International Journal of Basic and Applied Sciences, 14 (5) (2025) 840-852

International Journal of Basic and Applied Sciences

Website: www.sciencepubco.com/index.php/IJBAS
https://doi.org/10.14419/zc5xbc77
Research paper

Protecting Historical Treasures: Deep Learning for Structural Health Assessment

M.Snehapriya [1] *, A.Umamageswari [2]

[1] Research Scholar, Department of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India
[2] Associate Professor, Department of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India
**Corresponding Author email: snehaprm1@srmist.edu.in*

Received: July 27, 2025, Accepted: September 9, 2025, Published: September 25, 2025

Abstract

Preserving ancient monuments not only stands as evidence of the collective history but is also a fundamental aspect of conserving cultural heritage. As these architectural wonders age, they become susceptible to the encroachment of moss and the development of structural cracks, jeopardizing their overall integrity. In the present era of advanced technology, the use of deep learning algorithms for the automatic detection and classification of cracks, moss, and seepages arises as a promising solution. This study explores the deployment of deep learning techniques, particularly Single Shot Multibox Detector (SSD), for the identification and localization of moss and cracks on ancient monuments. The proposed work investigates the effectiveness of these algorithms in analyzing high-resolution images, offering a non-invasive and precise means of monitoring structural defects and biological growth with the help of the dataset Indian Ancient Monuments (Kaggle). To delve into the architectural significance of these structures, which underlines the importance of their preservation. The deep learning models, trained on the dataset, exhibit remarkable proficiency in distinguishing between normal surfaces and those bearing cracks, moss, and seepages. We assess the model's performance through metrics such as accuracy, precision, recall, and F1-score. The proposed work is compared with the state-of-the-art techniques like SVM, KNN, CNN, and Random Forest in the field. Furthermore, this study provides a glimpse into potential applications of the developed models, including real-time monitoring and alert systems for heritage conservationists and preservation authorities. By amalgamating advanced deep learning technology SSD with heritage preservation, this research underscores a proactive approach to safeguarding ancient monuments with the accuracy of detecting cracks, moss, and seepages as 96.9%, 97.5% and 97.5% respectively. The implications extend beyond the technical realm, transcending into cultural heritage and conservation practices, thereby ensuring the longevity of these historical treasures.

***Keywords**: Single Shot Multibox Detector (SSD), Cracks, Moss, Seepages*

1. Introduction

In a culturally diverse nation like India, monuments serve as essential windows into the past, offering profound insights into the customs, traditions, and societal norms that have evolved over centuries. These architectural marvels tell stories of ancient civilizations, dynasties, and empires that have left an indelible mark on the nation's collective identity. They stand as tangible evidence of the artistic and engineering prowess of their respective eras. These monuments are not just static structures; they are dynamic cultural ambassadors. They attract millions of tourists from around the world, fostering cross-cultural exchanges and promoting a deeper understanding of India's unique cultural heritage. The preservation and maintenance of these monuments are not just historical duties but are essential for the continuity of India's rich cultural narrative. India boasts a vast and diverse population, and its cultural heritage is a global treasure. The nation is home to an astounding 2.5 million historic sites, with the state of Tamil Nadu alone harboring an impressive 40,000 such sites. These relics from the past have withstood the test of time for centuries, but they are now facing the pressing issue of inadequate protection against environmental degradation.

These ancient artifacts are more than just remnants of a bygone era; they are repositories of invaluable knowledge encompassing the wealth, culture, and myriad facets of the nation's heritage. They hold within them the secrets of music, astronomy, medicine, dance, and much more. The architectural and construction practices of ancient India were so advanced that they gave rise to magnificent heritage structures. These architectural marvels stand as testaments to the genius of their creators and the rich history of the land.

However, the relentless changes in the environment over the centuries have taken a toll on these monuments and structures. They have weathered the forces of nature and the impact of time, which makes it all the more urgent to protect and preserve them for the benefit of current and future generations. These sites are a living link to India's illustrious past, and safeguarding them is not only a responsibility but a necessity to ensure that their invaluable insights into history and culture continue to shine brightly in the present and the future. While there exists a multitude of heritage structures scattered throughout the country, it is regrettable that only a limited of them are maintained by the Archaeological Society of India (ASI). The majority of these structures are closely tied to religious centers, where conservation efforts often lack a scientific approach [1]. To safeguard these cultural heritage treasures, it is imperative to introduce scientific interventions that can assess the types of weathering and extent of damage these structures endure. These assessments can serve as invaluable inputs for devising restoration and preservation plans. Such interventions must be non-destructive, conducted on-site, efficient in terms of time, and cost-effective, while providing graphical maps of damage locations and distributions [2]. Resorting to destructive techniques to evaluate a monument's state of decay is strongly discouraged, as it could inadvertently exacerbate the existing damage.

In this context, the proposed research offers a technological intervention to evaluate the deterioration of these monuments using non-destructive methods.

Monuments suffer structural damage from a range of weathering factors. There are three primary forms of weathering deterioration: chemical, physical, and biological. Chemical weathering leads to the replacement of molecular structures in rocks, causing the gradual collapse of the rocks' integrity. Biological weathering involves the deterioration of structures due to the actions of animals and plants. Moreover, the natural processes of weathering, which occur year-round and are influenced by the changing seasons, pose a constant threat to items exposed at heritage sites. The impact of rain and wind from the environment can result in peeling, exfoliation, and disintegration of monuments. Additionally, two significant contributors to the decay of monuments are cracks and the presence of moss [3]. Addressing these forms of deterioration is essential to preserving the cultural heritage.

The assessment of monument decay can be carried out through two distinct methods: Destructive and Non-destructive approaches [4-7]. Destructive techniques involve the use of diffraction analysis of X-rays and advanced microscopy techniques, such as Scanning Electron Microscopy and Transmitted Light Microscopy [8].

On the other hand, Non-destructive methods offer more gentle alternatives, which are crucial for preserving the integrity of these heritage structures. These non-destructive techniques include digital image processing, ultrasonic imaging, and infrared thermography [9].

Cultural heritage structures often succumb to deterioration primarily caused by moss growth and the formation of cracks. The task of assessing structural health and safety is particularly challenging, with the detection of cracks and seepage demanding careful attention [10]. Manual inspections for moss and cracks are laborious and can sometimes lead to arbitrary decisions on the part of investigators.

A promising development in this field is the creation of a hybrid crack detection model [11]. This innovative approach involves the measurement of crack properties such as area and direction, offering a more systematic and reliable means of assessing structural damage. Such advancements are pivotal in ensuring the preservation of the cultural heritage without causing further harm to these valuable monuments.

The manual identification of cracks and moss on heritage structures can be a challenging and time-consuming task. Gathering data and information on decay factors simultaneously presents significant difficulties. To address these challenges, deep learning techniques have emerged as a valuable tool, offering a more efficient and less labor-intensive approach.

In a recent development, a novel image-based machine-learning approach has been proposed for the classification of cracked and uncracked specimens. This approach offers a means to objectify and automate crack identification, effectively removing sources of uncertainty and inaccuracy that often result from manual experiment post-processing [12].

The objective of the study by Mehta and Shah is to enhance pipeline system engineering practices by predicting longitudinal dispersion coefficients in water pipelines using Extreme Learning Machine (ELM) concepts. Methodologically, the research employs ELM-based predictive modeling techniques on pertinent datasets related to water pipeline systems, integrating advanced data handling methodologies to ensure accurate predictions. The output of the study is precise estimations of longitudinal dispersion coefficients, providing valuable insights that can inform the optimization of water pipeline infrastructure design and maintenance strategies, ultimately contributing to the efficient and sustainable management of water resources [13]. This study aims to enhance prediction accuracy for scour depth around long contractions in hydraulic engineering contexts by employing Support Vector Machine (SVM) and Adaptive Neuro-Fuzzy Inference System (ANFIS) methodologies. It utilizes SVM and ANFIS techniques to analyze datasets related to scour depth around long contractions, integrating advanced computational approaches for precise prediction of scour depth in hydraulic engineering scenarios. The research outputs improved prediction accuracy for scour depth, thus contributing to a better understanding and management of scour phenomena in water infrastructure projects [14].

This study aims to enhance the accuracy of water quality parameter estimation through the introduction of a novel Multiple-Kernel Support Vector Regression (MK-SVR) algorithm. The research employs the MK-SVR methodology to analyze relevant datasets about water quality parameters, integrating advanced computational techniques for precise estimation. The study's output includes improved accuracy in estimating water quality parameters, offering valuable insights into the effectiveness of the MK-SVR algorithm in water resources management applications, thus contributing to a better understanding and management of water quality issues [15].

By leveraging these cutting-edge techniques, the proposed method can streamline the process of identifying and assessing decay factors, ultimately contributing to the preservation of the cultural heritage more accurately and efficiently. The primary objectives of deploying deep learning techniques, particularly Single Shot Multibox Detector (SSD), for the identification and localization of moss and cracks on ancient monuments are to automate the detection process, ensuring efficient monitoring of monument condition, accurately localize affected areas for targeted intervention, enable early detection to prevent further deterioration, facilitate continuous monitoring for proactive maintenance, and utilize data-driven insights to inform decision-making processes regarding monument preservation and management.

This paper is structured as follows: In Section 2, we delve into an exploration of the relevant works in this field. Section 3 provides an in-depth explanation of the proposed work, highlighting its contributions and innovations. Moving forward, Section 4 offers a comprehensive presentation of the results obtained and engages in discussions to provide a deeper understanding of the findings. Finally, in Section 5, we conclude and outline the potential directions for future research in this area.

2. Related Works

The Author introduces a novel DBST-LCM-CLAHE approach for image denoising, combining deep learning with traditional image processing techniques. This method effectively reduces noise and enhances image quality, demonstrating its versatility and potential for various applications. The results highlight the method's efficacy, making it a valuable contribution to image processing and denoising research. The method may encounter difficulties with computational complexity and limited evaluation across various datasets, which could impact its scalability and generalizability, even though it enhances noise reduction and detail preservation. despite these concerns, it exhibits great promise for uses like satellite and medical imaging that call for high-fidelity image restoration [16].

The Author introduces an advanced image denoising method using Convolutional Neural Networks (CNNs), taking advantage of CNNs' capacity to learn spatial features and adaptively eliminate noise. Their findings show that the model is successful in maintaining structural details while lowering noise, as evidenced by improvements in PSNR and SSIM over conventional filtering and transform-based techniques.

However, the CNN approach may have trouble with unseen noise types or varying intensities, and its performance is largely dependent on the quality and diversity of the training dataset. Furthermore, real-time deployment may be hampered by CNNs' computational requirements, and the study does not compare its results to those of more sophisticated techniques like GANs or autoencoders. The study highlights CNNs' potential for further improvements in image restoration tasks and establishes them as a solid baseline in deep learning-based image denoising, despite these drawbacks [17].

The Author presents a novel medical image enhancement algorithm that combines the Pelican Optimization Algorithm (POA) with Contrast Limited Adaptive Histogram Equalization (CLAHE), where POA adaptively adjusts enhancement parameters while CLAHE enhances local contrast. The outcomes demonstrate improved image quality when compared to traditional techniques, indicating great promise for medical imaging applications that demand precise visualization. However, the study primarily uses statistical and visual measures without substantial clinical validation, and optimization-based approaches are frequently computationally demanding, which may limit real-time deployment. Furthermore, consistency issues are brought up by relying solely on POA, since other optimization strategies may work better in certain circumstances. Despite these drawbacks, the study shows that combining bio-inspired optimization and adaptive contrast enhancement can improve the quality of diagnostic images [18]

The Author likely presents an enhanced U-Net image segmentation method for more accurate and efficient image segmentation. Its variants for medical image segmentation emphasize how well they like histopathology, CT, and MRI. The article classifies architectural enhancements such as residual connections and attention mechanisms, demonstrating how they solve issues with feature extraction and class imbalance. But as a review, it is still descriptive without empirical benchmarking, and it pays little attention to real-world problems like clinical deployment, data scarcity, and computational cost. despite these drawbacks, the study is an important resource that highlights U-Net's adaptability and long-standing significance in biomedical image analysis [19]. The Author proposed the approach, which involves the use of a fully convolutional encoder-decoder network, which is a deep learning technique commonly employed for image segmentation tasks. The author suggests advanced technology for automatic crack detection in concrete structures, which can be vital for maintenance and safety assessments in civil engineering and construction. Its emphasis on concrete structures, however, restricts its applicability to other materials, such as marble or stone, and it ignores environmental elements like moss or stains. Furthermore, the computational expense of the model might prevent widespread real-time implementation. [20].The Author likely introduces an image segmentation algorithm that is based on an improved multiscale random field model within the wavelet domain that aims to enhance image segmentation by incorporating a sophisticated multiscale model within the wavelet transform framework. This approach is expected to improve the accuracy and robustness of image segmentation. However, real-time applicability is restricted by its high computational cost, and scalability issues are raised by validation on small datasets. It is unclear if it can adapt to typical patterns in heritage sites, such as moss or cracks.[21].

The Author likely presents a novel approach to image segmentation by using fuzzy clustering techniques to extract transition regions. This approach aims to focus on fuzzy clustering, suggesting a method that can effectively handle uncertainty and ambiguity in image segmentation, making it a valuable contribution to the field of computer vision and image processing. However, weakens it in different textures or lighting conditions. Its scalability and adaptability are still constrained in comparison to deep learning [22]. The Author introduces a non-invasive method that provides an efficient way to monitor and evaluate the condition of historical structures, facilitating timely conservation efforts to significantly advance the field of monument preservation, emphasizing the role of image processing in safeguarding cultural heritage. The study has potential for non-invasive detection, but it is less robust and scalable than deep learning-based techniques due to its reliance on manually created features and thresholds, which restricts adaptability to changing circumstances. [23].

The Author presents a groundbreaking and non-invasive method for the accurate detection of moss and cracks in monuments using image processing techniques to preserve the cultural heritage by providing an efficient tool for assessing monument condition. This non-destructive approach holds promise for safeguarding historical landmarks and architectural structures. For large-scale or outdoor heritage monument applications, where non-invasive, image-only deep learning techniques might be more useful, the dependence on laser hardware limits scalability and raises costs and complexity.[24]. The Author introduces a novel approach that offers a non-destructive and accurate way to assess crack depths, improving safety and reliability in industries relying on steel materials. The research represents an innovative application that combines laser and image processing techniques to enhance structural health assessment [25].

The author aims to enhance the process of detecting and evaluating decay factors, thereby advancing the preservation efforts of our cultural heritage with heightened precision and effectiveness. The study concentrates on refining the precision and efficiency of leaf disease detection through the application of cutting-edge deep learning techniques [26-27]. The author presents a deep learning-based, non-invasive approach to monument decay analysis, showing how neural networks can more accurately identify and categorize deterioration patterns like moss and cracks than conventional image processing. Despite highlighting the potential of deep learning for heritage conservation, the study does not go into great detail about how to optimize hyperparameters, scale to different types of monuments, or deal with real-world deployment issues. Concerns regarding generalization across various materials and environmental conditions are also raised by the dependence on small datasets [28].

The assumption of homogeneity in decay patterns across different monuments is a limitation of the proposed work. It recognizes that variations in architectural styles, construction materials, and environmental exposures can lead to diverse decay patterns that may not be adequately addressed by a one-size-fits-all deep learning model. To address this limitation, dataset expansion with more augmentation in the dataset with a diverse range of architectural styles, construction materials, and environmental conditions, transfer learning to adapt pre-trained models to specific subsets of data corresponding to different types of monuments, and ensemble methods to combine predictions from multiple specialized models trained on different subsets of data. By aggregating diverse perspectives, it is expected to improve the overall robustness and generalization capability of the deep learning framework.

3. Method

Deep learning plays a crucial role in heritage preservation by automating defect detection, predicting future deterioration, and optimizing resource allocation. Its non-invasive approach, combined with data-driven decision-making, ensures efficient monitoring and proactive preservation efforts. Integrating deep learning into heritage conservation enables effective preservation of monuments, thereby extending their lifespan and cultural significance. Detecting cracks, moss, and seepages in images, especially in the context of ancient monuments or cultural heritage preservation, typically involves a combination of image processing techniques, computer vision, and machine learning. Figure 3.1 shows the architecture diagram for decay assessment in monuments.

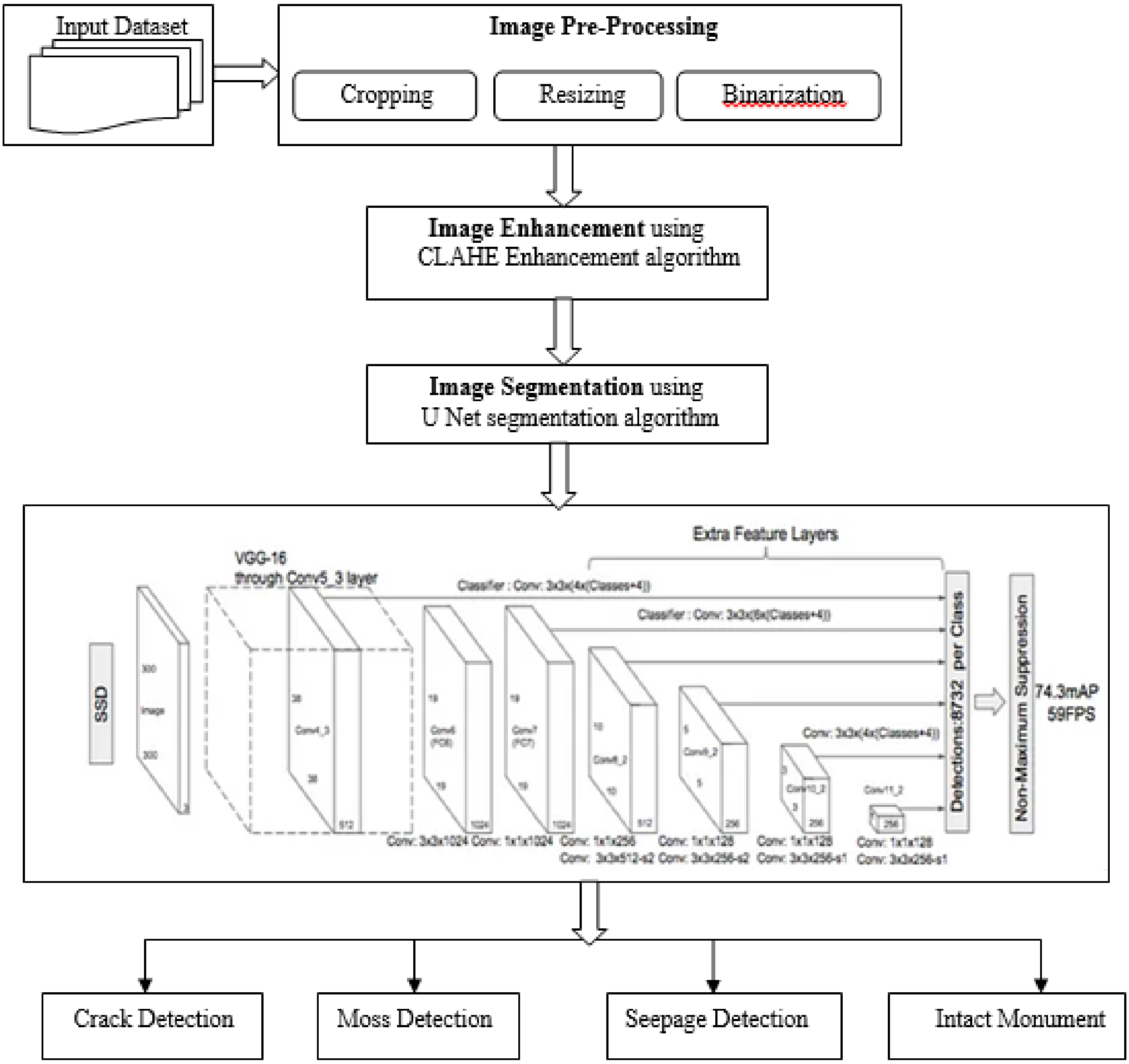


Fig. 1: Architecture Diagram

3.1 Data Preprocessing:

Image preprocessing is a critical step in decay analysis of ancient monuments. It enhances quality and prepares images for subsequent analysis, making it easier to detect and quantify decay. The following are some common preprocessing steps for decay analysis in ancient monuments.

3.1.1 Cropping:

Cropping is a common image preprocessing technique that involves removing unwanted portions of an image while retaining the region of interest (ROI). It is used to focus on a specific area of the image, eliminate distractions, and reduce computational overhead. Cropping of an image is illustrated in Fig. 2.

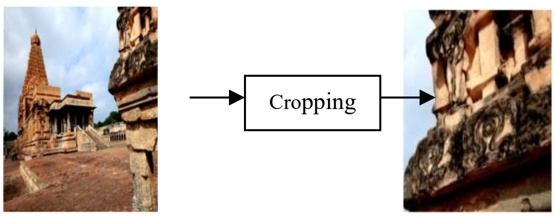


Fig 2. Cropping of an image

3.1.2 Resizing:

Resizing the image to a specific aspect ratio by cropping or adding padding. After resizing, I need to apply noise reduction filters to reduce image noise, which can improve the quality of decay detection.

3.1.3 Binarization:

Fig. 3 depicts the Binarization of an image, which is the process of converting a grayscale image into a binary image, where each pixel is classified as either "foreground" (usually represented as white) or "background" (usually represented as black) based on a specified threshold. In the resulting binary image, pixels are typically assigned one of two values, often 0 for background and 255 (or 1) for foreground, although the exact values can vary depending on the representation. The primary goal of binarization is to simplify and segment the image, making it easier to analyze or process for subsequent image processing tasks.

Fig 3. Binarization

3.1.4 Image Enhancement:

Image enhancement techniques are used to improve the visual quality of images by emphasizing certain features or reducing unwanted artifacts. The CLAHE (Contrast Limited Adaptive Histogram Equalization) algorithm is a sophisticated image enhancement technique designed to improve the visibility of fine details in images, particularly in regions with varying illumination or contrast. It operates by dividing the image into small tiles, calculating histograms for each tile, and then equalizing the histograms independently. This adaptive approach ensures that local contrast enhancement is tailored to the specific characteristics of different image regions, allowing the method to be highly effective in enhancing images affected by non-uniform lighting conditions. By controlling the degree of contrast enhancement through a contrast limit, CLAHE prevents over-amplification of noise, making it suitable for images with varying noise levels. CLAHE's ability to adapt to the local image context makes it a valuable tool in fields such as medical imaging, document processing, remote sensing, and quality control, where the preservation of detail and visibility is crucial for analysis and interpretation. Figure 4 depicts the enhancement of the image using the CLAHE technique.

Step 1: Gather a dataset of images of ancient monuments:
D = {I_1, I_2, ..., I_N}.
Step 2: Apply CLAHE to enhance contrast in each image:
I_i' = CLAHE(I_i), where I_i' is the enhanced image.
Step 3: Utilize image segmentation techniques to identify decay regions. Let the segmented regions be
represented as masks:
Segmentation_Masks = {M_1, M_2, ..., M_N}.
Step 4: Analyze the segmented regions to assess the extent and nature of decay. For instance, measure the size and shape of decayed areas:
Decay_Characteristics = {Size, Shape, Type, Severity}.
Step 5: Apply post-processing to refine the segmentation results or remove small artifacts if needed:
Segmentation_Masks_Refined = PostProcess (Segmentation_Masks).
Step 6: Visualize the segmented decay regions overlaid on the original images:
Visualized_Images = Overlay(I_i, Segmentation_Masks_Refined).
Step 7: Prepare reports summarizing the extent and location of decay based on the analysis:
Reports = {Visualized_Images, Decay_Characteristics}.

Fig 4: Image Enhancement

3.1.5 Image Segmentation using the U-Net method:

Image segmentation is a fundamental task in image processing. The goal of image segmentation is to partition an image into homogeneous regions, making it easier to analyze and extract information from the image. The U-Net segmentation algorithm is a groundbreaking convolutional neural network (CNN) architecture designed for semantic and instance segmentation tasks in computer vision. Its distinctive U-shaped architecture features a contracting path for feature extraction and an expansive path for precise pixel-wise segmentation. U-Net's ability to capture fine-grained details, coupled with its strong feature learning capabilities, has made it particularly valuable in medical image analysis, where it has excelled in applications such as organ segmentation and lesion detection. The architecture's skip connections between the contracting and expansive paths enable the model to recover spatial information lost during down-sampling, improving segmentation accuracy. U-Net has gained widespread adoption in the field of deep learning and image analysis, as its effectiveness in tasks requiring high-resolution segmentation is illustrated in Fig. 5.

Step 1: Begin with a dataset of images (I) and corresponding ground truth segmentation masks (M).

Step 2: Define the U-Net architecture, consisting of an encoder and decoder. The encoder extracts features, and the decoder generates segmentation maps.

Step 3: In the encoder, employ convolutional layers (Conv) with specified kernel sizes (k) and number of filters (n) to reduce spatial dimensions, increasing feature depth:

Y_encoder = Conv(I, k, n)

Apply ReLU activation functions (ReLU) after each convolutional layer.

Step 4: Optionally, use pooling layers (MaxPooling) with a specified pool size (p) for downscaling:

Y encoder = MaxPooling(ReLU(Conv(I, k, n)), p)

Step 5: Include a bottleneck layer that captures the most abstract features:

Y_bottleneck = ReLU(Conv(Y_encoder, k, n))

Step 6: In the decoder, employ transpose convolutional layers (TransConv) to upsample feature maps:

Y_decoder = TransConv(Y_bottleneck, k, n). Utilize skip connections by concatenating (Concat) feature maps from the encoder to preserve spatial information:

Y_decoder = Concat(Y_decoder, Y_encoder)

Step 7: Conclude the network with a convolutional layer having the appropriate number of Output channels for segmentation masks:

Y_output = Conv(Y_decoder, k, num_output_channels)

Use a softmax activation function (Softmax) for multiclass segmentation or a sigmoid (Sigmoid) for binary segmentation:

Segmentation_Mask = Softmax(Y_output) /

Segmentation_Mask = Sigmoid(Y_output)

Step 8: Select a suitable loss function (L) for your task, such as cross-entropy or Dice coefficient loss

Loss = L(Segmentation_Mask, M)

Step 9: Train the U-Net using backpropagation and an optimization algorithm (e.g., SGD or Adam): Adjust the model's parameters to minimize the loss:

Parameters = Parameters - learning_rate * Gradient(Loss)

Step 10: Fine-tune hyperparameters, including learning rate, batch size, dropout rate, and regularization strength, as needed.

Step 11: Use the trained U-Net for inference on new, unlabeled images: Input an image (I) into the network to obtain pixel-wise segmentation maps:

Segmentation_Mask = U-Net(I)

Step 12: Apply post-processing techniques, such as filtering or morphological operations, to refine segmentation results.

Step 13: Evaluate the U-Net's performance using appropriate metrics like Intersection over Union (IoU), Dice coefficient, or accuracy:

Evaluation_Metric = Evaluate(Segmention_Mask, M)

Step 14: Report the results, including visualizations of segmentation maps and a discussion of model performance.

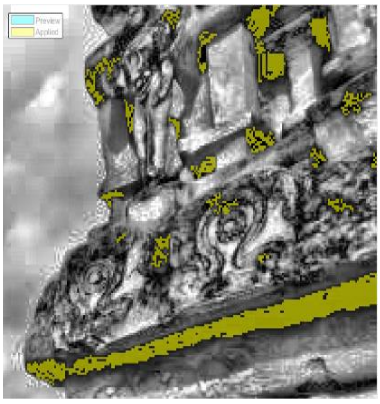


Fig 5: Image Segmentation

3.1.6 Image classification using the SSD method:

SSD is adept at object detection and classification in images, making it suitable for analyzing high-resolution images of monuments. By training SSD with labeled datasets containing examples of cracks, moss, and other defects, the algorithm can learn to identify and classify these features accurately. This approach enables efficient and scalable inspection of monuments, allowing for early detection of defects and timely intervention to mitigate further deterioration. Additionally, SSD can operate autonomously, reducing the need for manual inspection and enabling continuous monitoring of monument condition, thereby aiding in the preservation efforts of these valuable cultural heritage sites.

The following are the practical challenges of implementing the SSD (Single Shot MultiBox Detector) deep learning model in real-world heritage conservation scenarios. The first issue is computational resources; in conservation settings with limited resources, real-time monitoring of large heritage sites may not always be possible due to the need for powerful GPUs or cloud infrastructure. Second, implementing high-resolution imaging systems and ensuring reliable data collection across large monuments can be expensive, which may prevent widespread adoption. Third, ethical issues need to be taken into account to make sure AI-powered evaluations respect cultural sensitivities and don't take precedence over professional advice when it comes to heritage preservation. By serving as an auxiliary tool rather than a substitute for professional inspections, the suggested system can be incorporated into current heritage conservation workflows to overcome these obstacles. For instance, conservationists could receive automated detection results as annotated reports or visual heatmaps, which would enable them to prioritize areas of concern and more effectively allocate resources. This would bridge the gap between advanced AI technology and conventional preservation techniques.

The Single Shot MultiBox Detector (SSD) algorithm is a powerful object detection method that efficiently combines the tasks of object localization and classification in a single deep neural network. SSD accomplishes real-time object detection by employing a multi-scale approach, utilizing multiple feature maps at different resolutions to capture objects of varying sizes and aspect ratios. This versatility, combined with prior box predictions for bounding box generation, enables SSD to efficiently and accurately detect objects in complex scenes across multiple object categories. The algorithm has found applications in fields such as autonomous driving, surveillance, and image-based search, where robust real-time object detection is crucial for various tasks, including pedestrian and object detection in diverse environments.

Step 1: Gather images (D) with annotations for cracks and moss, and seepages.

Step 2: Assign class labels to the anomalies and annotate each image with bounding boxes or segmentation masks for these anomalies.

- Cracks: $C = \{c_1, c_2, ..., c_k\}$
- Moss: $M = \{m_1, m_2, ..., m_l\}$
- Seepages: $S = \{s_1, s_2, ..., s_m\}$

Step 3: Modify SSD architecture for classes: $C \cup M \cup S \cup$ Background.

Step 4: Train SSD to predict bounding boxes B and labels (C, M, S, Background). Define the predicted bounding boxes and class labels for each object in each image as:

$B_i = \{(b_1, c_1), (b_2, c_2), ..., (b_n, c_n)\}$

where b_j is a bounding box and c_j is the associated class label.

Step 5: For each test image I_{test}, the model produces a set of predicted bounding boxes and their associated class labels:

$B\ test = \{(b_1, c_1), (b_2, c_2), ..., (b_k, c_k)\}$

Step 6: Refine results (e.g, filter false positives).

Step 7: Visualize detected anomalies.

Step 8: Assess accuracy with metrics.

Evaluation = {Precision, Recall, F1-Score}

Step 9: Apply the system for heritage preservation and maintenance.

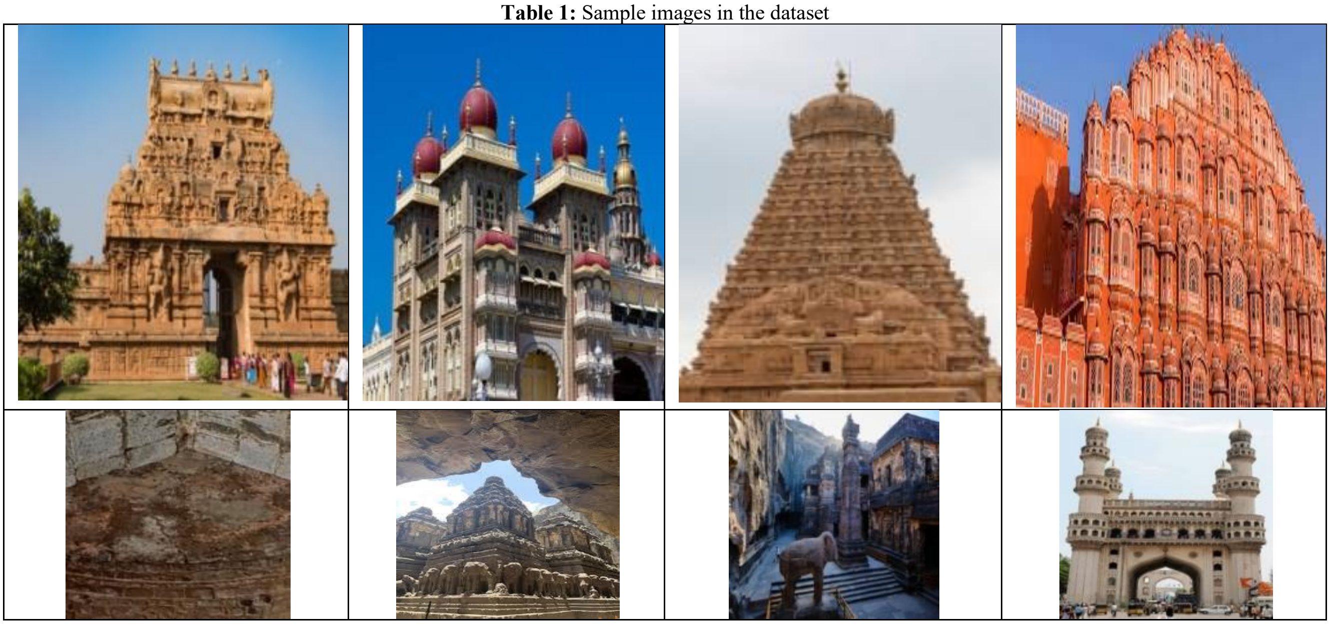
In heritage preservation applications, especially in real-time investigations, it provides more reliability. Moreover, in heritage preservation, it offers efficient object detection, real-time processing capabilities, adaptability to different environments, reduced false positives, scalability, and facilitation of interactive experiences, all of which contribute to more effective and sustainable preservation effects.

4. Results and Discussion

4.1 Dataset Description:

The Indian Monuments Image Dataset comprises high-resolution images from various ancient monuments of 24 directories with 5624 images with annotations for moss, cracks, and seepages. SSD architecture was employed and trained on 75% (4592 images) of the dataset, with the remaining 25% (1032 images) of the dataset used for validation or testing. The study's main drawback is the presumption that decay patterns are uniform across monuments, which might not accurately represent the variety of materials and environmental circumstances found in actual situations. The model can be trained to learn more generalized representations of cracks, moss, and seepages by expanding the dataset by including images from a greater range of monuments, locations, and materials. Another feasible option is transfer learning, which allows models that have already been trained on sizable, varied image datasets to be refined using heritage-specific data, increasing adaptability with fewer samples. Additionally, by capturing various facets of decay, ensemble approaches that integrate predictions from several models (such as Random Forest or SSD with CNN) can improve robustness. Due to variations in building materials (such as marble in Europe versus sandstone in India) and climate-induced weathering (such as tropical moss growth.

Table 1: Sample images in the dataset



4.2 Evaluation Metrics:

The deep learning model achieved significant success in detecting moss and cracks in ancient monuments. Evaluation metrics such as accuracy, precision, recall, and F1 score indicate the effectiveness of the model in distinguishing between different classes of image features.

4.2.1 Accuracy:

Accuracy is the ratio of correctly predicted instances to the total number of instances.

$$\text{Accuracy} = \frac{CorrectPredeictions}{Totalinstances} \quad (1)$$

Equation (1) shows the formula for finding the value of Accuracy.

Table 2: Comparative assessment of Accuracy in %

Algorithm Used	Accuracy		
	Crack	**Moss**	**Seepages**
SVM	91.3	83.1	88.6
KNN	88.5	89.1	88.2
CNN	85.3	91.2	84.5
Random Forest	87.8	86.4	87.2
Proposed Algorithm	96.9	97.5	97.5

Equation (1) shows the formula for finding the value of Accuracy.

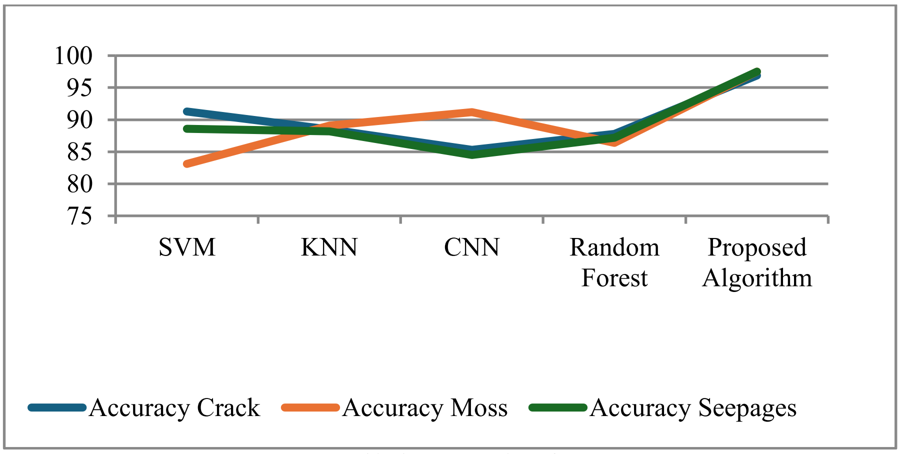


Fig. 6: Graphical representation of Accuracy

Table 2 and Fig. 6 show that the proposed method achieved the accuracy values of Crack, Moss, and Seepages are 96.9%, 97.5% and 97.5% respectively, indicating the percentage of correctly identified instances among the instances predicted as positive.

4.2.2 Precision:

Precision is the ratio of correctly predicted positive observations to the total predicted positives.

$$\text{Precision} = \frac{TruePositives}{TruePositives + FalsePositives} \quad (2)$$

Equation (2) shows the formula for finding the value of Precision.

Table 3: Comparative assessment of Precision in %

Algorithm Used	Precision		
	Crack	Moss	Seepages
SVM	82.4	86.5	75.4
KNN	86	89.4	76.2
CNN	85.1	87.2	79.1
Random Forest	84.7	91.1	82.4
Proposed Algorithm	95.9	90.9	80.4

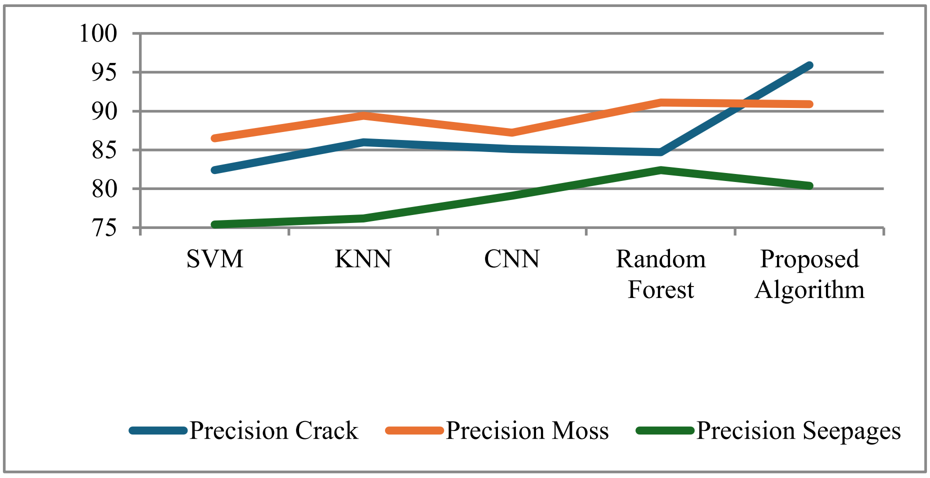


Fig 7: Graphical representation of Precision

Table 3 and Fig. 7 show that the proposed method achieved the precision values of Cracks, Moss, and Seepages are 95.9%, 90.9%, and 80.4% respectively, indicating the percentage of correctly identified instances among the instances predicted as positive.

4.2.3 Recall

Recall is the ratio of correctly predicted positive observations to all observations in the actual class.

$$\text{Recall}=\frac{TruePositivis}{TruePositives+FalseNegatives} \tag{3}$$

Equation (3) shows the formula for finding the value of Recall.

Table 4: Comparative assessment of Recall in %

Algorithm Used	Recall		
	Crack	Moss	Seepages
SVM	90.4	88.2	85.6
KNN	89	91.3	89.2
CNN	84.1	89.5	87.9
Random Forest	90.7	92.4	90.1
Proposed Algorithm	93	93.9	90.8

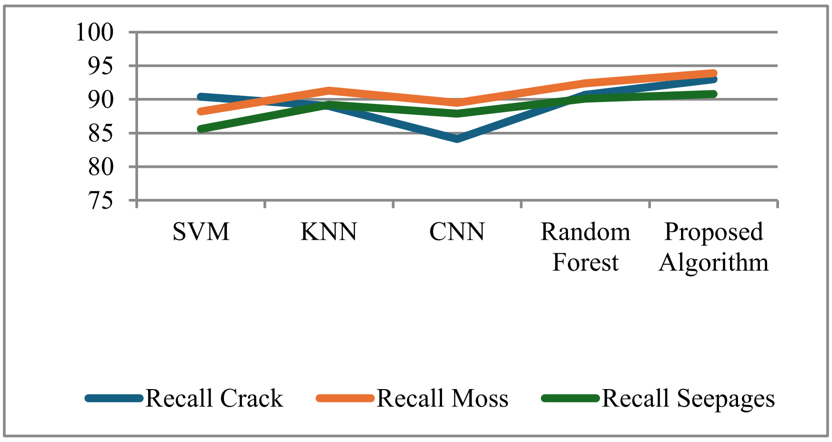


Fig. 8: Graphical representation of Precision

Table 4 and Fig. 8 show that the proposed method achieved the recall values of Cracks, Moss, and Seepages are 93%, 93.9%, and 90.8% respectively, indicating the percentage of correctly identified instances among all actual positive instances.

4.2.4 F1 Score:

F1 Score is the weighted average of Precision and Recall.

$$F1\ Score = \frac{2*precision*recall}{precision+recall} \tag{4}$$

Equation (4) shows the formula for finding the value for the F1 Score.

Table 5: Comparative assessment of F1 Score in %

Algorithm Used	F1		
	Crack	Moss	Seepages
SVM	90.4	902	81.2
KNN	89	89.2	84.2
CNN	84.1	91.4	85.1
Random Forest	90.7	90.1	89.6
Proposed Algorithm	94.4	92.4	85.3

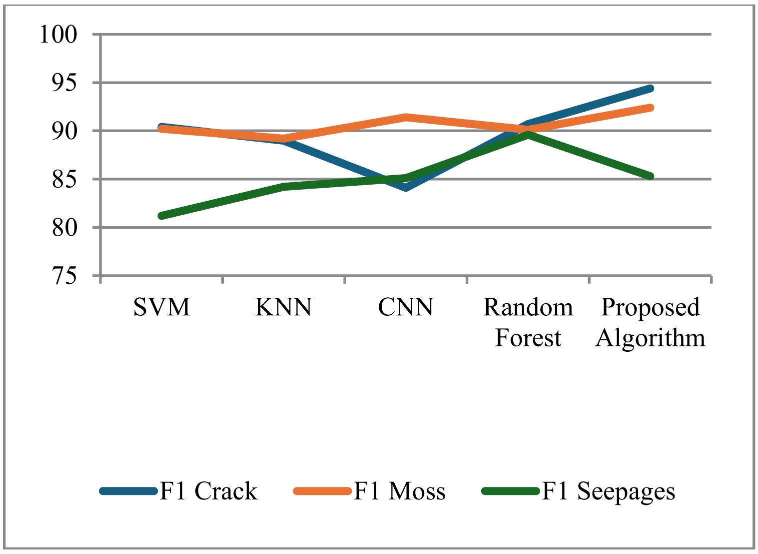


Fig 9: Graphical representation of F1 Score

Table 5 and Fig. 9 show that the proposed method achieved the F1 Score values of Cracks, Moss and Seepages are 94.4%, 92.4%, and 85.3% respectively, providing a balance between precision and recall, representing the harmonic mean of the two metrics.

Hyperparameters influence the performance and reproducibility of deep learning models. Hyperparameters such as learning rate [0.001, 0.01, 0.1], batch size [16, 32, 64], number of U-Net layers [4, 5, 6], dropout rate [0.2, 0.3, 0.4], and activation function [ReLU, Leaky ReLU] were systematically tuned using a combination of random search and manual fine-tuning, with evaluation metrics. The optimal hyperparameter configuration, determined through a stratified 5-fold cross-validation process, included a learning rate of 0.01, batch size of 32, five U-Net layers, dropout rate of 0.3, and Leaky ReLU activation function, achieving an average F1 score of 94.4, 92.4, and 85.3, respectively. These findings underscore the significance of hyperparameter optimization in enhancing the transparency, reproducibility, and accuracy of crack moss and seepage detection systems, contributing to efficient infrastructure maintenance.

The proposed method, which uses the Single Shot Multibox Detector (SSD) with carefully optimized hyperparameters, has given superior performance compared to traditional algorithms such as Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Convolutional Neural Networks (CNN), and Random Forest. Through appropriate and rigorous experimentation and hyperparameter tuning, the proposed model has achieved remarkable accuracy, recall, precision, and F1 score in detecting cracks, moss, and seepages in ancient monuments. The SSD model with optimized hyperparameters not only outperformed these baseline algorithms but also showcased advantages in terms of computational efficiency and scalability.

4.2.5 ROC-AUC Analysis:

Fig. 10 depicts the AUC-ROC curve analysis of ancient monument preservation utilizing Single Shot MultiBox Detector (SSD) deep learning outputs, which provides valuable insights into the model's performance in detecting anomalies like cracks, moss, and seepages. By plotting true positive rates against false positive rates at various detection thresholds, the curve illustrates the trade-off between sensitivity and specificity, crucial for evaluating the model's discrimination ability. A higher AUC value indicates superior performance in distinguishing between positive and negative instances, signifying the model's effectiveness in identifying anomalies accurately. Interpretation of the curve's shape and steepness offers insights into the model's robustness and potential areas for improvement. Additionally, comparing the AUC of the SSD model with other state-of-the-art methods like SVM, CNN, KNN, and random forest can inform decision-making regarding algorithm selection for monument preservation tasks, facilitating proactive conservation efforts and ensuring the longevity of cultural heritage.

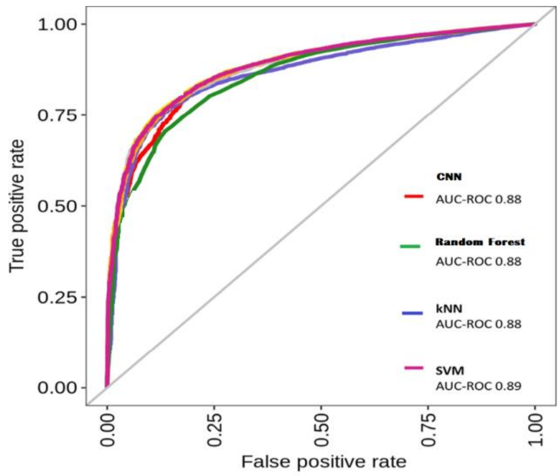


Fig. 10: Graphical representation of ROC-AUC Analysis

4.3 Confusion Matrix:

A multiclass confusion matrix is a table that is used to visualize the performance of the classification algorithm for a problem that contains more than 2 classes. Each row present in the matrix represents the actual class, and each column present in the matrix represents the predicted class. Figure 11 represents the multiclass confusion matrix.

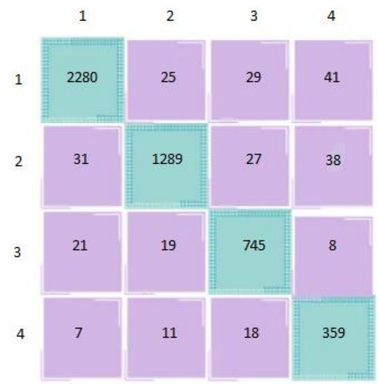


Fig. 11: Multiclass Confusion Matrix

Here, the paper contains 4 classes, namely normal, crack, moss, and seepage, which can be depicted as classes 1, 2, 3, and 4, respectively.
Actual Normal:
Out of 2375 instances, the model correctly predicts the normal as 2280 instances, misclassifies 25 instances as crack, misclassifies 29 instances as moss, and misclassifies 41 instances as seepage. Table 6 represents the confusion matrix for class 1.

Table 6: Confusion matrix for Class 1

		Predicted class	
Actual Class		1 Yes	2, 3 & 4 No
	1Yes	TP – 2280	FN-95
	2, 3 & 4 No	FP-59	TN-2514

Actual Crack:
Out of 1385 instances, the model correctly predicts crack in 1289 instances, misclassifies 31 instances as normal, misclassifies 27 instances as moss, and misclassifies 38 instances as seepage. Table 7 represents the confusion matrix for class 2.

Table 7: Confusion matrix for Class 2

		Predicted class	
Actual Class		2 Yes	1, 3 & 4 No
	2 Yes	TP 1289	FN 96
	1, 3 & 4 No	FP 55	TN 3508

Actual Moss:
Out of 793 instances, the model correctly predicts moss in 745 instances, misclassifies 21instances as normal, misclassifies 19 instances as crack, and misclassifies 8 instances as seepage. Table 8 represents the confusion matrix for class 3.

Table 8: Confusion matrix for Class 3

		Predicted class	
Actual Class		3 Yes	1, 2 & 4 No
	3 Yes	TP 745	FN 48
	1, 2 & 4 No	FP 74	TN 4081

Actual Seepage :
Out of 395 instances, the model correctly predicts seepage in 359 instances, misclassifies 7 instances as normal, misclassifies 11 instances as crack, and misclassifies 18 instances as moss. Table 9 represents the confusion matrix for class 4.

Table 9: Confusion matrix for Class 4

		Predicted class	
Actual Class		4 Yes	1, 2 &3 No
	4 Yes	TP 359	FN 36
	1, 2 & 3 No	FP 87	TN 4466

4.4 Comparative Assessment of execution time:

This paper experiments to compare the execution time of different models, and the results are given in the following table 10.

Table 10: Comparative Assessment of Execution Time

Algorithms Used	Execution Time (milliseconds)
SVM	50
KNN	30
CNN	40
Random Forest	35
SSD	25

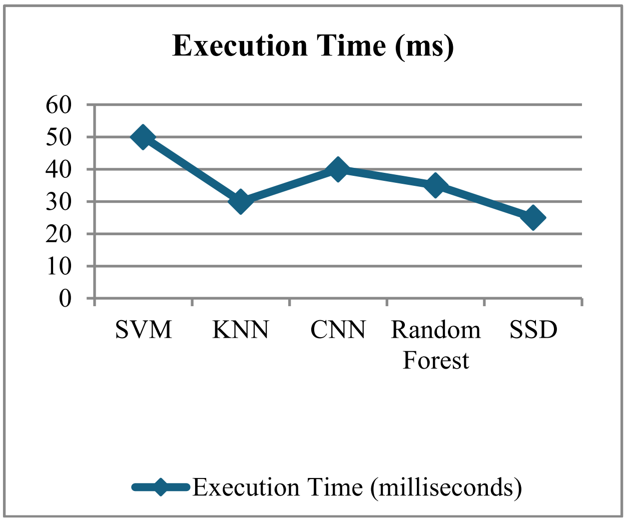


Fig. 12: Graphical representation of Execution Time

Table 10 and Fig. 12 show the comparative assessment of execution time. SSD exhibits the lowest running time among all the tested models, which in turn indicates that the SSD provides high computational efficiency.

5. Conclusion

The study focused on the application of deep learning to safeguard heritage by detecting moss, cracks, and seepages in ancient monuments. After evaluating various models such as SVM, KNN, CNN, Random Forest, and SSD, we identified SSD as the standout performer with superior computational efficiency and the shortest inference time. This positions SSD as the prime choice for real-time applications, ensuring timely and accurate preservation efforts. The proposed method provides the Accuracy values of Crack, Moss, and Seepages are 96.9%, 97.5% and 97.5%. The findings underscore the importance of achieving a balance between accuracy and efficiency in model selection for heritage preservation. The potential of deep learning, particularly exemplified by SSD, holds great promise in protecting ancient monuments from environmental deterioration. SSD may struggle with detecting subtle or nuanced features, particularly in images with complex backgrounds or varying lighting conditions. The ongoing advancements in deep learning technologies will further enhance the effectiveness of heritage conservation endeavors.

References

[1] Hatir, M.E., Barstuğan, M., İnce, İ. (2020). Deep learning-based Weathering type recognition in historical stone monuments. Journal of Cultural Heritage, 45: 193- 203.

[2] Sýkora, M., Diamantidis, D., Holický, M., Marková, J., Rózsás, Á. (2018). Assessment of compressive strength of historic masonry using non-destructive and destructive techniques. Construction and Building Materials, 193: 196-210.

[3] Kordatos, E.Z., Exarchos, D.A., Stavrakos, C., Moropoulou, A., Matikas, T.E. (2013). Infrared thermographic inspection of murals and characterization of degradation in historic monuments. Construction and Building Materials, 48: 1261-1265. 062

[4] Uchida, E., Watanabe, R., Cheng, R., Nakamura, Y., Takeyama, T. (2021). Non-destructive in-situ classification of sandstones used in the Angkor mo6

[5] Costamagna, E., Quintero, M.S., Bianchini, N., Mendes, N., Lourenço, P.B., Su, S., Paik, Y.M., Min, A. (2020). Advanced non-destructive techniques for the diagnosis of historic buildings: The Loka-Hteik-Pan temple in Bagan. Journal of Cultural Heritage, 43:108-117.

[6] Fais, S., Casula, G., Cuccuru, F., Ligas, P., Bianchi, M.G. (2018). An innovative methodology for the non-destructive diagnosis of architectural elements of ancient historical buildings. Scientific Reports, 8(1): 1-11.

[7] Rahimi Jafari, F., Habibi, F., Moazen, S. (2021). Introducing the non-destructive method of photogrammetry in the study and servey of historical monuments. Journal of Research on Archaeometry, 7(2): 135-158.

[8] Clemente, P. (2018). Extending the life-span of cultural heritage structures. Journal of Civil Structural Health Monitoring, 8: 171-179.

[9] Hoang, N.D. (2018). Detection of surface crack in building structures using image processing technique with an improved Otsu method for image thresholding. Advances in Civil Engineering, 2018: 3924120.

[10] Fitzner, B., Heinrichs, K. (2004). Photo atlas of weathering forms on stone monuments, Geological Institute, RWTH Aachen University–Working Group Natural stones and weathering.

[11] Müller, A., Karathanasopoulos, N., Roth, C.C., Mohr, D. (2021). Machine learning classifiers for surface crack detection in fracture experiments. International Journal of Mechanical Sciences, 209: 106698.

[12] Pansare, D., Panchwadkar, A. (2022). Crack detection using image processing. AIP Conference Proceedings, 2653(1): 030007.

[13] R. K. Mehta and D. N. Shah, "Prediction of Longitudinal Dispersion Coefficients in Water Pipelines Using Extreme Learning Machine Conceptions," Journal of Pipeline Systems Engineering and Practice, vol. 10, no. 4, pp. 04019020, Oct. 2019. DOI: 10.1061/(ASCE)PS.1949-1204.0000374.

[14] K. Sarkar and E. Nagaiah, "Prediction of Scour Depth around Long Contractions Using Support Vector Machine and Adaptive Neuro-Fuzzy Inference System," Journal of Hydrologic Engineering, vol. 23, no. 2, pp. 04018074, Feb. 2018. DOI: 10.1061/(ASCE)HE.1943-5584.0001524.

[15] S. Bhattacharjee and S. Parija, "A Novel Multiple-Kernel Support Vector Regression Algorithm for Estimation of Water Quality Parameters," Water Resources Management, vol. 34, no. 6, pp. 1903-1921, Jun. 2020. DOI: 10.1007/s11269-020-02505-3

[16] Chakraverti, S., Agarwal, P., Pattanayak, H. S., Chauhan, S. P. S., Chakraverti, A. K., & Kumar, M. (2023). De-noising the image using DBST-LCM-CLAHE: A deep learning approach. Multimedia Tools and Applications, 1-26.

[17] Ghose, S., Singh, N., & Singh, P. (2020), January). Image denoising using deep learning: Convolutional neural network. In 2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 511-517). IEEE.

[18] Haddadi, Y. R., & Mansouri, B. (2023). A Novel Medical Image Enhancement Algorithm Based on CLAHE and Pelican Optimization.

[19] Siddique, N., Paheding, S., Elkin, C. P., & Devabhaktuni, V. (2021). U-net and its variants for medical image segmentation: A review of theory and applications. Ieee Access, 9, 82031-82057.

[20] Islam, M. M., & Kim, J. M. (2019). Vision-based autonomous crack detection of concrete structures using a fully convolutional encoder–decoder network. Sensors, 19(19), 4251.

[21] Tang, W., Wang, Y., & He, W. (2016). An image segmentation algorithm based on improved multiscale random field model in wavelet domain. Journal of Ambient Intelligence and Humanized Computing, 7, 221-228.

[22] Parida, P., & Bhoi, N. (2018). Fuzzy clustering based transition region extraction for image segmentation. Engineering Science and Technology, an International Journal, 21(4), 547-563.

[23] Perumal, R., & Venkatachalam, S. B. (2021). Non invasive detection of moss and crack in monuments using image processing techniques. Journal of Ambient Intelligence and Humanized Computing, 12, 5277-5285.

[24] Shehata, H. M., Mohamed, Y. S., Abdellatif, M., & Awad, T. H. (2018). Depth estimation of steel cracks using laser and image processing techniques. Alexandria engineering journal, 57(4), 2713-2718.

[25] Loverdos, D., Sarhosis, V., Adamopoulos, E., & Drougkas, A. (2021). An innovative image processing-based framework for the numerical modelling of cracked masonry structures. Automation in Construction, 125, 103633.

[26] A. Umamageswari, S. Deepa, K. Raja, “An enhanced approach for leaf disease identification and classification using deep learning techniques”, Measurement: Sensors, Volume 24, 2022, 100568, ISSN2665-9174

[27] Umamageswari., Johnson, S.D., Sara, D., Kothandaraman, R. (2022), “An enhanced identification and classification algorithm for plant leaf diseases based on deep learning”, Traitement du Signal, Vol. 39, No.3, pp. 1013-1018.

[28] Ramani Perumal, Subbiah Bharathi Venkatachalam (2023), “Non Invasive Decay Analysis of Monument Using Deep Learning Techniques”, Traitement du Signal, Vol. 40, No. 2, 2023, pp. 639-646.