iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** IV **Month of publication:** April 2024

DOI: https://doi.org/10.22214/ijraset.2024.59955

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

Heritage Identification of Monument Using Deep Learning

Vedshree Labde[1], Renushree Wakode[2], Shruti Wakode[3], Gitanjali Bhaiswar[4], Tanaya Nawale[5], Prof. Swapnil Sawalkar[6]
[1,2,3,4,5]Scholar,[6]Assistant Professor, Department of Computer Science & Engineering, Sipna College of Engineering and Technology, Amravati, Maharashtra, India

Abstract: ***There are many beautiful and colorful countries in the world which are famous for their culture and heritage. The historical landmarks of such countries convey a great deal through their magnificent structures and rich heritage. The extensive past and legacy of such nations are apparent in the splendid palaces, fortresses, towers, shrines, and cathedrals that are scattered throughout the entirety of their land. We need to protect cultural monuments and landmarks as they are one of the sources of inspiration for future generations. Traditional methods for monument identification often rely on manual inspection or rudimentary computer vision techniques, which can be time-consuming and error-prone. We can provide the best solutions using modern machine learning and deep learning techniques. In our model, we used an approach based on deep learning for the automated identification of heritage monuments from images. We leveraged the power of convolutional neural networks (CNNs) to extract hierarchical features from monument images and train a classifier to recognize different architectural styles and structures. We evaluated our approach on a dataset comprising diverse monument images collected from public archives and online repositories. We applied various pre-processing techniques such as image resizing, normalization of pixel values, and data augmentation. We analyzed and further improved the model's performance for heritage identification of monuments.***
Keywords: Monument Identification, Deep Learning, Convolutional neural network, Image Resizing, Normalization of Pixel Values, Data Augmentation.

I. INTRODUCTION

Preserving our cultural heritage is incredibly important in a world where technology, environmental factors, and urbanization are constantly changing. Monuments are crucial in showcasing historical events, architectural prowess, and the evolution of societies, making them essential parts of our cultural heritage. By preserving these monuments, we can understand the development of our culture and society over time[1]. Many of the images available online are captured by non-professional photographers using their phone cameras. While such images are easy to download and use, they often lack clarity due to the absence of captions and proper categorization, unlike those taken by expert photographers. Accurately Identifying monuments is a crucial first step toward preserving cultural heritage, promoting tourism and education, supporting research, and ensuring responsible urban development. Monument identification refers to the process of recognizing and categorizing various types of monuments or historic structures. This paper proposes a classification approach based on monument image features. The primary objective is to develop a robust framework that can accurately recognize and categorize monuments based on their architectural features. Monument identification with deep learning involves using advanced machine learning techniques, specifically deep neural networks[2], to automatically recognize and categorize monuments or historical structures in images or data. This innovation allows for the development of a website that identify landmarks and provide engaging educational opportunities for both visitors and residents. This technology enables the creation of monument recognition applications that offer interactive learning experiences for tourists and locals alike.

II. LITERATURE SURVEY

TABLE 1
LITERATURE SURVEY

Authors	Approach	Dataset	Accuracy	Research Gap
Tajbakhsh N	Pre-trained CNNs for Medical images Analysis	Utilised pre-trained CNNs fine-tuned for medical Imaging tasks.	Outperformed or matched Scratch trained CNNs.	Demonstrated the potential of pre-trained models with fine-tuning in medical imaging, addressing data scarcity

©IJRASET: All Rights are Reserved

				challenges[3].
Takeuchi and Hebert	Image-Based Landmark Recognition	Utilized image sequences for landmark recognition in urban environments	Achieved robust performance under varied conditions.	Proposed an image- based approach for landmark recognition, addressing challenges like lighting variations and clutter[4].
He et al.	Deep Residual Learning Framework	Evaluated deep residual networks on ImageNet dataset.	Achieved state- of-the-art results in visual recognition tasks.	Introduced a framework for training deeper networks, improving accuracy and performance[5].

This concise summary outlines the approaches, datasets, accuracies and research gaps addressed in each study, providing insights into their contributions to deep learning techniques in their respective domains.

III. METHODOLOGY

In this section, you will find a detailed description of the concepts used to successfully identify monuments.

A. Deep Learning

Deep learning is a form of artificial intelligence categorized within the realm of machine learning. It works by replicating the way humans acquire different types of knowledge[6]. Deep learning models can be trained to perform classification tasks and identify patterns in various forms of data such as photos, text, and audio. Furthermore, it has the ability to execute assignments that formerly necessitated human intellect, such as depicting images or transcribing audio recordings

Deep learning uses neural networks made up of software nodes that work together like neurons in the human brain to learn from labeled data[2]. It learns by discovering complex patterns in data through multi-layered computational models that create multiple levels of abstraction[7].

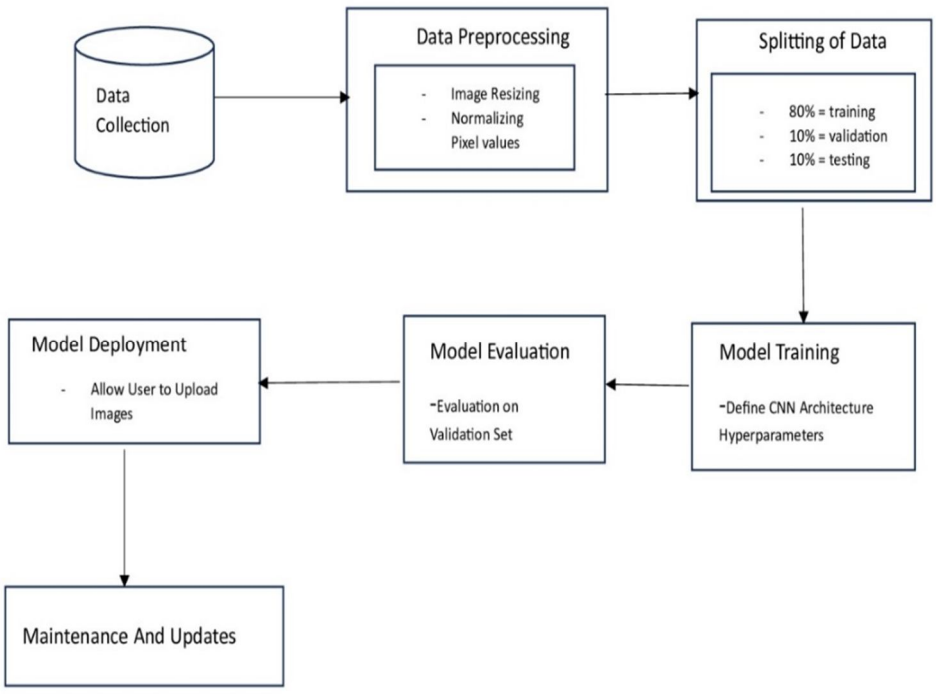


Fig. 1 Workflow

B. Dataset

The dataset comprises images capturing several renowned landmarks and architectural wonders from India. It features various sets of images showcasing Ajanta Caves, Charminar, Ellora Caves, Golden Temple, Mysore Palace, Qutub Minar, Taj Mahal, Tanjavar Temple, Fatehpur Sikri, Gateway of India, Hawa Mahal, India Gate, Lotus Temple, and Sun Temple Konark. From above dataset,80% of the data is allocated for training the model.10% of the data is reserved for testing the model's performance. Another 10% of the data is set aside for validation purposes. This split ensures that the model learns patterns effectively from the training data, while separate datasets for validation and testing help assess its performance on unseen data and tune hyperparameters as needed.

Each dataset provides a comprehensive visual representation of these iconic sites, offering diverse perspectives, angles, and lighting conditions. Researchers and developers can leverage this dataset for a wide range of applications, including image recognition, object detection, cultural heritage preservation, tourism initiatives, and educational projects.

By incorporating these datasets, the collection aims to encapsulate the rich cultural heritage, intricate architectural designs, and historical significance associated with these landmarks, serving as a valuable resource for studying and appreciating India's architectural diversity and historical legacy.

TABLE 2
DATASET OF MONUMENTS





C. Data Cleaning and Preprocessing

Data cleaning and preprocessing are essential for improving data quality and making it suitable for analysis or modeling. Below are the steps used to clean and preprocess image datasets:

1) Data Collection and Preparation :

To carry out this study, we first gathered and prepared the data. The dataset consists of images displaying various monuments, organized in a directory format where each subdirectory represents a different monument class or category. We utilized TensorFlow's tf.keras.preprocessing.image_dataset_from_directory utility to load the dataset, applying specifications for shuffling[8] , resizing[9], and batching. This process ensures that the dataset is formatted appropriately and ready for subsequent analysis. Additionally, we verified the integrity of the dataset and ensured that each image was correctly assigned its respective class label.

2) Data Augmentation :

In this study, we utilized various data augmentation techniques to enhance the quality of the training data[10]. These techniques include Rotation, which helps the model to become more invariant to orientation variations in the input data; horizontal and vertical flipping, which helps the model to learn features that are invariant to left-right or top-bottom orientations, respectively; zooming and scaling, which helps the model to generalize better to objects of different sizes present in the real-world data; and brightness and contrast adjustment, which helps the model to become more robust to variations in illumination. By applying these techniques, we enriched the training data with variations of existing images[11].

3) Preprocessing Layers Integration :

The neural network architecture includes a preprocessing pipeline[12] with two crucial layers to standardize input dimensions and pixel values, facilitating effective deep-learning model training. The pipeline comprises two layers:

- Resizing Layer: We used layers.Resizing to resize input images to 256×256256×256 pixels. Resizing images uniformly ensures consistent input dimensions, preventing distortions during training.
- Rescaling Layer: We rescaled pixel values to [0, 1] using the layers.Rescaling preprocessing layer. This normalizes the input data, standardizing pixel intensities across all images to improve model convergence during training by mitigating input scale issues.

With the assistance of these components, we can effectively and efficiently clean and preprocess data, paving the way for successful training of our neural network model.

4) Data Splitting :

We have divided our dataset into three subsets: training, testing, and validation[13]. The train-test-validation split is a model validation process that helps to check how model would perform with a new data set. The train-test-validation split is crucial for assessing how effectively a machine learning model can extrapolate fresh, unseen data.

This partitioning strategy also acts as a safeguard against overfitting[14], wherein a model may excel with the training data but struggles to generalize to new examples. By leveraging a validation set, practitioners can iteratively fine-tune the model's parameters to optimize its performance on unseen data samples[15].

During the process of data splitting, the dataset has been divided into three primary subsets: the training set (80% of the dataset), utilized to train the model; the validation set (10% of the dataset), employed to monitor model parameters and prevent overfitting; and the testing set (10% of the dataset), utilized to evaluate the model's performance on new data. Each subset had a distinct role in the iterative development of the model.

D. Convolutional Neural Network

The Convolutional Neural Network (CNN) is a class of deep learning network, and it is inspired by an Artificial Neural Network **(ANN)** [16]**.**A type of neural network, Convolutional Neural Network (CNN) is designed for the purpose of processing data such as images. CNNs are primarily used in computer vision tasks, including image recognition, object localization, image bracket, Facial Emotion Recognition. They're inspired by the mortal visual cortex. In 1998, Yann Lecun Developed First Neural Network at AT & TLabs. It can overlook the blank checks.

CNN is the Backbone of Computer Vision, where meaningful Information is uprooted from visual data to take specific conduct.CNN is a kind of Deep learning neural network which were designed from a biological it even models it how the humans brain work it has been through and very efficient for all the image processes pattern recognition kind of application as we can see in the diagram there are different layers in the diagram or let say I have an input image of zebra these image goes through the bunch of the layers of convolutional layer and pooling layer.

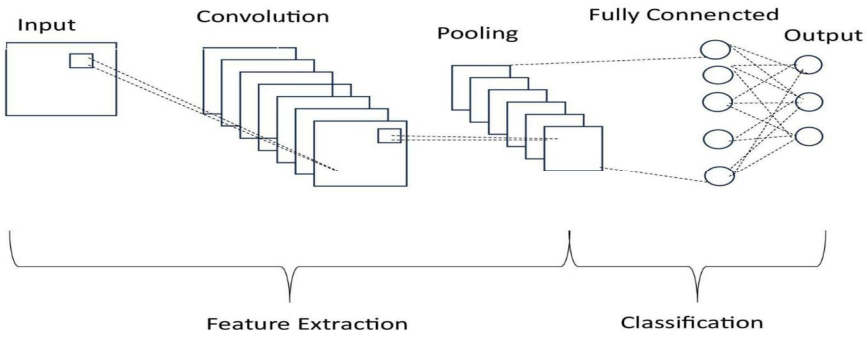


Fig. 2 Architecture Of CNN

In CNN architecture mainly images pass as an input because CNN is a featured based algorithm. It has mainly 3 layers :

1) *Convolutional layer:* The first layer of CNN Architecture is called Convolutional Layer,It has set of Filters. Main objective of the convolutional layer is to extract the useful feature from the images. These layers apply a set of filters. ,These filters also known as kernels to input images to extract features such as edges, textures, and patterns[16].
2) *Pooling layer:* Pooling layer is helping to decide what was the possible or necessary feature to forwarding and deciding the output it improves the efficiency of the CNN model it is used to reduce the dimension to decide[17].
3) *Fully Connected Layers:* The last layer of CNN model consists of fully connected layers.

The term “fully connected” stems from the fact that each neuron in the previous layer is linked to each neuron on the following one[18].The purpose of these layers is to carry out classification, based on the features discovered and output obtained from the previous layers. The feature maps obtained as output from the other two layers are three-dimensional vectors , flatten them to create a single dimensional vector pass to this fully connected layer for the final classification task this layer use the activation function like softmax,Relu to predict the probabilities of classes or for final prediction . Fully connected layer is responsible for making the high level decision such as image classification.

E. Model Evalutation

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

We have used Accuracy for assessing our model that tells the chance of correct prognostications made by a model delicacy score in machine learning is a metric that measures the number of correct prognostications the model made relative to the total number of prognostications. We calculate the standard divagation by dividing the number of correct prognostications by the total number of prognostications.

$$Accuracy = \frac{Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Predictions}$$

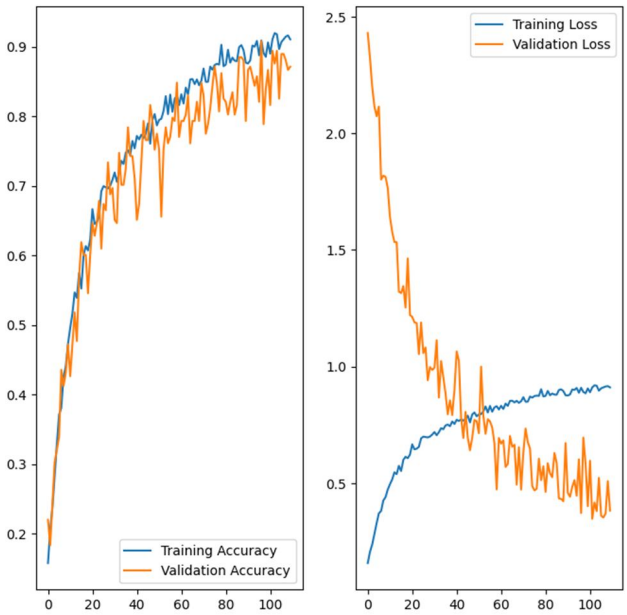


Fig. 3 Training and Validation accuracy and loss

F. Model Deployment

We use Streamlit to present our designs. Streamlit is a machine learning application built in Python. Streamlit is the fastest way to create and share apps. It is the latest model export tool that makes it easy to send all models online and allows you to export models quickly. When we used Streamlit, we came to the conclusion that it is a simple, fast and defined export model.

The following lines of code fully load the saved model, allowing it to be used for prediction, evaluation or further training: -

```
model = tf.keras.models.load_model(" ../saved_models/2.keras")
```

G. Extra Features

We have imported the class named Nominatim from the module named geopy.geocoders. By using this class, it has become possible to access the Nominatim geocoding service. We have used Nominatim geocoding service for converting addresses into latitude and longitude that is geographic coordinates.

Also, we have imported PyDesk library, which is a tool for creating interactive data visualizations, primarily for geospatial data. We have used this library for showcasing maps.

IV. OVERVIEW OF WORK

We need to write following code on command prompt for running our model on streamlit :-

```
streamlit run main.py
```

After entering command, Next image shows the home page of our project.

©IJRASET: All Rights are Reserved

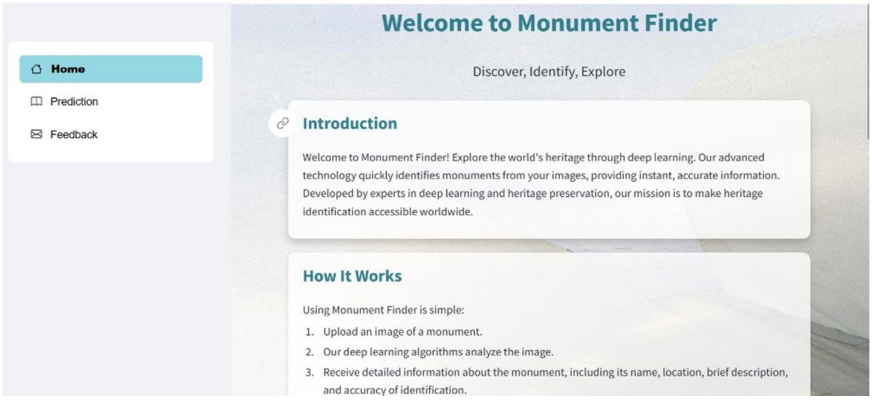


Fig.4 Home page

It gives direction that how the project works and gives the brief introduction to the project. After Click on the Prediction below screen is displayed.

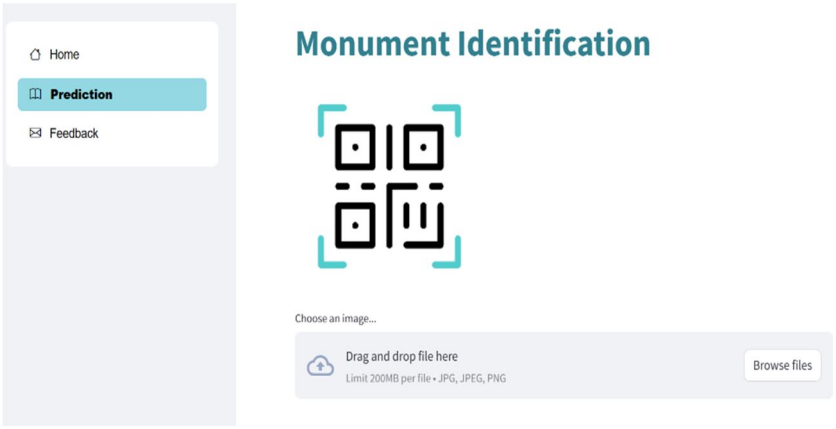


Fig.5 Drag and Drop option for uploading image

The above image represents the interface of our project which includes drag and drop option for uploading an image of monument that we need to identify and to get the related information about it like its name, address, latitude & longitude, and map.

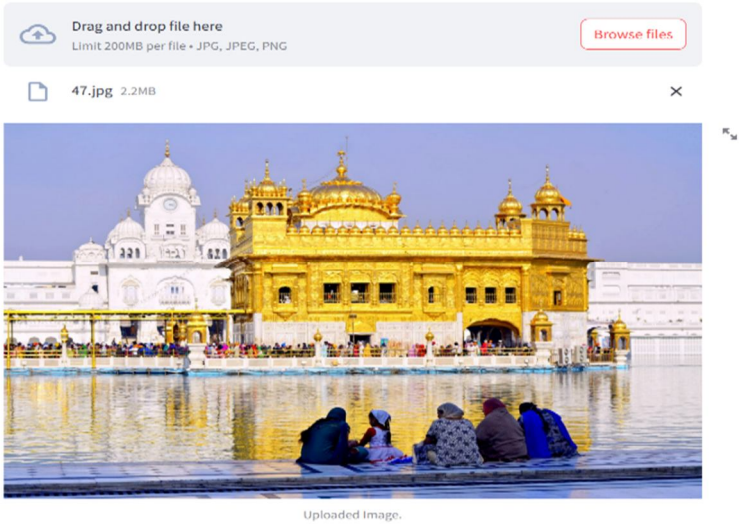


Fig.6 Uploaded Image

After uploading an image of monument, it shows the above screen representing the selected image.



Fig.7 Prediction of Monument

The above screen represents the output that we get, which includes prediction. In prediction, we get name of the monument of uploaded image (Golden Temple in this instance). It also shows accuracy which is the percentage that evaluates how well the model is performing for a particular image. In addition, we get latitude and longitude of the location of monument which is helpful in locating any place on maps.

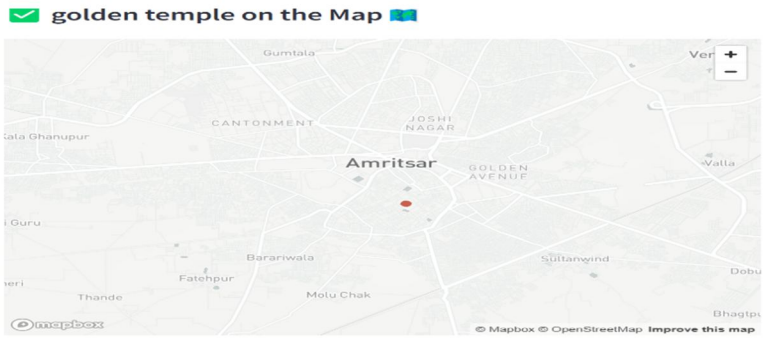


Fig.8 Location of Monument

Lastly, in the output we get the map of the monument's location which is helpful to easily reach the place for the ones who want to visit the place.

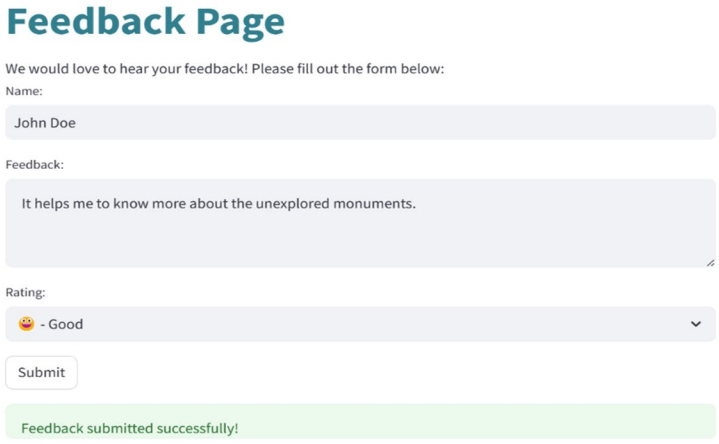


Fig. 9 Feedback for improving performance

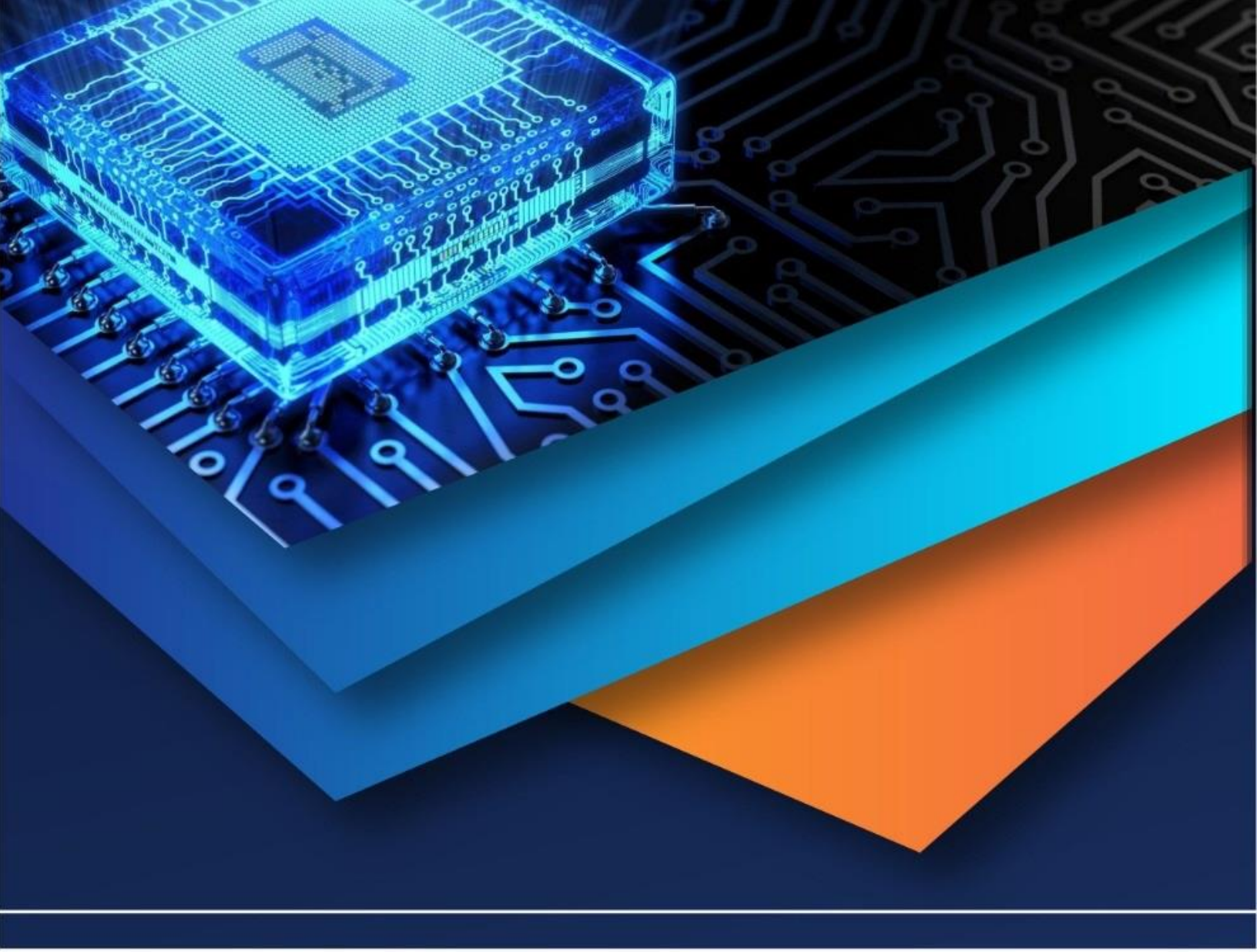
After the Prediction, We welcome your valuable feedback on our model's performance, and based on your observation, we will evaluate our model to further refine its efficiency.

V. CONCLUSION

In conclusion, our research presents innovatory deep learning-based method for automatic recognition of heritage monuments from images, objective for preserving the rich cultural heritage of India. By Utilizing the capabilities of convolutional neural networks (CNNs), we successfully extracted involved features from monument images, allowing accurate classification of architectural styles and structures. Through severe evaluation on a Global dataset, we demonstrated the effectiveness of our approach, achieving notable performance in monument identification. Moreover, we enhanced the model's robustness through accurate preprocessing techniques, including image resizing, pixel value normalization, and data augmentation. Our research shows that using advanced computer technology can help protect important historical sites for the future. This sets a foundation for more progress in preserving our heritage.

REFERENCES

[1] S. Murugesan and H. Mariam H, "HERITAGE IDENTIFICATION OF MONUMENTS USING DEEP LEARNING TECHNIQUES," *Journal of Data Acquisition and Processing*, vol. 38, no. 3, p. 1927, doi: 10.5281/zenodo.98549421.
[2] A. Tavanaei, M. Ghodrati, S. R. Kheradpisheh, T. Masquelier, and A. S. Maida, "Deep Learning in Spiking Neural Networks," Apr. 2018, doi: 10.1016/j.neunet.2018.12.002.
[3] N. Tajbakhsh *et al.*, "Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?," *IEEE Trans Med Imaging*, vol. 35, no. 5, pp. 1299–1312, May 2016, doi: 10.1109/TMI.2016.2535302.
[4] Y. 'Takeuchi, P. 'Gros, M. 'Heber, and K. 'Ikeuchi, "Visual Learning for Landmark Recognition," 1997.
[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," Dec. 2015, [Online]. Available: http://arxiv.org/abs/1512.03385
[6] E. J. Lee, Y. H. Kim, N. Kim, and D. W. Kang, "Deep into the brain: Artificial intelligence in stroke imaging," *Journal of Stroke*, vol. 19, no. 3. Korean Stroke Society, pp. 277–285, Sep. 01, 2017. doi: 10.5853/jos.2017.02054.
[7] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539ï.
[8] T. T. Nguyen *et al.*, "Why Globally Re-shuffle? Revisiting Data Shuffling in Large Scale Deep Learning," in *2022 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, IEEE, May 2022, pp. 1085–1096. doi: 10.1109/IPDPS53621.2022.00109.
[9] S. Saponara and A. Elhanashi, "Impact of Image Resizing on Deep Learning Detectors for Training Time and Model Performance," in *Lecture Notes in Electrical Engineering*, Springer Science and Business Media Deutschland GmbH, 2022, pp. 10–17. doi: 10.1007/978-3-030-95498-7_2.
[10] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *J Big Data*, vol. 6, no. 1, Dec. 2019, doi: 10.1186/s40537-019-0197-0.
[11] A. Mikołajczyk and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," in *2018 International Interdisciplinary PhD Workshop, IIPhDW 2018*, Institute of Electrical and Electronics Engineers Inc., Jun. 2018, pp. 117–122. doi: 10.1109/IIPHDW.2018.8388338.
[12] A. Isenko, R. Mayer, J. Jedele, and H. A. Jacobsen, "Where Is My Training Bottleneck? Hidden Trade-Offs in Deep Learning Preprocessing Pipelines," in *Proceedings of the ACM SIGMOD International Conference on Management of Data*, Association for Computing Machinery, Jun. 2022, pp. 1825–1839. doi: 10.1145/3514221.3517848.
[13] 'Zaccone and M. R. 'Karim, *Deep Learning with TensorFlow: Explore neural networks and build intelligent systems with Python*, 2nd ed. Birmingham: Packt Publishing Ltd, 2018.
[14] D. M. Hawkins, "The Problem of Overfitting," *Journal of Chemical Information and Computer Sciences*, vol. 44, no. 1. pp. 1–12, Jan. 2004. doi: 10.1021/ci0342472.
[15] Z. Liu *et al.*, "Improved Fine-Tuning by Better Leveraging Pre-Training Data." [Online]. Available: https://github.com/ziquanliu/NeurIPS2022_UOT_fine_tuning.
[16] M. Hamzah, A. Muntasir, A.-A. Zahir, and M. Hussain, "Architectural heritage images classification using deep learning Architectural heritage images classification using deep learning with CNN with CNN Architectural Heritage Images Classification Using Deep Learning With CNN." [Online]. Available: http://ceur-ws.org/Vol-2602/
[17] S. Jindam, J. K. Mannem, M. Nenavath, and V. Munigala, "Heritage Identification of Monuments using Deep Learning Techniques," *Indian Journal of Image Processing and Recognition*, vol. 3, no. 4, pp. 1–7, Jun. 2023, doi: 10.54105/ijipr.D1022.063423.
[18] T. Stougiannis, "Landmark and monument recognition with Deep Learning Theodoros Stougiannis Landmark and monument recognition with Deep Learning," 2021.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)