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: 7.53
Volume 4, Issue 4, March 2024

Heritage Harbor with YOLOv8:Preserving Monuments through Audio Tales

[1]Ms. Sophiya Sugantha Grace M, [2]Mr. Abishek C, [3]Mr. Harish Kanna P, [4]Mr. Varun G
[1]Assistant Professor, Department of Computer Science and Engineering
[2,3,4]Students, Department of Computer Science and Engineering
SRM Valliammai Engineering College, Chennai, Tamil Nadu, India

Abstract: *Heritage Harbor with YOLOv8" is a pioneering project aimed at conserving cultural heritage through a fusion of modern technology and traditional storytelling. Leveraging the advanced object detection capabilities of YOLOv8, this initiative focuses on the preservation of monuments by capturing their essence through audio narratives. By integrating audio tales with real-time object recognition, visitors are immersed in a rich, interactive experience that fosters a deeper appreciation for historical sites.This abstract outlines the methodology and objectives of the project, emphasizing the synergy between technology and cultural preservation. Through a combination of machine learning algorithms and curated storytelling, the project aims to revitalize interest in heritage sites while safeguarding them for future generations. Furthermore, it explores the potential for scalability and adaptability, envisioning a future where similar approaches can be applied to diverse cultural landscapes worldwide.*

Keywords: Heritage preservation, YOLOv8, object detection, audio storytelling, cultural conservation

I. INTRODUCTION

Vision based classification of objects has gained immense popularity in the past few years and is being used for a multitude of tasks ranging from face recognition to vehicle detection. While in the earlier days, this problem would have been solved by measuring the similarities between features of two images and classifying images with most matching features as belonging to the same class, automatic classification has now become possible, thanks to the sophisticated **machine** learning algorithms in use today. In this project, certain supervised machine learning classification techniques were applied to the exciting task of classification of different monument images based on their architecture. India, a country which has seen the rise and fall of several dynasties throughout the ages, has a rich architectural and cultural inheritance to behold. The history and heritage of every kingdom is greatly preserved in the monuments and buildings that were built during their reign. Decades later, while those dynasties no longer continue to exist, these preserved relics still flaunt the grandeur of those kingdoms through the uniqueness of their architecture. What is worth noting is that these architectural styles, though unique in themselves, are often inspired by others, which makes determining the architecture of a particular monument through visual observation a rather challenging task. In this project, an attempt to determine how well machine learning algorithms perform in such a situation was made. Local features have been found to be useful for learning the keypoints of an image in the past, and so the ORB local features were used in this project.

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: 7.53
Volume 4, Issue 4, March 2024

II. METHODOLOGY

The methodology for the proposed Detection of Gastrointestinal Lesion system would involve the following steps:

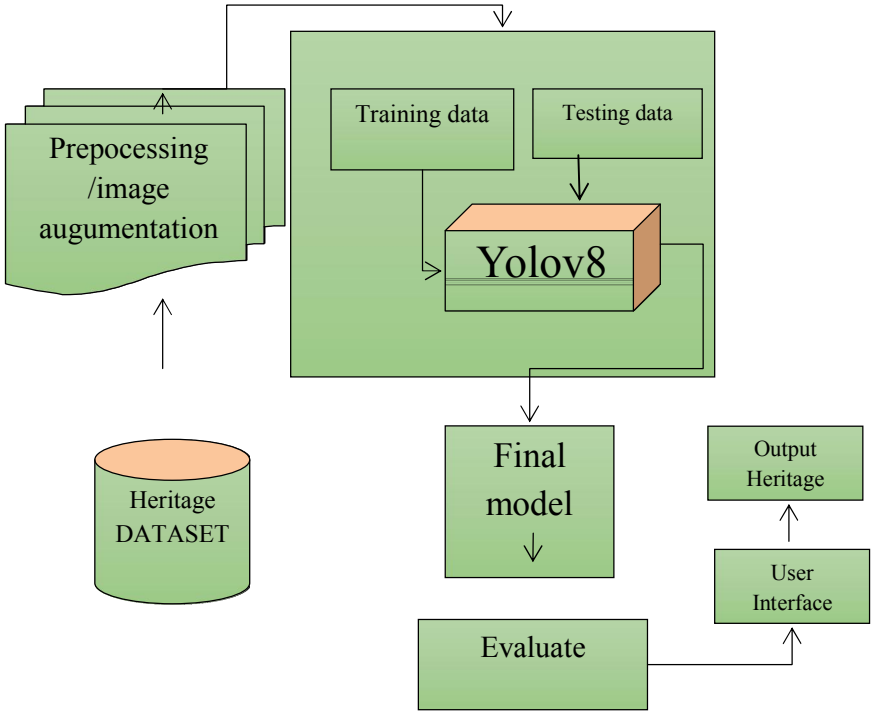


Fig. 1. Architecture diagram Heritage harbor with yolov8:preserving monumets through audio tales

III. PROPOSED SYSTEM

The proposed system is designed as an end-to-end solution for the precise detection and of monuments in various scenarios and conditions using an AI-based model. The key components of the system include YOLOv8 for object detection, Roboflow for data annotation, and Weights and Biases (Wandb) for experiment tracking. Additionally, a graphical user interface (GUI) based on Tkinter is integrated for real-time predictions and enhanced usability.

Data Collection and Annotation:

- Gather a diverse dataset of images containing different monuments from various angles, lighting conditions, and perspectives.
- Ensure that the dataset represents the real-world scenarios the model will encounter.
- Annotate the collected images with bounding boxes around the monuments. Each bounding box should include information about the class label (e.g., the type of monument) and possibly additional attributes.

Data Preprocessing:

- The dataset is then preprocessed, which involves tasks such as resizing, normalization, and augmentation to enhance the model's ability to generalize to different scenarios.
- The dataset is organized into training, validation, and test sets to train and evaluate the model effectively.

Model Architecture - YOLOv8: training

- Use the annotated and preprocessed dataset to train the YOLOv8 model.
- Split the dataset into training and validation sets to monitor the model's performance during training.
- Adjust hyperparameters, such as learning rate and batch size, based on validation performance.
- YOLOv8, a state-of-the-art object detection model, is chosen for its efficiency and accuracy.

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: 7.53
Volume 4, Issue 4, March 2024

- The YOLOv8 variant is fine-tuned to adapt specifically to the task of monument
- Various hyperparameters and data augmentation techniques are experimented with to optimize the model's performance.

YOLOv8 is the newest state-of-the-art YOLO model that can be used for object detection, image classification, and instance segmentation tasks. YOLOv8 was developed by Ultralytics, who also created the influential and industry-defining YOLOv5 model. YOLOv8 includes numerous architectural and developer experience changes and improvements over YOLOv5.

The YOLO (You Only Look Once) series of models has become famous in the computer vision world. YOLO's fame is attributable to its considerable accuracy while maintaining a small model size. YOLO models can be trained on a single GPU, which makes it accessible to a wide range of developers. Machine learning practitioners can deploy it for low cost on edge hardware or in the cloud.

YOLO has been nurtured by the computer vision community since its first launch in 2015 by Joseph Redmond. In the early days (versions 1-4), YOLO was maintained in C code in a custom deep learning framework written by Redmond called Darknet. YOLOv8 achieves strong accuracy on COCO. For example, the YOLOv8m model -- the medium model -- achieves a 50.2% mAP when measured on COCO. When evaluated against Roboflow 100, a dataset that specifically evaluates model performance on various task-specific domains, YOLOv8 scored substantially better than YOLOv5. More information on this is provided in our performance analysis later in the article.

Furthermore, the developer-convenience features in YOLOv8 are significant. As opposed to other models where tasks are split across many different Python files that you can execute, YOLOv8 comes with a CLI that makes training a model more intuitive. This is in addition to a Python package that provides a more seamless coding experience than prior models.

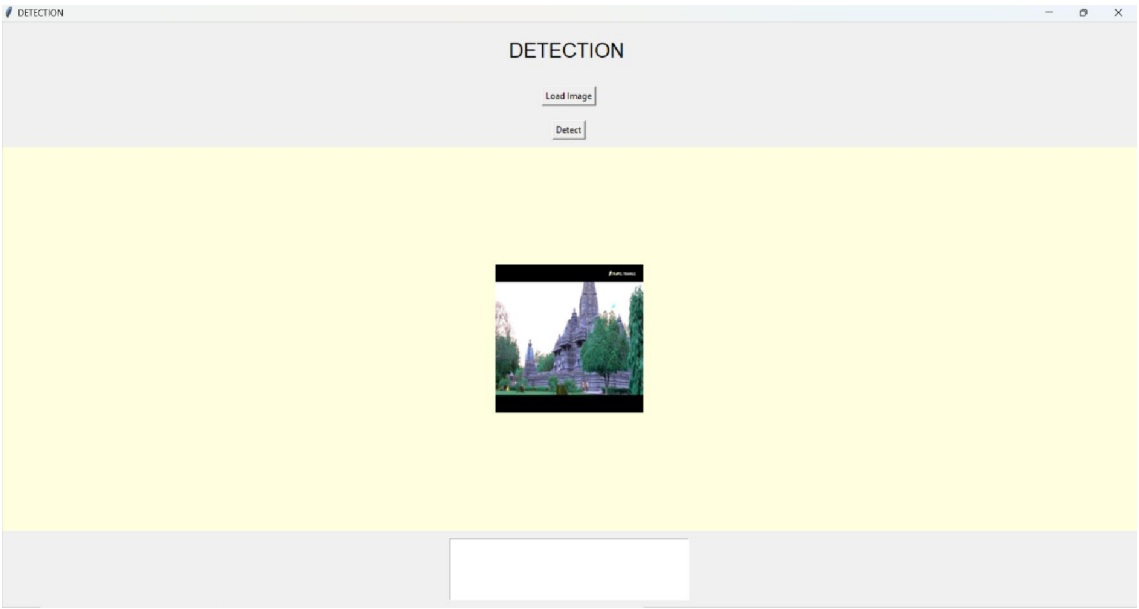
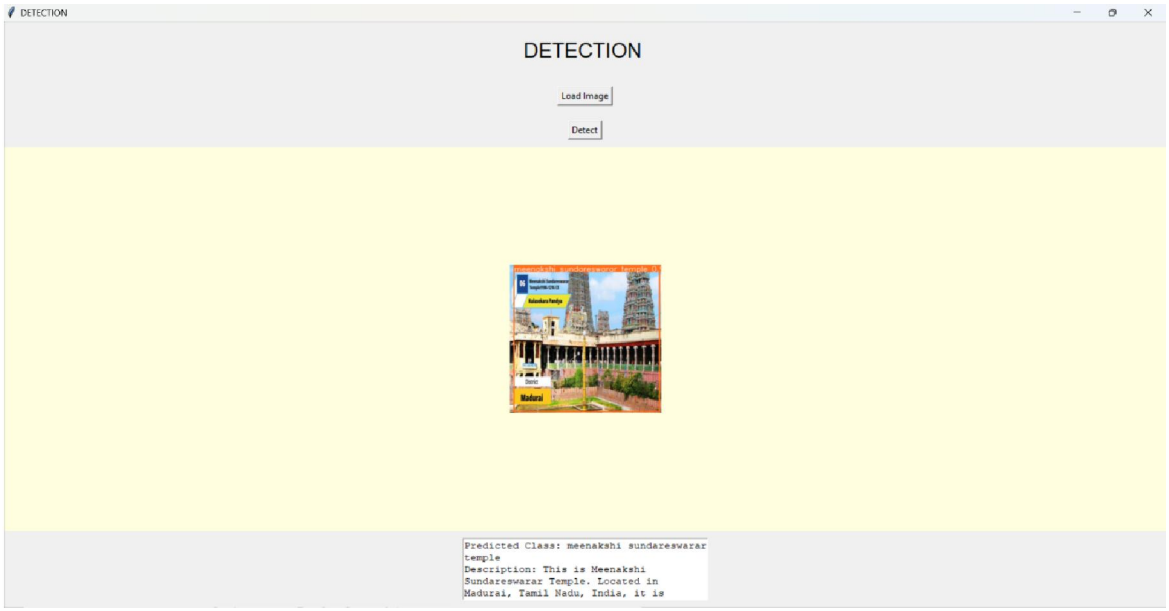


Fig. 2. The User Interface of Heritage harbor with yolov8:preserving monuments through audio tales

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: 7.53
Volume 4, Issue 4, March 2024

IV. RESULTS



Heritage Harbor with YOLOv8: Preserving Monuments Through Audio Tales" is a project that merges cutting-edge technology with cultural preservation. YOLOv8, a state-of-the-art object detection algorithm, is employed to recognize and identify monuments within Heritage Harbor. Once identified, these monuments are then linked to audio tales or narratives that delve into their historical significance, cultural context, and anecdotes associated with them. This innovative approach not only enhances the visitor experience by providing engaging audio content but also contributes to the preservation of heritage sites by fostering a deeper connection and understanding among visitors. Through this project, Heritage Harbor aims to safeguard its rich cultural heritage while embracing modern technological advancements.

V. CONCLUSION

This study introduces a groundbreaking application of AI technology, specifically YOLOv8, in the domain of monument detection. The integration of this advanced object detection model into the identification of historical landmarks represents a significant leap forward in the efficient preservation, study, and understanding of our cultural heritage. The central aim of the research is to establish a comprehensive AI-driven system capable of detecting and precisely localizing more than 10 distinct monuments within diverse and complex historical landscapes. By harnessing the state-of-the-art object recognition abilities of YOLOv8, the approach facilitates rapid identification and assessment, thereby enhancing heritage preservation and research practices. The system's performance is rigorously evaluated using advanced metric assessment techniques, providing continuous monitoring and real-time insights into detection accuracy, model behavior, and potential biases. This eliminates the need for external platforms, streamlining the process of making quick adjustments and refinements. This ensures an optimized and reliable monument detection solution. Ultimately, this research contributes significantly to the field of monument detection, heritage preservation, and historical research.

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", Proceedings of the 17th International Conference on Pattern Recognition 2004. ICPR 2004, vol. 2, pp. 1-6, 2004.

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: 7.53
Volume 4, Issue 4, March 2024

[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.

[4]. A. Torralba, K. P. Murphy, W. T. Freeman and M. A. Rubin, "Context-based vision system for place and object recognition", Radiology, vol. 239, no. 1, pp. 301, 2006.

[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.

[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.

[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.

[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.

[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.

[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.

[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.

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

[13]. J. Lim, Y. Li, Y. You and J. Chevallet, "Scene Recognition with Camera Phones for Tourist