79.58	52.92
V	90.83	68.75	72.5	50.42
Average	**86.25**	**69**	**78.29**	**54.08**

Precision, Recall, and F1-Score:

Precision is computed as the ratio of correctly predicted positive samples and the total predicted positive samples. A higher precision indicates a fewer false positives. The precision results of the proposed model and other models are presented in Table 3. It is clear that SCULPNET has achieved the highest precision of 0.86. This indicates that SCULPNET performs well in prediction by minimizing false positives. MobileNet is also competitive, whereas the other two models show poorer performance with a higher number of false positives.

Table 3. Precision values of Deep learning algorithms

Iteration	SCULPNET	ANN	MobileNet	Resnet50
I	0.86	0.75	0.84	0.58
II	0.85	0.71	0.8	0.54
III	0.82	0.72	0.82	0.59
IV	0.9	0.76	0.8	0.58
V	0.87	0.72	0.7	0.51
Average	**0.86**	**0.73**	**0.81**	**0.56**

Table 4 presents the recall values of the evaluated models. Recall is calculated using correctly predicted positives and the actual positive cases. It measures the model's ability to identify the positive instances. Among the evaluated models, SCULPNET achieved the highest recall of 0.85. This indicates that SCULPNET is more effective in reducing false negatives compared to ANN, ResNet50, and MobileNet.

Table 4. Recall values of Deep learning algorithms

Iteration	SCULPNET	ANN	MobileNet	Resnet50
I	0.85	0.74	0.85	0.57
II	0.85	0.57	0.8	0.54
III	0.8	0.7	0.81	0.56
IV	0.9	0.75	0.8	0.53
V	0.87	0.69	0.72	0.5
Average	**0.85**	**0.69**	**0.8**	**0.54**

To evaluate a model's performance in reducing both false positives and false negatives, the F1-score is used. It is computed as the harmonic mean of precision and recall. The F1-score values of SCULPNET and the other models are presented in Table 5. SCULPNET outperforms all other models, achieving a higher F1-score in each iteration as well as on average. The model attained the highest F1-score (reflecting the combined effect of precision and recall) of 0.857, whereas ANN, MobileNet, and ResNet50 achieved 0.709, 0.794, and 0.55, respectively. These results indicate that the proposed SCULPNET is more effective in minimizing both false positives and false negatives compared to ANN, MobileNet, and ResNet50.

Table 5. F1-Scores of Deep learning algorithms

Iteration	SCULPNET	ANN	MobileNet	Resnet50
I	0.855	0.745	0.845	0.575

II	0.850	0.631	0.800	0.540
III	0.810	0.710	0.815	0.574
IV	0.900	0.755	0.800	0.554
V	0.870	0.705	0.710	0.505
Average	**0.857**	**0.709**	**0.794**	**0.550**

Learning Curves:

Figure 2 illustrates the training and validation accuracies plotted as learning curves. These curves help determine whether the model is overfitting, underfitting, or achieving a good fit. As the number of epochs increases, the training and validation accuracies converge, indicating that the model achieves a good fit.

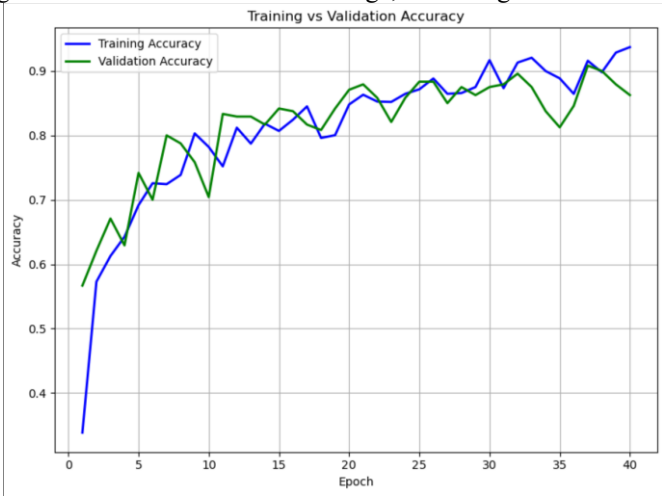


Figure 2 Training Accuracy vs Validation Accuracy for SCULPNET

Similarly, Figure 3 presents the training and validation loss curves, both of which show a decreasing trend as the number of epochs increases. This consistent reduction in loss further supports the model's stability during training. Overall, the learning curves confirm the good generalization ability of the proposed SCULPNET model.

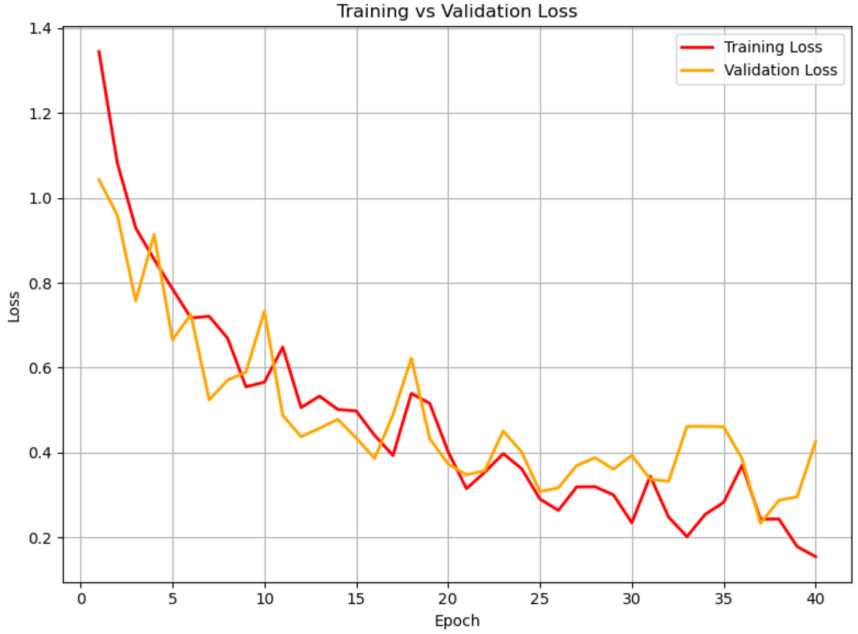


Figure 3 Training Loss vs Validation Loss for SCULPNET

Receiver operating characteristic (ROC) curve and Area Under ROC (AUROC):

Furthermore, the Receiver Operating Characteristic (ROC) curve of the proposed SCULPNET model is plotted to illustrate its effectiveness in multi-class classification, as shown in Figure 4. For strong generalization and class discrimination, the ROC curves are expected to be concentrated near the top-left corner of the plot. Since the curves for all classes are located close to the top-left corner, the proposed SCULPNET model demonstrates strong capability in discriminating among different sculpture architectures.

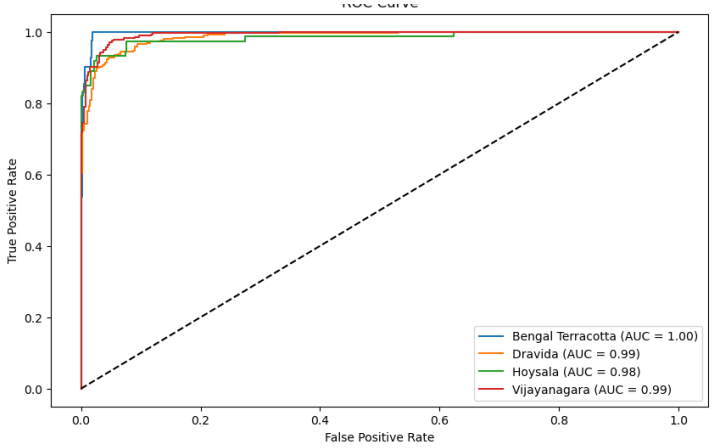


Figure 4 ROC Curve of SCULPNET

To quantify the discriminative ability of the models the Area Under the Receiver Operating Characteristic curve (AUROC) is computed. A well-performing model is expected to have an AUROC value closer to 1. The AUROC values of all models are presented in Table 6. It is evident that SCULPNET achieves the best classification performance with an average AUROC value of 0.97 among the compared models. MobileNet also performs well with an AUROC value of 0.912, whereas the other two models exhibit comparatively lower performance. It is noted that SCULPNET shows higher AUROC in each of the iterations as well.

Table 6. AUROC values of Deep learning algorithms

Iteration	SCULPNET	ANN	MobileNet	Resnet50
I	0.96	0.91	0.94	0.8
II	0.98	0.82	0.91	0.78
III	0.96	0.68	0.92	0.79
IV	0.98	0.76	0.91	0.77
V	0.97	0.83	0.88	0.76
Average	**0.97**	**0.8**	**0.912**	**0.78**

Sculpture classification System Interface:

The User Interface, based on the proposed model, is shown in Fig. 5, wherein a random image of a sculpture can be classified into one of the four architectures. It allows the user to select the image to be classified from the system folder using File manager (Fig. 6). Further, the system displays the details of the architecture with description for the selected sculpture image (Fig. 7).

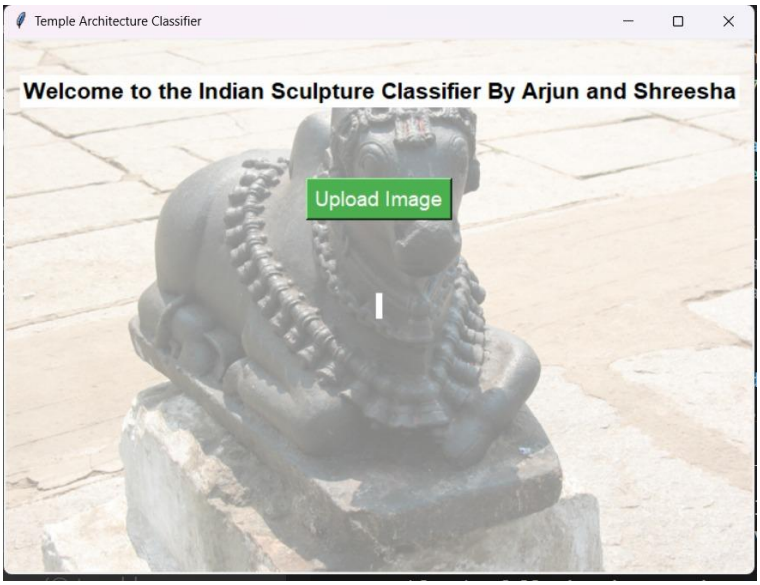


Figure 5 Simple UI to upload test images

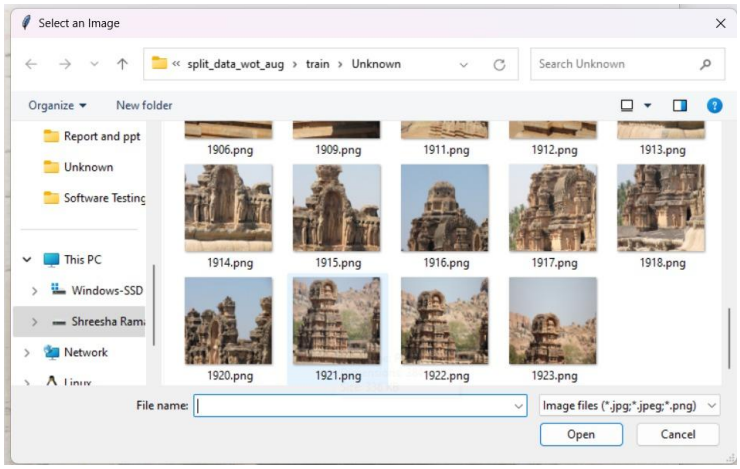


Figure 6 Upon clicking 'Upload image', File Manager is opened to select an image

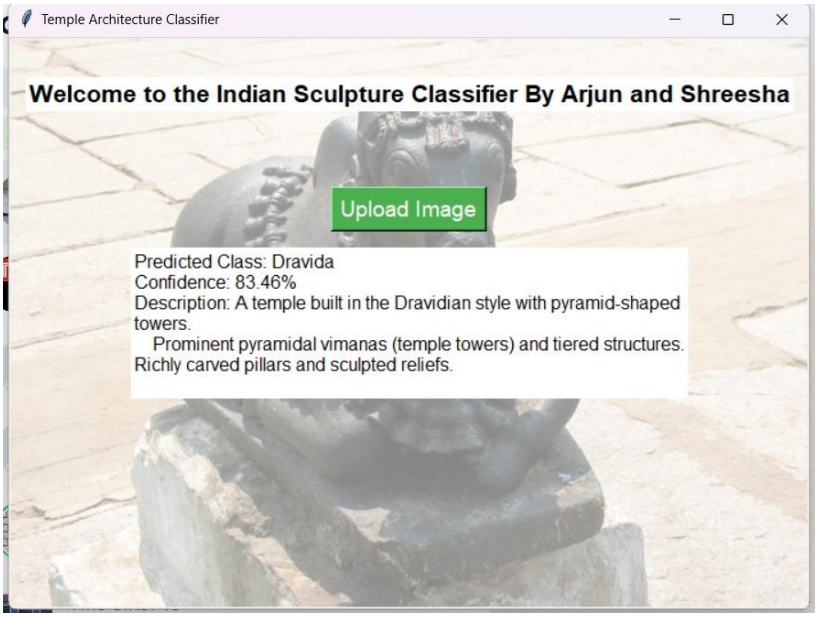


Figure 7 Image classification and the display of class details

6. Conclusion

The proposed SCULPNET deep network successfully demonstrates the application of deep learning for classifying sculptures into four distinct categories. By leveraging a custom-designed SCULPNET architecture, the model effectively extracts meaningful features from images, thereby enabling accurate classification. The proposed model achieves an average accuracy of 86.25%, which is higher compared to the other evaluated models. It also effectively reduces false positives and false negatives, attaining an F1-score of 0.857. Furthermore, it demonstrates superior class-discrimination capability with an AUROC value of 0.97.

This study explores the potential of deep learning in the domains of art and archaeology, offering a practical solution for automatic sculpture classification. The proposed approach contributes to the digital preservation and identification of Indian sculptures, facilitating quicker and more reliable classification for museum curators, researchers, and art collectors. Moreover, this initiative provides a foundation for further advancements in AI-powered cultural heritage categorization.

Author contributions statement

ASR and SBR formulated the problem, implemented the idea, obtained results, and drafted initial manuscript; TSP and VK involved in supervision, analysis of the results, reviewing and editing the manuscript, DVR reviewed the manuscript, and carried out literature review.

Additional information

The authors declare no competing interests.

Data Availability

Data is available as indicated in the manuscript [14].

References

1. Rao, A., C. Sindhu, A. Suhail, A. Mehta, and S. Dube. "Panchadeva: Sculpture image classification using CNN-SVM." Journal of Population Therapeutics & Clinical Pharmacology 30, no. 9 (2023): 332-344.

2. Hong, Dajeong, and Jongweon Kim. "Sculpture Detection Method using the Convolution Neural Network." In *Proceedings of the 2017 International Conference on Information Technology*, pp. 148-151. 2017.

3. Sharma, Saurabh, Priyal Aggarwal, Akanksha N. Bhattacharyya, and S. Indu. "Classification of Indian monuments into architectural styles." In *National Conference on Computer Vision, Pattern Recognition, Image Processing, and Graphics*, pp. 540-549. Singapore: Springer Singapore, 2017.

4. Ganapathi, Ananya, Deivanai Saravanan, and Sneha T. Raghavan. "Classification of Structures and Monuments based on Indian Architectural Styles." In *2025 3rd International Conference on Intelligent Systems, Advanced Computing and Communication (ISACC)*, pp. 1157-1163. IEEE, 2025.

5. Ishvarya, G., Kavya Sree Kammari, Sharanya Vanraj Thambi, and Manju Khanna. "ProGAN-Based Architectural Style Transfer: Cross-Cultural Adaptation of Indian Architecture." In *International Conference on Information and Communication Technology for Intelligent Systems*, pp. 497-509. Singapore: Springer Nature Singapore, 2024.

6. Mahesh, Vanishree Kavi, Lubna Ansari, and Aprna Tripathi. "AI-Based Image Classification for Indian Temple Iconography: A CNN-Based Approach." In *2025 International Conference on Artificial Intelligence and Data Engineering (AIDE)*, pp. 67-72. IEEE, 2025.

7. Kavitha, Srinivasan, S. Mohanavalli, B. Bharathi, C. H. Rahul, S. Shailesh, and K. Preethi. "Classification of Indian monument architecture styles using bi-level hybrid learning techniques." In *Inventive Systems and Control: Proceedings of ICISC 2022*, pp. 471-488. Singapore: Springer Nature Singapore, 2022.

8. Shivanee, Nikhil Kumar Rajput, and Ajay Jaiswal. "A Machine Learning Approach for the Classification of the Buddha Statues of Borobudur (Indonesia)." In *Data Analytics and Management: Proceedings of ICDAM*, pp. 891-900. Singapore: Springer Singapore, 2021.

9. Dalara, Ayush, C. Sindhu, and R. Vasanth. "Entity Recognition in Indian Sculpture using CLAHE and machine learning." In *2022 First international conference on electrical, electronics, information and communication technologies (ICEEICT)*, pp. 1-12. IEEE, 2022.

10. Wang, Bo, Sulan Zhang, Jifu Zhang, and Zhenjiao Cai. "Architectural style classification based on CNN and channel–spatial attention." *Signal, Image and Video Processing* 17, no. 1 (2023): 99-107.

11. Avudaiamal, R., Sam Varghese George, and Bhargavi YR. "Cost-Effective Multiclass Model: Unveiling the Legacy of Tamil Nadu's Iconic Temples." 2025 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE).

12. Karegowda, Asha Gowda, R. Pooja, K. N. Tara, and A. Leena Rani. "Identification of Deity Images Using CNN and Transfer Learning Models." In International Conference on Robotics, Control, Automation and Artificial Intelligence, pp. 153-167. Singapore: Springer Nature Singapore, 2023.

13. Khandelwal, Sarika, Arunesh Prasad, Anshuman Kumar, Jaanhavi Gautam, and Aishwarya Patle. "A Study on Efficient Image Classification of Historical Monuments Using CNN." In 2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA), pp. 220-225. IEEE, 2023.

14. Indian Heritage Image Retrieval Dataset (IHIRD). Visual Information Processing (VIP)Laboratory, Department of Computer Science & Engineering Indian Institute of Technology Kharagpur, West Bengal, India, https://facweb.iitkgp.ac.in/~jay/ihird/index.html.