iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** IV **Month of publication:** April 2023

DOI: https://doi.org/10.22214/ijraset.2023.50168

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 11 Issue IV Apr 2023- Available at www.ijraset.com

Heritage Identification of Monuments

Atharva Ravindra Chonkar[1], Shubham Rajesh Giri[2], Shubham Deepak Shirsat[3], Ganesh Gopal Reddy Guntaka[4], Surekha Khot[5]
Information Technology, A.C. Patil College of Engineering Navi Mumbai, Kharghar

Abstract: ***The creation of an automated system for the identification and classification of historical sites is the aim of the Heritage Identification of Monuments project. The major objective of the initiative is to conserve and protect the diverse array of buildings and monuments that make up the rich cultural and historical legacy of the world. In this study, characteristics and representations of numerous monuments will be extracted from a large collection of images using convolutional neural networks (CNNs), one of the most sophisticated deep learning approaches. These representations will then be used to train machine learning models to recognize and classify specific monuments based on their characteristics and traits, such as architectural style, location, and historical context. As part of the initiative, a smartphone application that will let people contribute images of monuments and obtain information about their historical and cultural significance will also be created Using the created machine-learning models, the smartphone application will categorize the provided photos and provide information on the recognized monuments. The ultimate goal of the Heritage Identification of Monuments project is to aid in the protection and promotion of the world's cultural and historical heritage by developing an automated system for the identification and categorization of monuments using computer vision techniques.***

Convolutional neural networks, deep learning, historical context, mobile applications, and monument recognition are some relevant keywords.

I. INTRODUCTION

In today's fast-paced world, it is essential to preserve the rich and diverse cultural and historical heritage of the world. Archaeologists and historians have spent a significant deal of time and effort researching the various monuments and architectural styles by traveling to the locations and making first-hand observations. The classification of monuments, the segmentation of particular architectural styles, and many other applications of computer vision techniques are now being utilized to examine the monuments. These methods expand on their work and streamline and scale up a portion of the process. Monument categorization is the process of identifying and classifying photographs of monuments into sub-categories based on their architectural design.

The more general topic of landmark identification includes the categorization and acknowledgment of monuments. Even though landmark recognition is a well-researched topic of computer vision, recognizing monuments is challenging. This is brought on by a variety of problems, such as the scarcity of annotated datasets of monuments in non-English speaking regions, subtle changes in the architectural styles of monuments, and image samples with different perspectives, resolutions, lighting, scales, and viewpoints. These important barriers significantly increase the difficulty of monument recognition in a diverse nation like India. Automatic monument designation has advantages in many areas, including but not limited to educational, historical, conservation, and tourist attractions.

II. PROBLEM STATEMENT

Recent years have seen a rapid development of deep learning algorithms due to the widespread use of large image datasets and previously unheard-of computing power (DL). Convolutional neural networks (CNNs) have become one of the most used DL methods in computer vision, with applications in many different industries. This study describes an ongoing investigation into CNN techniques in the field of architectural heritage, a yet under-researched subject. The development of a smartphone app to identify monuments' first procedures and results is explained. While AI is only beginning to interface with the built world through mobile devices, heritage technologies have long produced and examined digital models and spatial archives.

III. EXISTING SYSTEM

The classification of photos taken while measuring an architectural asset is a crucial task in the digital documentation of cultural assets. Classifying images is a tedious process that frequently requires a lot of time due to the regular handling of many photos (and is consequently prone to mistakes) The availability of automatic methods to make these sporting activities simpler would improve a crucial step in the digital documenting process. Also, correct categorization of the accessible pictures provides better management and more efficient searches using specific terms, aiding in the duties of investigating and interpreting the disputed heritage item.

©IJRASET: All Rights are Reserved

The major objective of this research is to categorize photos of a historically significant building using deep learning techniques, notably convolutional neural networks. the effectiveness of creating these networks from scratch vs simply optimizing ones that have already been created is evaluated. All of this has been utilized to group intriguing details in images of buildings with significant architectural heritage. No datasets of this kind that are suitable for network training exist, hence a new dataset has been created and made available. It is believed that the application of these methodologies can significantly contribute to the digital documentation of architectural history. In terms of accuracy, promising findings have been generated.

IV. LITERATURE SURVEY

Researchers have been studying landmark classification, a subset of monument classification, over the last few decades. They have used a range of approaches that can either be global feature-based or local feature-based. Edges, textures, and colours are among the most fundamental and resource-intensive global properties. Linde et al. [2] demonstrated the superiority of higher-order composite field histograms by performing efficient computation on sparse matrices. Ge et al. [3] demonstrated a covariance descriptor-based approach using a Support Vector Machine (SVM) for the combined classifier and voting technique. Contextual priming can be utilized to use scene information for object detection, as demonstrated by Torralba et al.'s[4] use of visual context for location recognition and categorization. In 2020, a novel approach that made use of an ensemble of subcentre ArcFace models [5] with dynamic margins and just global features was successful in winning the Google Landmark Recognition challenge on the GLDv2 dataset [6]. Due to their lack of granularity and inability to focus on Regions of Interest, global features are often used in conjunction with local features to solve object detection problems in general, such as monument detection. (ROIs) Local features are focused on Points of Interest (POI) or Regions of Interest and are robust to partial occlusion, illumination fluctuation, and changes in viewpoint. (ROI). Common approaches include affine-invariant features [8] and scale-invariant Feature Transform (SIFT) [7]. These techniques commonly model the visual words that are grouped and represent local qualities using a Bag-of-Words (BoW) model [9,10]. Numerous such methods have been put forth, such as the use of a probability density response map to assess the likelihood of local patches [11], the estimation of patch saliency using contextual data [12], the estimation of patch importance using non-parametric density estimation [13], spatial pyramid kernel-based BoW methods (SPK-BoW) [14,15], and scalable vocabulary trees [16].

V. METHODOLOGY AND TECHNICAL BACKGROUND

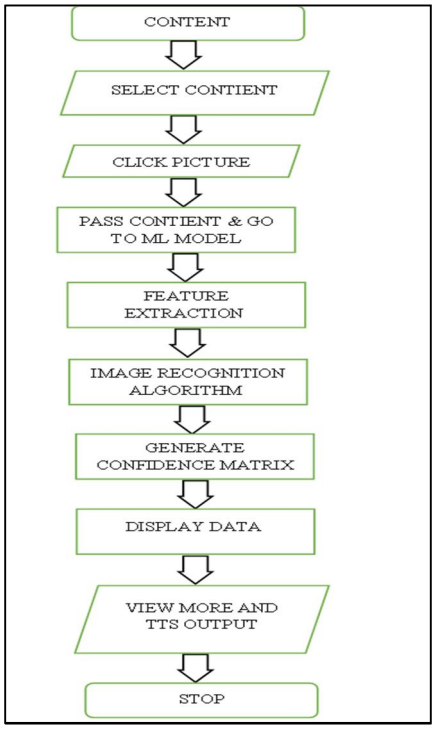


Figure 1 Flowchart

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IV Apr 2023- Available at www.ijraset.com

A. System Methodology

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.

VI. PREPROCESSING: ONCE THE DATA IS COLLECTED, IT NEEDS TO BE PREPROCESSED TO IMPROVE THE QUALITY OF THE IMAGES. THIS INVOLVES RESIZING, CROPPING, AND REMOVING NOISE FROM THE IMAGES

1) *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.
2) *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.
3) *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.
4) *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.
5) *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.

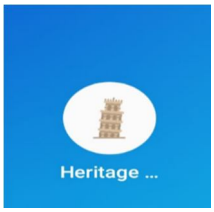
The Image Processing Module uses OpenCV to capture photographs of historical sites using image utilities like cameras or cell phones. Once the picture has been taken, it is sent via the machine learning model that has been trained to recognize the monument.

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

©IJRASET: All Rights are Reserved

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.

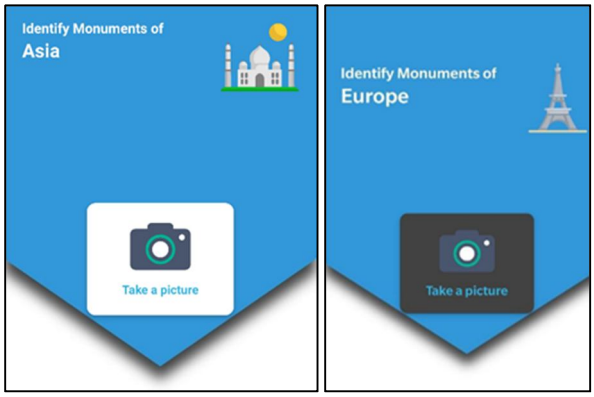
E. ANN: (Artificial Neural Network)

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.'

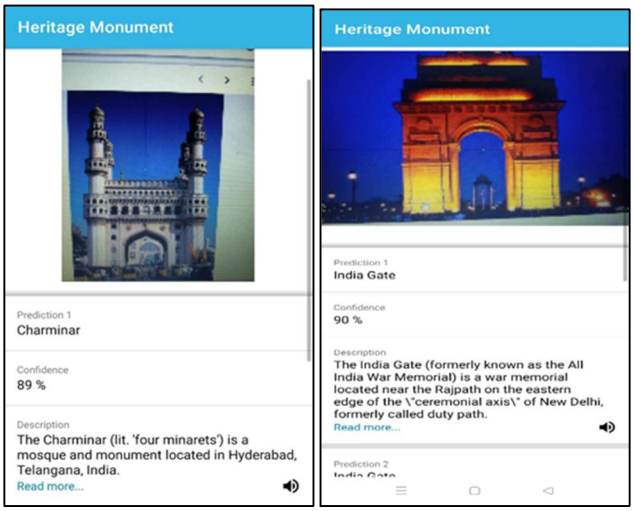
VII. RESULT AND ANALYSIS

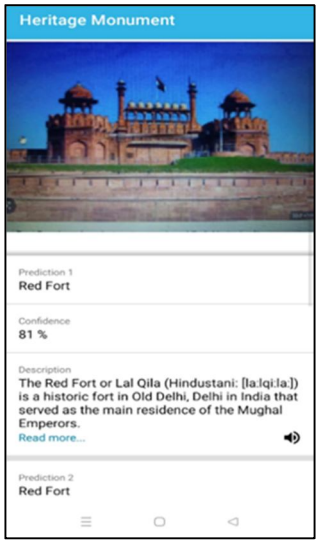


The continents are displayed to the user on our app's UI. The user has the option to look for monuments on the continent in which he is now located by choosing that continent. When locating a monument, location is important.



After selecting the continent, we are provided with this screen. The next step is to click on a picture. The camera is activated when we choose the image. The next step is handled by the app after we click the monument's picture.





The image we clicked is displayed on this screen, which serves as the final product. When an image is predicted, the name of the image, in this case, the India Gate, is revealed. The degree to which a machine is confident that it is that monument is referred to as confidence. Via the machine learning algorithm, confidence is obtained. Using the available photographs and the percentage of certainty that it is the same image, expresses its level of confidence. The selection then provides us with the fundamental details

about the monument in a disc. We can access the Google page for the monument by clicking read more if we require more details. With the help of ml, a lot of predictions are made at the app's back end. In the output screen, the top three forecasts are shown. Predictions and percentages can occasionally diverge from one another.

VIII.CONCLUSION

Also, developing an automated method for categorizing and identifying historical sites has a number of potential benefits. Computer vision and machine learning methods can greatly simplify the process of classifying and recording cultural assets, which can contribute to the preservation of important historical places. This technology can enhance the educational and tourism industries by providing visitors with a more interactive and immersive experience. As computer vision and machine learning develop, it is possible that we may see more advancements in this field, with even more sophisticated algorithms capable of precisely classifying and identifying monuments and landmarks.

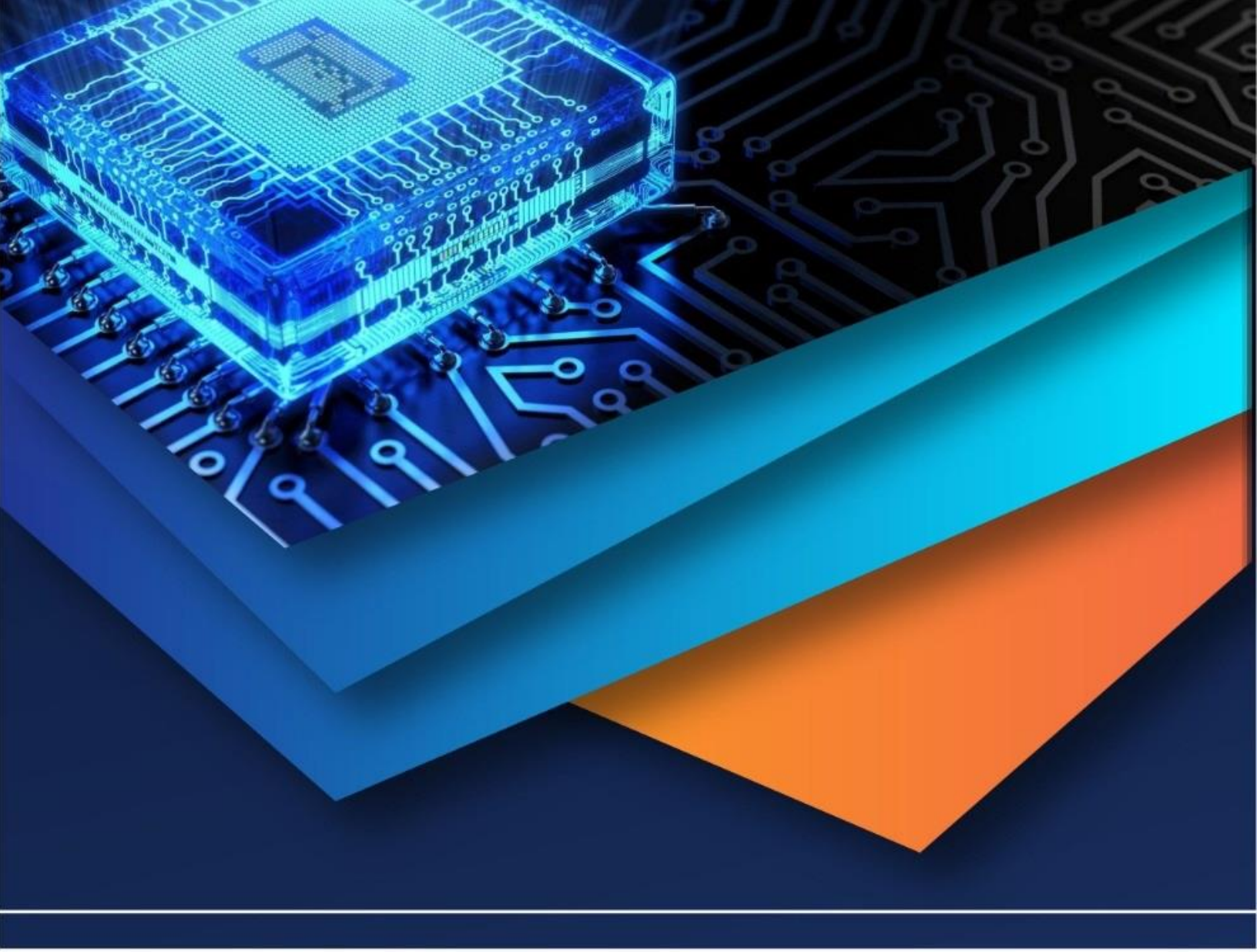
IX. ACKNOWLEDGEMENT

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.

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



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)