International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 04 | April - 2023 Impact Factor: 8.176 ISSN: 2582-3930

HERITAGE IDENTIFICATION OF MONUMENTS

Atharva Ravindra Chonkar
Information Technology
A.C. Patil College of Engineering Navi Mumbai, Kharghar
atharvarchonkar@acpce.ac.in

Shubham Rajesh Giri
Information Technology
A.C. Patil College of Engineering
Navi Mumbai, Kharghar
shubhamrgiri@acpce.ac.in

Shubham Deepak Shirsat
Information Technology
A.C. Patil College of Engineering
Navi Mumbai, Kharghar
sdeepakshirsat@acpce.ac.in

Ganesh Gopal Reddy Guntaka
Information Technology
A.C. Patil College of Engineering
Navi Mumbai, Kharghar
ganeshrguntaka@acpce.ac.in

Surekha Khot
Information Technology
A.C. Patil College of Engineering
Navi Mumbai, Kharghar
sakhot@ acpce. ac. in

***Abstract*— The goal of the Heritage Identification of Monuments project is to create an automated system for the recognition and classification of historical sites. The project's main goal is to preserve and maintain the rich cultural and historical heritage of the world, which comprises a variety of structures and monuments. Convolutional Neural Networks (CNNs), one of the most advanced deep learning techniques, will be used in this research to extract features and representations of various monuments from a big dataset of photographs. Then, depending on their features and qualities, such as architectural style, location, and historical context, these representations will be utilized to train machine learning models to identify and categorize particular monuments. A mobile application that will enable users to upload pictures of monuments and access details about their historical and cultural significance will also be developed as part of the project. The smartphone application will classify the supplied photographs and provide details on the recognized monuments using the developed machine-learning models. By creating an automated system for the identification and categorization of monuments using computer vision techniques, the Heritage Identification of Monuments project overall wants to contribute to the preservation and promotion of the world's cultural and historical heritage.**

Keywords - Monument identification, Computer vision, Deep learning, Convolutional Neural Networks, Historical context, Mobile application, Machine learning

I. INTRODUCTION

The world has a rich and varied cultural and historical history, whose preservation is crucial in today's fast-paced world. By visiting the locations and conducting first-hand observations, archaeologists and historians have invested a great deal of time and energy into researching the many monuments and architectural styles. Computer vision techniques have now made a foray and are also being used to study the monuments, with applications including the classification of monuments, segmentation of specific architectural styles, and much more. These approaches build upon their work and make a part of that process more efficient and scalable. The process of cognizing and grouping photos of monuments into sub-categories depending on their architectural style is generally referred to as monument categorization.

The classification and acknowledgment of monuments fall within the more general category of landmark identification. Monument recognition is difficult despite the fact that landmark recognition is a well-researched area of computer vision. This is due to a number of issues, including the dearth of annotated datasets of monuments in non-English speaking regions, subtle variations in the architectural styles of monuments, and image samples that vary in perspective, resolution, lighting, scale, and viewpoint. Monument recognition is made substantially more difficult in a varied country like India by these significant obstacles. Several fields, including but not limited to archaeological, historical, conservation, tourist attractions, and education, can benefit from automatic monument designation.

II. PROBLEM STATEMENT

Deep learning algorithms have developed rapidly in recent years as a result of the widespread use of big image datasets and previously unheard-of computer power (DL). Convolutional neural networks (CNNs) have emerged as one of the most successful DL techniques in computer vision, finding use in a variety of fields. This paper presents an ongoing attempt to investigate CNN methods in the domain of architectural heritage, a largely undeveloped research area. The initial processes and outcomes of creating a mobile app to identify monuments are discussed. Heritage technologies have long produced and explored digital models and spatial archives, whereas AI is just starting to interact with the built world through mobile devices.

III. EXISTING SYSTEM

An important duty in the digital documenting of cultural assets is the classification of the photographs that are captured during the measuring of an architectural asset. Since a lot of photos are often handled, classifying them is a laborious operation that frequently takes a long time (and is consequently prone to mistakes). An essential step in the digital documentation process would be enhanced by the availability of automatic approaches to simplify these sporting activities. Also, proper classification of the photos that are available enables better administration and more effective searches through certain terms, aiding in the responsibilities of researching and interpreting the contested heritage asset. This project's main goal is to use deep learning methods for the categorization of photographs of historically significant architecture, specifically by means of convolutional neural networks. The effectiveness of building these networks from scratch or just optimizing already-built networks is assessed for this. All of this has been used to categorize interesting features in pictures of structures having architectural heritage value. As there aren't any datasets of this kind that are appropriate for network training, a new dataset has been produced and made public. In terms of accuracy, promising results have been produced, and it is thought that the use of these approaches can make a substantial contribution to the digital documentation of architectural history.

IV. LITERATURE SURVEY

Throughout the last few decades, researchers have been working on landmark classification, a subset of monument classification, employing a variety of methodologies that can be either global feature-based or local feature based. Global characteristics like edges, textures, and colors are the most basic and need the least amount of resources. Higher-order composite field histograms were shown to be superior by Linde et al. [2] using effective computation on sparse matrices. A covariance descriptor-based method utilizing a Support Vector Machine (SVM) was demonstrated by Ge et al. [3] for the combined classifier and voting strategy.

Torralba et al[4] .'s use of visual context for location

recognition and classification showed how contextual

priming can be used to use scene information for item detection. The Google Landmark Recognition challenge on the GLDv2 dataset [6] was won in 2020 by a unique strategy that used an ensemble of subcentre ArcFace models [5] with dynamic margins and just global features. For monument detection or the object detection class of issues in general, global features are typically employed in conjunction with local features because they lack granularity and cannot focus on Regions of Interest (ROIs). Local characteristics are robust to partial occlusion, illumination variance, and changes in viewpoint and are concentrated on Points of Interest (POI) or Regions of Interest (ROI). Scale-invariant Feature Transform (SIFT) [7] and affine-invariant features [8] are examples of common methodologies. These methods frequently apply a Bag-of-Words (BoW) model [9,10] to the visual words that are packed together and stand in for local attributes. Numerous such approaches have been proposed, including the use of a probability density response map for determining the likelihood of local patches [11], the estimation of patch saliency using contextual information [12], the estimation of patch importance using non- parametric density estimation [13], spatial pyramid kernel- based BoW approaches (SPK-BoW) [14,15], and scalable vocabulary trees [16].

V. METHODOLOGY AND TECHNICAL BACKGROUND

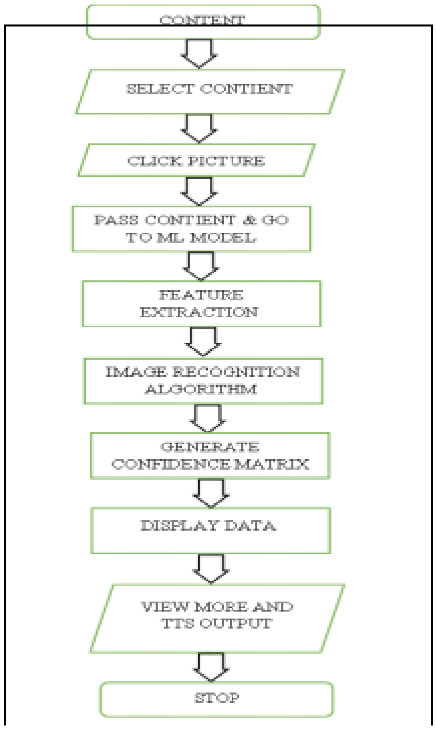

Figure 1 Flowchart

Refer to fig to comprehend the workflow described below:

1. Data Collection: The first step in the process is to gather images of historical monuments. These images will be used to train the model and extract features from them.

2. Feature Extraction: After preprocessing, the features of the images are extracted using Convolutional Neural Networks (CNNs). CNNs are widely used in image recognition and classification tasks due to their ability to learn complex patterns from images.

3. Model Development: A model is developed based on the extracted features, which will classify the images into different categories based on the architectural style of the monument.

4. Validation: The developed model needs to be validated to ensure its accuracy. For this, a test dataset is used, and the model's performance is evaluated based on different metrics like accuracy, precision, recall, and F1 score.

5. Image Utilities: The system uses image utilities to capture images of monuments. These utilities are built using libraries like OpenCV, which helps in identifying the monument in the image.

6. Identification and Information Display: The captured image is passed through the OpenCV module, which identifies the monument in the image. Once the monument is identified, the system displays information related to the monument like its name, location, historical significance, and architectural style.

The system implementation consists of the following modules:

A. Machine Learning

A machine learning model is a mathematical illustration of a system or process used in real life that has been trained on a dataset to make predictions or judgments based on fresh incoming data. The machine learning model is trained on a collection of photographs of historical monuments as part of the project to become capable of identifying and classifying various monument kinds.

Convolutional neural networks (CNNs), a class of neural networks frequently employed for image recognition tasks, are used to build the model. Convolutional, pooling, and fully connected layers are some of the layers that make up CNN, and they all work together to learn and extract features from the input images.

The model learns to link the visual characteristics of the photos with their corresponding labels as it is trained by being shown a series of labeled images. To reduce the discrepancy between the model's predicted labels and the actual labels of the training images, the model's weights and biases are optimized.

Once trained, the model can be used to forecast outcomes for fresh, unexplored photos of monuments. The trained model is fed the input image and, using the characteristics and associations it has learned, returns a predicted label or class for the image.

The amount and quality of the training dataset, the complexity of the CNN architecture and the method of optimization and regularization used all affect the accuracy and performance of the machine learning model.

B. Image Processing

The image Processing Module uses a variety of Python modules and technologies, including OpenCV, to take pictures of historical sites and identify them using a trained machine learning module.

OpenCV (Open Source Computer Vision Library) is a well-known open-source toolkit for computer vision and image processing that offers a number of tools and features for processing images and videos, including picture capture, filtering, segmentation, feature identification, and object recognition.

The Image Processing Module uses OpenCV to capture photographs of historical sites using image utilities like cameras or cellphones. Once the picture has been taken, it is sent via the machine learning model that has been trained to recognize the monument.

Before sending the photos through the machine learning module, the images must first go through the image processing module. Tasks like image scaling, noise removal, and feature extraction may fall under this category.

Once the monument has been located in the photograph, pertinent details about it, like its name, location, historical significance, and other pertinent information, can be shown

C. Android App:

The Heritage Identification of Monuments project requires the Android app because it offers a user-friendly interface for taking pictures of monuments and learning more about them. By utilizing the camera on their phone or tablet to take a picture of a landmark, users will be able to use the app to identify the monument and show pertinent information about it using a trained machine learning model and image processing module.

The Java programming language and the Android Studio development environment will be used to create the Android app. It will support various screen sizes and resolutions and be made to function on a variety of Android devices. Also, the app will be enhanced for speed and dependability with features like caching to reduce network usage and error handling to avert crashes and other problems.

D. CNN: (convolutional neural network)

CNNs are employed in the heritage identification of monuments because they are efficient in analyzing picture data and identifying visual traits and patterns that are particular to each monument.

One method for identifying heritage monuments is to use visual representations of the monuments, such as photographs, and identify distinctive aspects of each site. CNN may then be trained to identify each monument based on its visual qualities using these attributes.

CNNs are particularly well-suited for this task because they are able to learn complex representations of image data without requiring hand-engineered features. This means that the CNN can automatically identify features that are relevant for recognizing different heritage monuments, such as the shape and structure of the monument, as well as any unique patterns or textures.

E. ANN: (Artificial neural network)

In a similar way to how CNNs are frequently employed for this task, ANNs may also be used to identify photographs or pictures of historical monuments.

In this instance, a sizable dataset of photos that includes instances of various heritage landmarks we are interested in identifying can be used to train the ANN. The ANN gains the ability to identify patterns and features in the photos that are connected to each monument, such as certain architectural styles, distinctive features, or distinguishing colors, throughout training.

As the ANN is trained, it may recognise fresh photographs of historical sites by comparing the characteristics retrieved from these images with the features discovered during training. Based on the retrieved features, the ANN can then predict which monument is visible in the image.

When the dataset is small or there aren't enough resources to train a CNN, ANNs can be especially helpful for identifying monuments that are historically significant. The identified photos can also be subjected to more complex analysis using ANNs, such as detecting specific architectural characteristics or classifying monuments according to their style or historical significance.

VI. RESULT AND ANALYSIS

It is the interface of our app which provides the user with the continents. A user can select the continent in which he is present and search for the monuments located in that continent. Location plays an important role in locating a monument.

This is the screen that we get after we select the continent. After that, we have to click on take a picture. When we select that it goes to the camera. Then we have to click the picture of the monument and then the app does the next.

This is the final output screen in this screen we get the image that we have clicked. The prediction of the image in the prediction of the image we get the name of the image(India gate) in this case. Confidence is the percentage that the machine is sure that it is that monument. Confidence is achieved through the machine learning algorithm. It tells us the confidence through the images available and the percentage it is sure that it is the same image. Then in a disc, the option gives us the basic information about the monument. And if we need more information we click on read more which takes us to the google page for that monument. There are many predictions that happen in the back end of the app with the help of ml. It displays the top 3 predictions on the output screen. With the predictions and percentages are sometimes different than one another.

VII. CONCLUSION

There are several potential advantages to creating an automated system for classifying and identifying historical sites. The process of categorizing and recording cultural assets can be made much simpler with the help of computer vision and machine learning tools, which can aid in the preservation of significant historical sites. By giving visitors a more participatory and immersive experience, this technology can help improve the educational and tourism sectors. It is conceivable that we will see additional breakthroughs in this area as computer vision and machine learning continue to grow, with even more advanced algorithms capable of precisely classifying and identifying monuments and landmarks.

ACKNOWLEDGEMENT

I thank my college Principal Dr. V. N. Pawar sir for providing the required resources for the development of the project. I would also like to thank HOD Dr. V. Y. Bhole for suggesting such a great project topic for departmental purposes. My sincere thanks to my Project Guide Prof. S. A Khot for helping, suggesting new ideas, and guiding me throughout the semester. I am also grateful to all the faculty members for their support and encouragement.

REFERENCES

[1] J. S. Hawley, The Bhakti Movement--from Where? Since When? India International Centre, 2009.

[2] O. Linde and T. Lindeberg, "Object recognition using composed receptive field histograms of higher dimensionality," in Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004., 2004, vol. 2, pp. 1-6 Vol.2, doi: 10.1109/ICPR.2004.1333965.

[3] Y. Ge and J. Yu, "A scene recognition algorithm based on covariance descriptor," 2008 IEEE Int. Conf. Cybern. Intell. Syst. CIS 2008, pp. 838– 842, 2008, doi: 10.1109/ICCIS.2008.4670816.

[4] A. Torralba, K. P. Murphy, W. T. Freeman, and M. A. Rubin, "Contextbased vision system for place and object recognition," Radiology, vol. 239, no. 1, p. 301, 2006, doi: 10.1148/radiol.2391051085.

[5] J. Deng, J. Guo, T. Liu, M. Gong, and S. Zafeiriou, "Sub-center ArcFace: Boosting Face Recognition by Large-Scale Noisy Web Faces," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 12356 LNCS, pp. 741–757, 2020, doi: 10.1007/978-3-030-58621-8_43.

[6] T. Weyand and B. Leibe, "Visual landmark recognition from Internet photo collections: A large-scale evaluation," Comput. Vis. Image Underst., vol. 135, pp. 1– 15, 2015, doi: 10.1016/j.cviu.2015.02.002.

[7] D. G. Lowe, "Object recognition from local scale-invariant features," Proc. Seventh IEEE Int. Conf. Comput. Vision, 1999, vol. 2, pp. 1150– 1157, 1999, doi: 10.1109/ICCV.1999.790410.

[8] J. Sivic and A. Zisserman, "Video google: A text retrieval approach to object matching in videos," Proc. IEEE Int. Conf. Comput. Vis., vol. 2, pp. 1470– 1477, 2003, doi: 10.1109/iccv.2003.1238663.

[9] A. Bosch, A. Zisserman, and X. Muñoz, "Scene classification using a hybrid generative/discriminative approach," IEEE Trans. Pattern Anal. Mach. Intell., vol. 30, no. 4, pp. 712–727, 2008, doi: 10.1109/TPAMI.2007.70716.
[10] R. Fergus, P. Perona, and A. Zisserman, "Object class recognition by unsupervised scale-invariant learning," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2, 2003, doi: 10.1109/cvpr.2003.1211479.

[11] L. Lu, K. Toyama, and G. D. Hager, "A two level approach for scene recognition," Proc. - 2005 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognition, CVPR 2005, vol. I, pp. 688–695, 2005, doi: 10.1109/cvpr.2005.51.

[12] D. Parikh, C. L. Zitnick, and T. Chen, "Determining Patch Saliency Using Low-Level Context," in Computer Vision -- ECCV 2008, 2008, pp. 446–459.

[13] J. Lim, Y. Li, Y. You, and J. Chevallet, "Scene Recognition with Camera Phones for Tourist Information Access," Jul. 2007, doi: 10.1109/ICME.2007.4284596.

[14] T. Chen and K. H. Yap, "Discriminative BoW framework for mobile Landmark recognition," IEEE Trans. Cybern., vol. 44, no. 5, pp. 695– 706, 2014, doi:

10.1109/TCYB.2013.2267015.

[15] J. Cao et al., "Landmark recognition with sparse representation classification and extreme learning machine," J. Franklin Inst., vol. 352, no. 10, pp. 4528–4545, 2015, doi: 10.1016/j.jfranklin.2015.07.002.

[16] D. Nistér and H. Stewénius, "Scalable recognition with a vocabulary tree," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2, pp. 2161–2168, 2006, doi: 10.1109/CVPR.2006.264.