ST. TERESA'S COLLEGE(AUTONOMOUS), ERNAKULAM

AFFILIATED TO MAHATMA GANDHI UNIVERSITY

HERITAGE VISION:A novel CNN classification model for Archaeological Heritage Monuments in India

PROJECT REPORT

In partial fulfilment of the requirements for the award of the degree of

BACHELOR OF SCIENCE IN COMPUTER APPLICATIONS [TRIPLE MAIN]

Submitted By:

DILIN MARIA

III B.Sc. Computer Applications [Triple Main]

Register No: SB22CA011

Under the guidance of

Ms Janeena Shaju

Department of Computer Applications 2022 - 2025

ST. TERESA'S COLLEGE (AUTONOMOUS), ERNAKULAM

AFFILIATED TO MAHATMA GANDHI UNIVERSITY

ST. TERESA'S COLLEGE
(AUTONOMOUS)
ESTD 1925

CERTIFICATE

This is to certify that the project report entitled "**HERITAGE VISION:A novel CNN classification model for Archaeological Heritage Monuments in India**" is a bona fide record of the work done by **DILIN MARIA** (SB22CA011) during the year 2022 – 2025 and submitted in partial fulfilment of the requirements for the degree of Bachelor of Science in Computer Applications (Triple Main) under Mahatma Gandhi University, Kottayam.

17/03/2025

Head of the Department

ST. TERESA'S COLLEGE
DEPT. OF COMPUTERS
ERNAKULAM

17/3/2025

Internal Examiner

17/3/'25

External Examiner

Date:

DECLARATION

I DILIN MARIA (Register no: SB22CA011), B.Sc. Computer Applications [Triple Main] student of St. Teresa's College (Autonomous), Ernakulam, hereby declare that the project submitted for Bachelor's Degree in Computer Application is my original work. I further declare that the said work has not previously been submitted to any other university or academic body.

Date: 14/03/2025

Place: Ernakulam

DILIN MARIA

ACKNOWLEDGMENT

I would like to convey my heartfelt gratitude to **Rev. Dr. Sr. Vinitha (CSST) Manager, Director Rev. Sr. Emeline (CSST) and Principal Dr. Alphonsa Vijaya Joseph** for providing me with this wonderful opportunity to work on a project with the topic HERITAGE VISION : A novel CNN classification model for Archaeological Heritage Monuments in India.

I would like to express my profound gratitude to the Head of the Department of Computer Applications **Sheeba Emmanual** and **my project guide Ms Janeena Shaju** and all other faculty of the department for their contributions to the completion of my project. The completion of the project would not have been possible without their help and insights.

I would also like to Thank my project guide at Ms. Sangeetha Chandran Research Scholar, from CUSAT for training me well to develop this project.

Finally, I take this opportunity to Thank all them who has directly or indirectly helped me with my project.

DILIN MARIA

Paper was presented at **The International Conference on Artificial Intelligence and Machine Learning Conducted by Bharata Mata College Thrikkakara** on 29th November 2024.

ABSTRACT

India with its magnificent and deep-rooted culture can, not only, pride of its rich cultural heritage but also, is home to many historic national monuments. Safeguarding and sustaining these structures is crucial in understanding history and enhancing the sense of cultural belonging. This research paper, named Heritage Vision, seeks to provide effective deep learning architectures and image processing methods to enable accurate recognition and categorization of the selected classes of Indian heritage monuments. The system also contributes to the dataset preservation process by allowing for more in depth classification and also acts as an encyclopaedia for the historians and the tourists for future research. Using a unique dataset, which contained images of selected Indian heritage monuments .we developed and trained a novel model of Convolutional Neural Network (CNN) and its architecture such as ResNet, AlexNet to classify and distinguish these monuments with regards to their geometry ,shape and history. Experimental results show a greater accuracy, indicating the significance of deep learning techniques on the preservation of heritage sites of archeaological importance in india. This study achieved the results of 83.37% of validation accuracy for ResNet, 86% for AlexNet and 93.44% for basic Cnn model respectively.

Contents

1. INTRODUCTION

1.1 About Project

The heritage of Indian monuments is the well-rooted cultural heritage of the Indian nation that carries the centuries of tradition, craftsmanship, and value tracing back to the dawn of time in the realms of ancient civilizations, religions, and art forms. The Taj Mahal, Sanchi Stupa each represents a reflection of extraordinary architecture and innovation which can speak of the stories of the respective aeras they were built in. These sites do not only glimpse across the variety of Indian culture but also act as historical records through inscriptions and reliefs that preserve valuable insights into what has happened in the past, traditions, and lifestyles. Beyond their cultural and historical importance, lies the contribution to these sites make to India’s economy, specifically through tourism. These monuments provide the knowledge to the artists, historians, and students have easy access to available resources where they can learn about the design and artistry of antiquity that would really animate India’s heritage among the youth. It is, therefore, through these monuments that preservation is found in sustaining cultural continuity as they make generations to come connect with this country’s past, which further settle cultural identity in this globalized world. Besides, tourism for conservation attempts ecological and socio-economic balance for the quest of preservation and developments. Preservation and promotion of these monuments and artifacts call for classification and analysis for promotion of research and education. Recently, DL algorithms have attracted massive attention due to their wide application in the classification and analysis of cultural heritage objects. Algorithms are a subgroup of machine learning. They can learn very complex, hierarchical representations of data by processing large amounts of it. These capabilities make DL especially suitable for classifying cultural heritage objects based on attributes such as style, material, and age and for analysis that might be reflecting patterns and relationships between the different artifacts. Several Deep Learning architectures are used to recognize monuments and ensure good performance. Our method employs rich visual features drawn out from satellite images and other images taken using phones in order to automatically discover and classify monuments. We apply CNNs for the classification and heritage recognition. For this purpose, they can learn complex pattern of Indian monuments. thus, they suit themselves to handle diverse and detailed visual characteristics of any heritage site. In this paper, we have elaborated on the different architectures of CNNs, including basic CNN,AlexNet and ResNet.

1.2 Objective of the project

The project's general goal is to show the significance of digital preservation and up keep of cultural heritage data. Through deep learning techniques, namely Convolutional Neural Networks (CNNs), the project intends to classify and evaluate cultural heritage monuments on parameters of style, material, and age class. This way, the recognition of monuments has been made quick by designing complex visual features from satellite images, and furthermore, contributes to the preservation and academic study of heritage.

2.LITERATURE REVIEW

2.1 Image Classification with CNN

A new area of interest in the work of data analysts is aggregation of heritage images. One of the most time consuming areas related to the heritage images is the growing volume of images. This research paper addresses the challenges by classifying the monuments within the studied images through the application of Convolutional Neural Networks. The current images categorization enriches the dataset of images as well and helps in searching for monuments simultaneously, so this enhances study and analysis of cultural heritage. Pre-trained convolutional neural networks such as GoogLeNet, ResNet18, and ResNet50 are recommended in classifying images related to the cultural heritage of the dataset. This work demonstrates a deep learning approach meant for the classification of digital documentation, mainly utilizing images that are colored. It makes use of a convolutional neural network (CNN) in an effective extraction of features and relevant information from images to classify them. Ultimately, the research does recommend the use of a Gabor filter for extracting features coupled with a support vector machine for automatic architectural style recognition.

2.2 Deep Learning Technique

This paper aims to protect the monuments and, at the same time, allow local people and tourists to visit them for educational purposes. The dataset is prepared by combining images taken from Google images and photographs. The extracted images have allowed the model to extract features using Local Binary Patterns and Mean Standard Deviation. A model was designed using deep learning techniques like CNN. The project started by uploading satellite images. These images were further analyzed to determine if a monument could be declared heritage. Further, the model also issues an accurate judgment with regard to the image identified and the visual representation of the monument is provided along with its relevant details. Monuments have proven to be rich sources of historical as well as political knowledge. In a way, they breathe to life our past. They are of immense importance in proving history and are useful in tourism also.

2.3 Monument Recognition

The Heritage Identification of Monuments initiative is to develop a framework for automating the identification and the classification of historical locations. A principal focus of the project is the protection and preservation of the wonderful variety of cultural and historical heritage found throughout the world. This work utilizes one of the most advanced techniques in the scope of deep learning, such as Convolutional Neural Networks (CNNs), to learn features and representations of the several monuments from a very huge set of images. It also enabled the creation of the mobile application in which users could upload images of monuments and retrieve information on their history and cultural importance. The Heritage Identification of Monuments initiative aims to enhance the preservation and sharing of worldwide cultural and historical heritage. Recent developments in computer vision methodologies have recently started contributing a lot in the analysis of monuments, and the particular applications that involve such classification into one of these categories. Though landmark recognition is a subdomain that has been well researched within the computer vision field, identifying monuments is yet challenging. Identification of monuments multicultural country like India. Many fields, like history,conservation, tourism, and education, would surely gain if automatic monument designation were possible.

Title	Authors	Techniques Used	Results	Dataset Size	Year
Detecting surface defects of heritage buildings based on deep learning	Xiaoli Fu and Niwat Angkawisittpan	CNN, YOLOv5, Swin Transformer	mIoU of 90.96 and mAcc of 95.78%	720-Train, 180-Test	2024
Survey Study: Monument Recognition using Artificial Intelligence	MA Hassan, A Hamdy, M Nasr	VGG16, KNN, ResNet, DenseNet	AI techniques recognized the monuments with high accuracy	4708	2023
Heritage identification of monument using deep learning	Dr S Murugesan, Dr N Ramshankar, Hiba Mariam, Kalapoorani P, Kalpika K	MobileNet V1, MobileNet V2	MobileNet V1_acc=97.0%, MobileNet V2_acc=93.3%	1466-Train, 367-Test	2023
A Hybrid Deep Learning Approach for Multi-Classification of Heritage Monuments Using a Real-Phase Image Dataset	Vinay Kukreja, Rishabh Sharma, Satvik Vats	Hybrid CNN and LSTM model	Monuments have accuracy of 92.37%, four classes of Indian monuments have accuracy of 95.89%	3000 images	2023
The Classification of Cultural Heritage Buildings in Athens	Konstantina Siountri and Christos-Nikolaos Anagnostopoulos	YOLO	Average Precision (mAP) metric measures model's accuracy on validation data	6500 images	2023
Machine Learning for Cultural Heritage: A Survey	Marco Fiorucci et al.	Convolutional Neural Networks, PCA, K-means Clustering	Precision, recall, F1 score, AUC-ROC, MSE, MAE	100 or 1000 images	2020

Figure 1

3.METHODOLOGY

3.1 Flow Chart

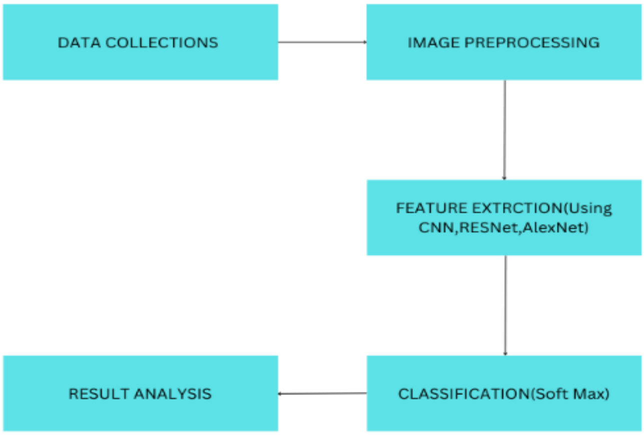

Figure 2

The proposed method in this project involves classifying and identifying Indian monuments through deep learning methods such as Convolutional Neural Networks (CNNs) and its advanced layers like ResNet and AlexNet. The objective is to address the problems inherent in handling large and heterogeneous datasets of heritage images in such a way that the work is done in an efficient and well-organized way, utilizing the data augmentation techniques and deeper CNN architectures like AlexNet and ResNet for feature extraction and classification. The approach centers around the workflow with a view to incorporating data augmentation for enhanced robustness, contemporary CNN architectures for meticulously picking up the details, and stringent validation in order to assess the performance of the entire model. This scientific methodology not only avoids loss of cultural heritage but also utilizes advanced technology in the service of conservation and research towards sustainability.

3.2 Data Collection and Preparation

Images of the following Indian monuments were obtained from online archives, museum collections, and field expeditions: Sanchi Stupa, Taj Mahal, Rani Ki Vav, Hampi, and Sun Temple. Data augmentation techniques in the form of random cropping, flipping, and rotation have been applied to make the model more robust and to expand the dataset to approximately 10,000 images per monument. Image Preprocessing steps included conversion to grayscale, normalization, resizing, and random selection of images. The dataset was then split 80:20 into training & testing data subset .

Figure 3

3.3 Image Preprocessing

Such preparation of data for analysis and model training involve image processing and various methodologies that improve and standardize input images. Data processing refers to the transformation of raw image data into a format suitable for machine learning applications. Data Processing includes grayscaling, normalization and random transformation.

Grascaling: This procedure is the conversion into grayscale, meaning to change a color image into gray tones by getting rid of the color information but retaining the intensity values.

Normalization: Normalization rescales all pixel values into some common range, generally within [0, 1] or [-1, 1], for the sake of improved model stability and performance. A training dataset is thereby made richer and diverse overfitting avoided through the application of random transformations like flipping, cropping, or rotation-that is, through data augmentation. Collectively, these image processing techniques result in a computer vision model that is efficient and accurate.

3.4 Feature Extraction

Deep Convolutional Neural Networks (CNNs), such as ResNet and AlexNet, were used to perform feature extraction. The convolution layers used in such architectures are sensitive to the texture and pattern of monuments. In addition to this, a MSD was applied using deep features extracted from the CNNs to account for the statistical properties of intensity.

3.5 Architectural Network

Model architecture involve an integration of ResNet, AlexNet, and a simple CNN which will be used in classifying of the Indian monuments. ResNet is constructed to use the skip connection which helps in avoiding vanishing gradients. The aligning convolution layers with max-pooling capture many levels of features.It contains an efficient CNN for extracting vital features of images that are classified using fully connected layers, and the output layer is softmax, which demonstrates probabilistic results. The skills of the transfer learning techniques help improve adaptation models toward an Indian monument and upgrade classification performance. It combines three methods- those advanced capabilities of ResNet, speed of AlexNet, and some effects of a simple CNN- achieving very good results for monument classification.

ResNet: This is another deeper neural network architecture that employs residual learning with the objective of solving issues arising from the impairment of deep networks, such as vanishing gradients. ResNet will be helpful in the capturing of complex features from images, and it will aid in high-level understanding of monument structures.

AlexNet: Another deep CNN architecture effective for the high-scale image classification task by taking a large number of convolutional layers and applying max-pooling on them so as to capture different features rapidly and use the fully connected layers to finally classify on the features explored.

Very simple CNN: A simpler architecture suitable for learning simpler feature representations. Though it is simple, it rides high on flexibility in tuning those parameters, eventually leading to faster training of the less complex model.

Combined Strengths: With this stack of ResNet + AlexNet + Simple CNN, the aim is to combine the strengths of different networks so that it will give stronger performance on varieties of images for robust classification.

3.4 Models Used

3.4.1 CNN

A Convolutional Neural Network is thus a deep learning model that has been constructed especially to deal with structured grid-like information types, such as color pixel images. CNNs retain the spatial relationships between pixels, while classical neural networks consider images to be one-dimensional vectors, hence are more suitable for image classification, object detection, and segmentation.

Architecture of CNN:

A CNN have many layers, as follows:-

Convolutional Layers: These layers filter the input image edge, texture, and pattern extraction, ranging from low-level to high-level features. The convolution operation is preferred because it uses less memory and computes fewer parameters than a fully connected layer.

Activation Function (ReLU): The Rectified Linear Unit introduces non-linearity by putting zero for negative pixel values. It helps the model learn complex features with relative ease.

Pooling Layers: These layers are used for down-sampling feature maps while keeping the most significant information. Max Pooling uses the maximum value in the region, while an average pooling takes the average.

Fully Connected Layers: These do involve conducting classification on inputs that are flattened feature maps after feature extraction. They use activation functions such as Softmax for the output probability distributions for different classes.

Dropout Layers (Regularization): Regularization in a convolutional neural network refers to methods by which one can configure the network in order to limit overfitting using techniques such as sample dropout during training. CNNs further initiated major transformations in computer vision that are widely applicable in real-life settings, such as facial recognition and medical imaging, and autonomous vehicles.

3.4.2 AlexNet

AlexNet AlexNet was the first in 2012 to introduce the architecture of a CNN, leading the way for deeper neural networks with convolution and pooling layers. It showed the enormous potential of deep learning in big datasets of images. Key contributions of AlexNet included the usage of ReLU activation functions, where very few neurons become active, hence allowing faster convergence, and the use of dropout layers to prevent overfitting. So this brought a lot of performance gains over its predecessors. So, it became the mother of all subsequent breakthroughs in computer vision.

AlexNet represents a deep convolutional neural network (CNN) that triumphed in the ImageNet competition in 2012.

Architecture of AlexNet:

The architecture comprises a total of 8 layers, including 5 convolutional layers and 3 fully connected layers.

Conv1(Convolutional Layer)-Extracts the low-level features such as edges and textures with the help of 96 11x11-sized filters with ReLU activation. Max pooling (3x3, stride 2) after that.

Conv2 (Convolutional Layer)-Detected more complex structures with 256 5x5-size filters. Then ReLU activation followed by max pooling (3x3, stride 2).

Conv3 (Convolutional Layer)-Utilizes 384 3x3 filters to improve mid-level feature extraction. ReLU activation is used.

Conv4 (Convolutional Layer)- Uses 384 3x3 size filters to filter features learned in earlier layers. ReLU activation used.

Conv5 (Convolutional Layer)- Uses 256 3x3 filters for terminal feature extraction. Following max pooling (3x3, stride 2).

FC6 (Fully Connected Layer)- Consists of 4096 neurons to process high-level features. Uses ReLU activation and dropout methods for regularization.

FC7(Fully Connected Layer)-And yet another 4096-neuron layer with ReLU activation and dropout for better generalization.

FC8 (Fully Connected Output Layer)- It has 1000 neurons for classification into 1000 classes with the softmax activation.

3.4.3 ResNet

ResNet ResNet was developed to deal with the problems of training more profound networks, like the vanishing gradient problems. This paper introduced residual connections, thereby allowing parts of the network to skip layers and thus make deeper learning possible without suffering degradation in performance. ResNet made it possible to train truly deep networks with great accuracy, while still ensuring some effective gradient propagation by stacking multiple layers. Its efficacy in highly complex recognition tasks, such as image classification and cultural heritage analysis, makes the model a key one in deep learning. The methodology of using CNN architectures like ResNet allows for the categorizing of cultural heritage artifacts by combining information flow from various sources while still respecting temporal and contextual provenance, all in support of digital preservation**.**The ResNet uses the following layers.

Architecture of ReNet:

ReNet Layer 1 - Scans image patches row by row using RNNs to mine local spatial relations.

ReNet Layer 2 - Processes the output generated by Layer 1 in a vertical orientation to effectively capture global dependencies.

Stacked ReNet Layers -Other layers further sharpen hierarchical feature extraction, e.g., in deep CNNs.

Fully Connected Layer- Flattens the features extracted and prepares them for classification.

Softmax Output Layer -Calculates probability distribution across classes for ultimate decision-making.

3.5 Training and Testing

The data was split into 80% for training and 20% for testing.

The models were trained on a high-performance system running on the Google Colab notebook.

Hyperparameters are:

Epochs: 10

Learning rate: 0.001

Batch size: 128

While testing, the evaluative metrics include accuracy, precision, recall, F1, and test loss. AlexNet produced a high test accuracy of 84.44%, while ResNet gave a high test loss of 33.19%, thus showing to be a more stable model**.**

3.6 Result Analysis

Image dataset includes 6 monuments with a total of 10,000 images per monument. The dataset is divided into the training dataset with features of 8000 images and testing dataset with 2000 images.

A simple CNN, an AlexNet, and a ResNet-based CNN were evaluated and determined to have performance-based measures towards the learning and classification of Indian monuments. It was trained with hyperparameters that are 10 epochs, the learning rate of 0.001, and the batch size of 128. The testing models were evaluated based on the training accuracy, test accuracy, validation accuracy, precision, recall, F1 score , and test loss.

AlexNet produced an accuracy of 84.44% with precision of 86%, with test loss of 53.65% as its performance. Test accuracy of ResNet was at a deep low-tier standard at 83.17%, having the least test loss at 33.19%, giving superior generalization and stability. The CNN model has highest validation accuracy at 93.44%, with test accuracy of AlexNet at an edge with 83.49%.

In terms of high test accuracy ranking with precision and recall amongst the Indian monuments, AlexNet is a robust model due to this property which makes it capable for implementation in real life like a mobile application. RestNet model has room for improvement with its little test loss if some more epochs or some more data for training is used.

The AlexNet is effective and efficient. AlexNet is the best model.It attains high test accuracy and precision. On further fine-tuning ResNet, it could outperform, stating that it is a more excellent candidate for research for the future.

4.ALGORITHMS

4.1 Image Processing

Input: Dataset Path D path
Output: Preprocessed Images
1. Mount Google Drive and set dataset path.
2. Retrieve all image paths I list from D path
3. Randomly sample 25 images for visualization.
4. For each image in I list .
• Load image using skimage.
• Convert to NumPy array.
• Display sample images using Matplotlib.
5. Compute image dimensions (height, width).
6. Find minimum and maximum image heights and widths.
7. Define resized image path R path
8. For each image in I list
• Read image.
• If image is not RGB, skip it.
• Resize image to 256 * 256
• Convert resized image to uint8 format.
• Save resized image to R path .
9. Compute mean and standard deviation of resized images.
10. Normalize images using:
I norm = (I - u)/sigma
11. Display sample normalized images.

4.3 CNN

1 Mount Google Drive
2 Load dataset paths for training and testing
3 Preprocess Images
4 Initialize Image DataGenerator for augmentation
5 Normalize pixel values by rescaling to [0,1]
6 Verify Image Integrity
7 for each image in dataset do
8 if image is corrupted then
9 Remove corrupted image
10 end if
11 end for
12 Create Data Generators
13 Set batch size and image dimensions
14 Generate training and testing batches
15 Build CNN Model

16 Initialize a sequential model
17 Add convolutional layers with tunable filters and kernel sizes
18 Apply MaxPooling after each convolution
19 Add dropout for regularization
20 Flatten and add dense layers
21 Compile model with categorical cross-entropy loss and Adam optimizer
22 Train the Model
23 Fit the model using training data
24 Evaluate the model using testing data
25 Generate Performance Report
26 Compute classification metrics: Accuracy, Precision, Recall, F1-score
27 Save results to CSV file
28 Output Model evaluation metrics and classification report

4.2 ResNet

1 Import Required Libraries
2 Install and import necessary packages
3 Mount Google Drive and define data directories
4 Set image size (150, 150) and batch size 180
5 Data Preprocessing
6 Apply ImageDataGenerator with rescaling
7 Create training, validation, and test data generators
8 Define the Model using ResNet50
9 Load ResNet50 with pre-trained ImageNet weights
10 Freeze base model layers
11 Add Global Average Pooling layer
12 Add Dense layer with tunable units (128 to 512)
13 Add Dropout layer with tunable dropout rate (0.2 to 0.5)
14 Add output Dense layer with softmax activation
15 Compile the model with Adam optimizer and categorical crossentropy loss
16 Hyperparameter Tuning with Keras Tuner
17 Define RandomSearch tuner with objective as validation accuracy
18 Run tuner search with training data for 5 epochs
19 Retrieve best model from tuner
20 Fine-tune the Model
21 Unfreeze last 20 layers of ResNet50
22 Recompile the model with a learning rate of 10−4
23 Train the model with training data for 5 epochs
24 Evaluate the Model
25 Compute test loss and test accuracy on test data

4.3 AlexNet

1 Import Required Libraries
2 Install and import necessary packages
3 Mount Google Drive and define data directories
4 Set image size (227, 227) and batch size 256
5 Data Preprocessing
6 Apply ImageDataGenerator with rescaling

7 Create training and test data generators
8 Define the AlexNet-based Model
9 Create a Sequential CNN model with:
10 - Convolution layers with ReLU activation
11- MaxPooling layers
12 - Flatten layer
13 - Dense layer with 128 neurons and ReLU activation
14 - Output Dense layer with softmax activation for classification
15 Compile the model with Adam optimizer and categorical crossentropy loss
16 Train the Model
17 for each epoch in range(5) do
18 for each batch in training data do
19 Extract batch data
20 Train on the batch
21 end for
22 end for
23 Evaluate the Model
24 Compute test loss and test accuracy
25 Make Predictions
26 Predict class labels on test data
27 Compute classification report and confusion matrix
28 Visualize confusion matrix using heatmap
29 Plot Training History
30 Plot accuracy and loss curves for training and validation

5 .RESULTS

These three models were applied to classify images of Indian monuments. The models were trained from a dataset with 6 classes of monuments, which contains 10,000 images, and the data split into 80% and 20% respectively for training and testing. The training was run for 10 epochs with a learning rate of 0.001 and a batch size of 128. The evaluation metrics were accuracy, precision, recall, F1 score, and test loss.

5.1 CNN

CNN model reached impressive validation accuracy at 93.44%, indicating that it could efficiently learn the key features of the monuments from the training data. The test accuracy of the CNN model was a little lower at 83.49%, indicating a lesser ability to generalize upon seeing new, unseen data compared to other models such as AlexNet and ResNet.

Precision: The CNN model attained a precision of 91.25%, indicating its ability to reduce the number of false positives and be correct whenever it classifies something as a monument.

Recall: The model's recall was at 83.64%, implying that it appropriately covered most relevant instances; nevertheless, there was still a scope for improvement in catching all relevant images.

F1 Score: The F1 score of 86.68% also indicated a healthy balance between precision and recall, suggesting that the model managed reasonably well in classification.

Test Loss: The test loss of 47.06% suggested that predictions by this model were further away from their actual values than some other models. This results in some of the learning inefficiencies during the test phase

5.2 ResNet

ResNet is a bit slower in test accuracy (83.17%) than AlexNet, leading to high generalizability and stability, as shown by its least test loss of 33.19%. This meant the model made the accurate predictions pertaining the testing set.

Precision: 84.51%, which is lower than AlexNet but suggests good performance in avoiding false positives.

Recall: 83.86%, indicating ResNet detected most of the positive instances, still down slightly from AlexNet.

F1 Score: The F1 score was 83.52% and fairly close to the F1 score for AlexNet, implying balanced performance for precision and recall.

Test Loss: The amount for test loss was 33.19% and the lowest of the three, suggesting ResNet is better than the previous two in adipose stability and with very minimum deviations between the average predicted and actual observation during testing.

The better generalization of ResNet achieved by it having a lower test loss makes it a strong contender for work that is taking place in the future involving classification of a monument. Although it has a slightly lower test accuracy than AlexNet, when supplied with more training data or with more epochs, it may outperform its competitor.

5.3 AlexNet

AlexNet proved superior to the CNN model in the case of test accuracy in that it produced an accuracy of 84.44%. Thus, it suggests that its ability to generalize and recognize features in the images is far better than that of the CNN model. Its precision of 86% demonstrates that it gave a relatively smaller false positive prediction.

Precision: 86%; exhibiting great performance and low false-positive rate for the classification.

Recall: 85%; this means that AlexNet captured a relatively higher percentage of the relevant images in the dataset.

F1 Score: The F1 score of 84% was slightly lower than the CNN model, but indeed, it indicated a good trade-off between precision and recall.

Test Loss: A test loss of 53.65% was higher than Resnet, which shows while AlexNet, in reality, produced reasonably high-class predictions in some instances, they were not very close to the actual values. AlexNet is quite suitable for practical applications like mobile apps, where precision and accuracy are considered to be of great importance. Even though its test loss is slightly high, it still presents the prospect of achieving a greater test precision and hence being the most reliable model when to classification of monuments.

Model	Train.acc	Test.acc	Val.acc	Precision	Recall	F1Score
CNN	0.8359	0.8349	0.9344	0.9125	0.8364	0.8668
AlexNet	0.8454	0.8444	0.86	0.86	0.85	0.84
RexNet	0.8435	0.8317	0.8337	0.8451	0.8386	0.8352

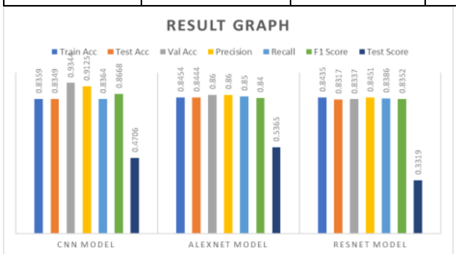

Figure 4

6.CONCLUSION

We successfully developed a system for identifying and also recognizing the Indian Monuments . In this work, models like CNN, ResNet and AlexNet were used for the monument images classification and a comparison study was carried out. The total dataset contained 10,000 images of heritage objects. This dataset was preprocessed and divided into 80% for training and 20% testing. The models trained as well as tested. The best performing model AlexNet achieved test accuracy of 84.44% with precision of 86% indicating its applicability for monument classification. The test accuracy of the ResNet model was relatively lower at 83.17%, but it provided the best generalization with the least test loss of 33.19%. From there, it is assumed that there is a possibility of improvement with further more tuning. The CNN model had the best validation accuracy (93.44%), however, the test result of this model was not as good as the result of AlexNet. This research demonstrates how computer vision and machine learning techniques can facilitate the process of Indian monuments classification and promises more on conserving and enhancing the significance of such historical structures.

7.REFERENCE

[1] Mohammed Hamzah Abed, Muntasir Al-Asfoor, and Zahir M Hussain. Architectural heritage images classification using deep learning with cnn. 2020.

[2] Emmanouil Alexakis, Ekaterini T Delegou, Philip Mavridis, Antonis Rifios, Dimosthenis Kyriazis, and Antonia Moropoulou. A novel application of deep learning approach over irt images for the automated detection of rising damp on historical masonries. Case Studies in results. Remote Sensing, 14(3):553, 2022. Construction Materials, 20:e02889, 2024.

[3] Emmanouil Alexakis, Kyriakos Lampropoulos, Nikolaos Doulamis, Anastasios Doulamis, and Antonia Moropoulou. Deep learning approach for the identification of structural layers in historic monuments from ground penetrating radar images. Scientific Culture, 8(2), 2022.

[4] Mark Altaweel, Adel Khelifi, Zehao Li, Andrea Squitieri, Tasnim Basmaji, and Mohammed Ghazal. Automated archaeological feature detection using deep learning on optical uav imagery: Preliminary

[5] Ariele Cˆamara, Ana de Almeida, David Cac¸ador, and Joao Oliveira. Automated methods for image detection of cultural heritage: Overviews and perspectives. Archaeological Prospection, 30(2):153–169, 2023.

[6] Atharva Ravindra Chonkar, Shubham Rajesh Giri, Shubham Deepak Shirsat, Ganesh Gopal Reddy Guntaka, and Surekha Khot. Heritage identification of monuments. 2023.

[7] Marco Fiorucci, Marina Khoroshiltseva, Massimiliano Pontil, Arianna Traviglia, Alessio Del Bue, and Stuart James. Machine learning for cultural heritage: A survey. Pattern Recognition Letters, 133:102–108, 2020

[8] Xiaoli Fu and Niwat Angkawisittpan. Detecting surface defects of heritage buildings based on deep learning. Journal of Intelligent Systems, 33(1):20230048, 2024

[9] Mennat Allah Hassan, Alaa Hamdy, and Mona Nasr. Survey study: Monument recognition using artificial intelligence. FCI-H Informatics Bulletin, 5(2):1–6, 2023.

[10] Sowjanya Jindam, Jaimini Keerthan Mannem, Meena Nenavath, and Vineela Munigala. Heritage identification of monuments using deep learning techniques. Indian Journal of Image Processing and Recognition (IJIPR), 3(4):1–7, 2023.

[11] Jayanth JT et al. Deep learning applications for identifying and detecting heritage sites. In 2024 International Conference on Inventive Computation Technologies (ICICT), pages 1112–1119. IEEE, 2024.

[12] Alexandra Karamitrou, Fraser Sturt, Petros Bogiatzis, and David Beresford-Jones. Towards the use of artificial intelligence deep learning networks for detection of archaeological sites. Surface Topography: Metrology and Properties, 10(4):044001, 2022.

[13]Classification of heritage monuments using a real-phase image dataset. In 2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA), pages 29–32. IEEE, 2023.

[14] MP Pavan Kumar, B Poornima, HS Nagendraswamy, C Manjunath, BE Rangaswamy, M Varsha, and HP Vinutha. Image abstraction framework as a pre-processing technique for accurate classification of archaeological monuments using machine learning approaches. SN Comput. Sci., 3(1):87, 2022.

[15] Mayank Mishra and Paulo B Lourenc¸o. Artificial intelligence-assisted visual inspection for cultural heritage: State-of-the-art review. Journal of Cultural Heritage, 66:536–550, 2024.

[16] S Murugesan, N Ramshankar, H Hiba Mariam, P Kalapoorani, and K Kalpika. Heritage identification of monuments using deep learning techniques. Journal of Data Acquisition and Processing, 38(3):1927,2023.

[17] Valerio Palma. Towards deep learning for architecture: A monument recognition mobile app. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42:551–556, 2019.

[18] Valerio Palma. Towards deep learning for architecture: A monument recognition mobile app. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information

Sciences, 42:551–556, 2019.
[19] The classification of cultural heritage buildings in athens using deep learning technique Heritage, 6(4):3673– 3705, 2023.
[20] K Zhang, S Teruggi, F Fassi, et al. Machine learning methods for unesco Chinese heritage: Complexity and comparisons. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 46(2):543–550, 2022