IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 2, August 2023

Smart Landmark Recognition Using Machine Learning

Salma Itagi, Janavi TNJ, Parthasarathi JN, Pranav Bhaskar, Tejas GS
Assistant Professor, Dept of Computer Science & Engg.[1]
8th Semester UG Students, Dept of Computer Science & Engg.[2,3,4,5]
Sai Vidya Institute of Technology, Bengaluru, India

Abstract*: Ease of tourism and travel has lead to humans capturing digital images of these landmarks and monuments. And recognition and classification of monuments in images is a challenging task due to the vast variations in their architecture and orientation. In recent years, machine learning approaches, particularly deep learning-based methods, have shown promising results in the recognition and classification of monuments. Through the aid of internet this identification process can be massively simplified. The proposed framework leverages a sophisticated processing approach that utilizes various features such as shapes, textures, and light interaction to classify different landmarks. Through a cross-platform website to provide identification and additional historical and cultural information on these sites. Our proposed system has significant implications for the tourism industry by providing a reliable and efficient means of identifying landmarks professionals who specialize in studying them are typically responsible for informing visitors about their architectural and historical significance. However, inaccurate information may be shared, and it may be difficult to present the details of these monuments in an engaging and appealing way by a tour guide. So a chatbot can provide an interactive means of communication*

Keywords: Image Processing, Machine Learning, CNN, Inception v3, Web-Based

I. INTRODUCTION

Vacations are a fundamental component of human existence, providing an opportunity to unwind and recharge from the daily stresses of work and life. During vacations, people often venture to new destinations, immerse themselves in diverse cultures, and capture memories through the art of photography. Despite the ubiquity of photographic mementos, individuals frequently struggle to recall the names of prominent landmarks or locations they visited. As human memory is inherently fallible and prone to decay over time, individuals may mis-remember or entirely forget the names of the landmarks they captured in their vacation photos.

To address this common conundrum, a landmark recognition tool can be an invaluable solution. This landmark recognition tool can be constructed using advanced machine learning models and supporting technologies. However, this tool does present certain challenges, including the limited dataset of landmark images available for training the machine learning model. Auto-annotating landmarks in images can be a challenging task as landmarks can possess various shapes and sizes, requiring a high degree of accuracy in the annotations. Furthermore, landmarks can display significant variations in appearance due to distinct lighting conditions, occlusions, environmental changes, and differences in scale and viewpoint. The computational complexity of landmark detection can be overwhelming, necessitating the use of high-end hardware and specialized software. Additionally, as training data may be limited, the model may overfit and exhibit poor performance on unseen data.

One of the primary issues with machine learning models in identifying landmarks is their inability to generalize well to new landmarks or landmarks in different environments. Lighting conditions can also affect landmark recognition, making it challenging to identify landmarks accurately. Moreover, landmarks can be partially or entirely obscured by objects in the foreground or background, rendering them difficult to recognize. Furthermore, some landmarks may appear similar to each other, making it difficult to distinguish between them. Finally, the appearance of a landmark may vary significantly depending on the viewpoint from which it is observed.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 2, August 2023

Nonetheless, overcoming these challenges can yield a highly robust tool for identifying landmarks and famous places from user-provided images, obviating the need to recall the name of the place from memory. The Landmark Recognition tool works by analyzing patterns and hidden classifiers in vast amounts of images to create a model for identifying landmarks through the iterative process of learning and improvement.

In summary, the Landmark Recognition tool offers an innovative application of neural networks in image analysis, addressing the pervasive issue of forgetting the names of land-marks. This paper aims to contribute to the field of computer vision and machine learning, showcasing the practical utility of these technologies in real-life applications.

In our approach we are using the Inception v3 model which uses transfer learning instead of training the model from scratch. It leverages the knowledge learned by a pre-trained model on a large and complex dataset to improve the performance on a smaller, dataset as our landmark dataset. By starting with a pre-trained model and fine-tuning it on the new dataset, transfer learning can help improve the accuracy and speed of training while reducing the amount of labeled data required.

The Inception v3 architecture is composed of multiple convolutional layers, which are designed to extract features from images at different levels of abstraction. The convolutional layers use small filters to perform convolutions over the input image, which helps the network learn local patterns and relationships between pixels.

During training, the dataset is divided into three sets: training, validation, and testing. The majority of the dataset is used for training, and images are entered into the network to revise the weights of the model. A portion of the dataset is kept apart for validation to prevent overfitting, and the remaining portion is used for testing to evaluate the performance of the model.

The first stage of the Inception v3 model involves computing bottleneck values for each image in the dataset. The final layer performs classification, while the layer preceding it calculates bottleneck values. This layer is trained to produce a set of values that can differentiate between different classes of images.

After generating the bottleneck values, the uppermost layer of the network is trained. The training process produces outputs for both training and validation, as well as a cross-entropy loss parameter that measures the model's learning ability. The goal is to minimize the loss, which indicates whether the model is learning correctly or not.

II. LITERATURE SURVEY

The paper by S. Hesham et al. [1] describes the implementation of a vision system to classify monuments in images using traditional and non-traditional machine learning techniques, such as deep learning. The authors used Resnet50 and VGG16 as deep learning algorithms, and KNN as the machine learning algorithm. To validate their approach, the authors surveyed 141 people to investigate their views on the proposed system. The survey results showed that most of the respondents had problems with tour guides and preferred using an application instead. 93.5% of respondents believed that using an application would be easier than relying on a human guide. Overall, the proposed system to classify monuments in images using machine learning techniques has the potential to improve tourists' experiences and help them get the most out of their tours.

The paper by Aradhya Saini et al. [2] presents a framework that relies mostly on Deep Convolutional Neural Networks (DCNN) for recognizing monument images. The use of DCNN has enabled the framework to achieve a much better accuracy of 92.7% compared to the hand-crafted features. The experiments conducted in the paper have shown the importance of using representations of monument images to build an effective recognition system. The framework can be further improved by incorporating the Graph Based Visual Saliency (GBVS) method as a preprocessing step to find the saliency of an image.

The proposed approach by Jaimala Jha and Sarita Singh Bhaduaria [3] focuses on improving the accuracy of monument image retrieval by reducing the semantic gap of feature vectors. The approach uses handcrafted features extraction techniques that generate a small size feature. The proposed approach achieves an accuracy of 97.11% in terms of F-scores. After applying the k-means clustering algorithm, the monuments image cluster is produced as output with cluster indices. The cluster indices give the cluster numbers for each image, indicating that each image belongs to a particular cluster. Overall, the proposed approach obtains better outcomes than other recent works in this area.

The proposed framework by Mehdi Etaati, Babak Majidi and Mohammad Taghi Manzuri [4] experimental results show that the deep neural network can extract high-level visual features and identify landmarks. The VGGNet model is used as the deep neural network architecture. The paper investigates the application of Scale Invariant Feature Transform (SIFT), Speeded Robust Feature (SURF), and Oriented FAST and Rotated BRIEF (ORB) for monument recognition but finds their accuracy lacking. The SVM and random forest classifiers achieve an accuracy of 85% and 95%, respectively, for landmark recognition. The proposed framework is evaluated on the tourist attractions of Iran, and the experimental results show that the proposed system can recognize the historical landmarks with a precision of 95%.

The proposed system by Neha Himesh et al. [5] based on Convolutional Neural Networks (CNN) for detecting and classifying different types of monuments. The system includes several modules such as creating a customized dataset, using a neural network with 96 convolutional layers and a pool size of 2x2, training the network with different learning rates and iterations, and testing the model on images. The loss rate and accuracy of the model with respect to the number of epochs increase. Overall, the proposed system has the potential to accurately detect and classify different types of monuments using CNNs with an accuracy of about 90%.

III. PROPOSED SYSTEM

This proposed methodology focuses on improving image classification performance by leveraging pre-trained Inception v3 models and TensorFlow. The approach involves extracting

bottleneck features from images using a pre-trained Inception v3 model, which are then used to classify the images into different categories. The system uses a combination of functions to create and read bottleneck files for images, thereby improving the retraining process.

To further enhance the performance of the image classification model, the system includes functions for adding image distortions during training. This involves creating a network of operations to apply different image distortions such as cropping, scaling, flipping, and changing brightness to the training images. Random images for the requested category are chosen, and the distortion graph and full graph are run to get the bottleneck results for each.

The system also includes functions for evaluating the ac-curacy of the model's predictions by comparing them with the ground truth labels. This involves inserting the necessary operations to calculate the cross-entropy loss function and optimizer for training, as well as the operations required to evaluate the accuracy of the model's predictions.

To enable visualization of the training process, the system uses the variable summaries function to attach summaries to a Tensor, tracking the mean, standard deviation, minimum, maximum, and histogram of a given variable.

Overall, this proposed methodology provides a robust and efficient approach to image classification using pre-trained models and TensorFlow. It enables the creation and reading of bottleneck files, as well as the application of image distortions during training, thereby improving the accuracy of the model's predictions.

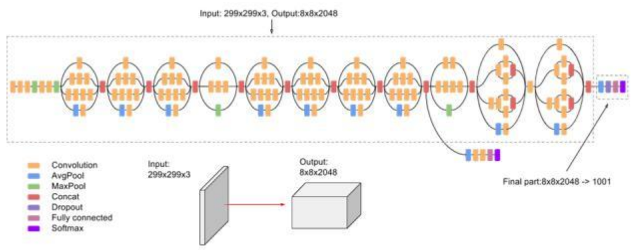

Fig. 1. Inception v3 Model

IV. METHODOLOGY

Front-End Framework:

Website

The development of a user-friendly graphical user interface (GUI) has been achieved through the utilization of HTML and JavaScript. This interface includes a "choose file" button, which upon activation prompts the system file explorer to enable the user to navigate and select their desired image file. And also a "Process image" button to upload and identify the landmark in the image.

It dispalys the image that was uploaded and finally gives the name of the landmark along with the accuracy of the prediction for this landmark

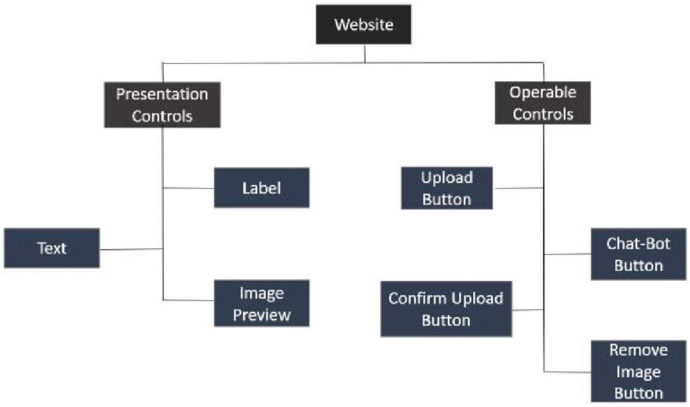

Fig. 2. Framework of Website

Back-End Technologies:
Using Firebase:
Store the image uploaded by user.
Send the image to the processing server.
Receive the landmark name from processing server.
Send the landmark name to the website to be displayed to the user.
Machine Learning:
Inception-v3
The proposed system is executed by following the subsequent steps:
Step 1: The user opens the landmark detection website on their mobile phone or computer or laptop.
Step 2: The user selects the picture which they want to identify by click on the ”choose file” button.
Step 3: The system file browser opens up where the user can select the landmark image that they want to identify.
Step 4: The user selects the image that they want to identify. Step 5: It shows the name of the image that the user has selected to be identified, if the user needs a different image to be recognized, they can click the ”choose file” button to select a different image.
Step 6: The user needs to click the ”Process image” button to upload the image to be identified.
Step 7: The image will be downloaded in the processing server and the landmark is identified.
Step 8: The name of the landmark along with the accuracy of the landmark is displayed on the website.
Step 9: Along with the name of the landmark, additional information about the landmark will be displayed on the website along with a text-to-speech option to read out this information.

V. RESULT AND ANALYSIS

Analysis of the Smart Landmark Recognition application is as shown below

Fig. 3. Main Page for Smart Landmark Recognition

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 2, August 2023

The main page of the Smart Landmark Recognition system that is loaded up when the user loads it up on their mobile phone / laptop / computer.

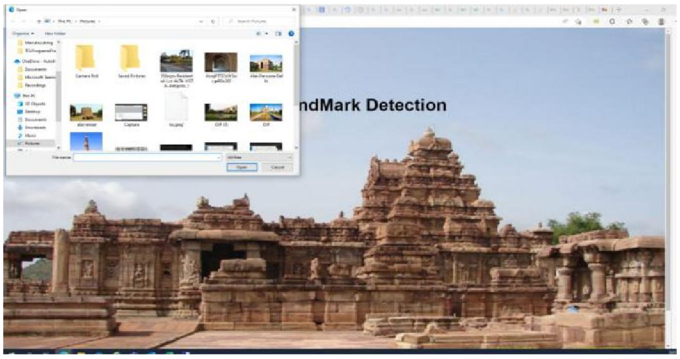

Fig. 4. File Browser to select image

To select the landmark image to be identified, the user can click on the "choose file" button, which will open the file system browser, enabling them to select the desired picture to be uploaded with ease.

Fig. 5. Displays the name of image

To ensure the selection of the correct image by the end user for the landmark processing and identification, the website displays the name of the image that was chosen by them in the previous step

Fig. 6. During the upload and processing part

The website provides visual indicators during the image up-loading and identification process, keeping the user informed of the ongoing background processing.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 2, August 2023

Fig. 7. After the successful identification of the landmark

The Smart Landmark Recognition system displays the name of the landmark along with the corresponding accuracy of recognition of the landmark

VI. CONCLUSION

In summary, our web-based application employing an image processing algorithm, classification algorithm, and recommender system has accomplished a commendable degree of precision in recognizing landmarks from photographs. This innovative solution effectively bridges the existing gap in landmark detection and recognition, rendering a valuable tool for travelers to carry on their mobile devices while traversing new territories. The exceedingly swift processing and analysis time, coupled with cross-platform compatibility, render it a highly efficient and feasible option for real-world applications. In essence, our application possesses the potential to enrich the travel experience for users by providing them with reliable and accurate landmark identification.

VII. FUTURE SCOPE

The future scope of our landmark recognition project includes various enhancements and features aimed at improving the accuracy and usability of the application. These include expanding the range of landmarks that can be recognized and improving the accuracy of the recognition algorithm beyond the 95%, incorporating natural language processing to enhance the chatbot's ability to provide information about the recognized landmark and related landmarks.

Incorporation of augmented reality (AR) technology to provide a more immersive experience for users, such as overlaying information about the landmark onto the camera feed in real-time or providing a 3D model of the landmark that can be explored.

Other enhancements include incorporating user-generated content to enhance the application's database and increase the variety of landmarks that can be recognized, implementing a feedback mechanism that allows users to provide feedback on the accuracy of the landmark recognition and suggest corrections for any errors, enhancing the recommendation system by incorporating user preferences, location data, and social media data to provide more personalized and relevant recommendations, implementing multilingual support to enable the application to cater to a broader audience and enhance its usability for travelers from different countries, integrating with social media platforms to enable users to share their travel experiences and landmarks with their friends and followers.

Finally even partnering with travel agencies and tour opera-tors to help promote the application and enhance its visibility among potential users can be a considered as well.

These enhancements and features will allow our landmark recognition project to provide a more comprehensive and engaging experience for users, as well as enable it to cater to a broader audience and improve its accuracy and usability over time.

REFERENCES

[1] S. Hesham, R. Khaled, D. Yasser, S. Refaat, N. Shorim and F. H. Ismail, "Monuments Recognition using Deep Learning VS Machine Learning," 2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC), NV, USA, 2021, pp. 0258-0263, doi: 10.1109/CCWC51732.2021.9376029.

[2] JA. Saini, T. Gupta, R. Kumar, A. K. Gupta, M. Panwar and A. Mittal, ”Image based Indian monument recognition using convoluted neural networks,” 2017 International Conference on Big Data, IoT and Data Science (BID), Pune, India, 2017, pp. 138-142, doi: 10.1109/BID.2017.8336587.
[3] Jha, Jaimala. (2020). A Novel approach for Retrieval of Historical Monuments Images using Visual Contents and Unsupervised Machine Learning. International Journal of Advanced Trends in Computer Science and Engineering. 9. 3563-3569. 10.30534/ijatcse/2020/162932020.
[4] M. Etaati, B. Majidi and M. T. Manzuri, ”Cross Platform Web-based Smart Tourism Using Deep Monument Mining,” 2019 4th International Conference on Pattern Recognition and Image Analysis (IPRIA), Tehran, Iran, 2019, pp. 190-194, doi: 10.1109/PRIA.2019.8785975.
[5] ”Indian Monument Recognition using CNN Algorithm”, International Journal of Emerging Technologies and Innovative Research (www.jetir.org), ISSN:2349-5162, Vol.6, Issue 5, page no.286-289, June-2019
[6] C. Termritthikun and P. Muneesawang, “NU-LiteNet: Mobile Landmark Recognition using Convolutional Neural Networks: Mobile Landmark Recognition using Convolutional Neural Networks”, ECTI-CIT Transactions, vol. 13, no. 1, pp. 21–28, Oct. 2019.
[7] A. J. Paul, S. Ghose, K. Aggarwal, N. Nethaji, S. Pal and A. Dutta Purkayastha, ”Machine Learning Advances aiding Recognition and Classification of Indian Monuments and Landmarks,” 2021 IEEE 8th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON), Dehradun, India, 2021, pp. 1-8, doi: 10.1109/UPCON52273.2021.9667619.
[8] G. Amato, F. Falchi and P. Bolettieri, ”Recognizing Landmarks Using Automated Classification Techniques: Evaluation of Various Visual Features,” 2010 Second International Conferences on Advances in Multimedia, Athens, Greece, 2010, pp. 78-83, doi: 10.1109/MMEDIA.2010.20.
[9] M. S. Bhatt and T. P. Patalia, ”Genetic programming evolved spatial descriptor for Indian monuments classification,” 2015 IEEE International Conference on Computer Graphics, Vision and Information Security (CGVIS), Bhubaneswar, India, 2015, pp. 131-136, doi: 10.1109/CGVIS.2015.7449908.
[10] P. Desai, J. Pujari, N. H. Ayachit and V. K. Prasad, ”Classification of archaeological monuments for different art forms with an application to CBIR,” 2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Mysore, India, 2013, pp. 1108-1112, doi: 10.1109/ICACCI.2013.6637332.
[11] B. Ghildiyal, A. Singh and H. S. Bhadauria, ”Image-based monument classification using bag-of-word architecture,” 2017 3rd International Conference on Advances in Computing,Communication & Automation (ICACCA) (Fall), Dehradun, India, 2017, pp. 1-5, doi: 10.1109/ICACCAF.2017.8344728.
[12] Palma, Valerio. “TOWARDS DEEP LEARNING FOR ARCHITECTURE: A MONUMENT RECOGNITION MOBILE APP.” The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences (2019): n. pag.
[13] P. Shukla, B. Rautela and A. Mittal, ”A Computer Vision Framework for Automatic Description of Indian Monuments,” 2017 13th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), Jaipur, India, 2017, pp. 116-122, doi: 10.1109/SITIS.2017.29.
[14] S. Gada, V. Mehta, K. Kanchan, C. Jain and P. Raut, ”Monument Recognition Using Deep Neural Networks,” 2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Coimbatore, India, 2017, pp. 1-6, doi: 10.1109/ICCIC.2017.8524224.
[15] T. Chen and K. -H. Yap, ”Discriminative BoW Framework for Mobile Landmark Recognition,” in IEEE Transactions on Cybernetics, vol. 44, no. 5, pp. 695-706, May 2014, doi: 10.1109/TCYB.2013.2267015.
[16] A. Del Bimbo, W. Nunziati and P. Pala, ”David: Discriminant analysis for verification of monuments in image data,” 2009 IEEE International Conference on Multimedia and Expo, New York, NY, USA, 2009, pp. 334-337, doi: 10.1109/ICME.2009.5202502.
[17] H. H. Alalawi, M. S. Alsuwat, S. M. Alhazmi, R. A. Bagido and H. S. Alshanbari, ”Makkah Landmark Recognition Using Deep Learning,” 2022 2nd International Conference on Computing and Information Technology (ICCIT), Tabuk, Saudi Arabia, 2022, pp. 260-264, doi: 10.1109/ICCIT52419.2022.9711612.

[18] T. Chen, K. -H. Yap and D. Zhang, ”Discriminative Soft Bag-of-Visual Phrase for Mobile Landmark Recognition,” in IEEE Transactions on Multimedia, vol. 16, no. 3, pp. 612-622, April 2014, doi: 10.1109/TMM.2014.2301978.

[19] Y. Kong, W. Liu and Z. Chen, ”Robust ConvNet Landmark-Based Visual Place Recognition by Optimizing Landmark Matching,” in IEEE Access, vol. 7, pp. 30754-30767, 2019, doi: 10.1109/ACCESS.2019.2901984.

[20] T. Chen, K. Wu, K. -H. Yap, Z. Li and F. S. Tsai, ”A Survey on Mobile Landmark Recognition for Information Retrieval,” 2009 Tenth International Conference on Mobile Data Management: Systems,