



ENHANCING TOURIST EXPERIENCE AND ENGAGEMENT THROUGH AI-DRIVEN REAL-TIME OBJECT DETECTION AND INTERACTIVE MOBILE APPLICATION

Marko Mihajlović¹,
Angelina Njeguš^{2*}

¹Ph.D. Student,
Singidunum University,
Belgrade, Serbia

²Singidunum University,
Belgrade, Serbia

Abstract:

Artificial intelligence has the potential to revolutionize the tourism industry, particularly by enhancing tourist experiences and providing valuable insights into tourist behavior. Our research identified a significant opportunity to enhance tourist satisfaction and engagement and enable advanced visitor analytics through an AI-powered real-time object detection system. In this study, we propose an AI-driven mobile application to address the gap in the digital infrastructure in Serbia by facilitating the real-time detection of tourist points of interest (PoI), allowing users to upload images instead of performing text searches. The uploaded images are analyzed by an AI system using the YOLO model trained on a customized dataset created specifically for this purpose. When a PoI is recognized, relevant information is displayed on the user's mobile phone in their preferred language, using advanced AI translation tools. This paper outlines how such a system not only enhances the overall tourist experience by delivering instant and pertinent information, but also provides tourism organizations with advanced analytics on visitor behavior, improved customer interaction capabilities, and enhanced promotional opportunities for lesser-known attractions and local businesses. Moreover, this study explores how the proposed application aligns directly with the tourism development strategy of Serbia and the UNDP recommendations, aiming to enhance digital infrastructure and support sustainable tourism initiatives across the country.

Keywords:

AI-driven Tourism, Real-time object detection, Visitor Behavior Analytics, Interactive Mobile Application.

INTRODUCTION

Tourists often find themselves at attractions with limited knowledge and without the means to easily perform detailed text searches to gain more information. Additionally, tourist organizations frequently lack robust tools for effectively tracking tourist behaviors and routes. Real-time information exchange, facilitated by information and communication technologies, is crucial for tourists to make informed decisions, and for service providers to offer timely and personalized services. The ability to access updated on-go information supports a seamless travel experience from planning and booking to navigating destinations (Verhun, 2022). However, the lack of accessible information can lead to various challenges. Tourists may face difficulties such as language barriers and navigation problems, which can result in a less satisfactory

Correspondence:
Angelina Njeguš

e-mail:
anjegus@singidunum.ac.rs



travel experience (Kumar *et al.*, 2021). For tourist organizations, the lack of real-time data can hinder their ability to understand and respond to tourists' preferences and movement patterns, potentially leading to inefficient resource allocation and missed opportunities for targeted marketing (Lee *et al.*, 2019). The significance of real-time information in tourism is underscored by the role of digital technologies in enhancing travel efficiency and experience (Gupta *et al.*, 2024).

Therefore, this paper presents a new proposal for tackling these challenges using sophisticated artificial intelligence technologies. The core of this solution revolves around the adoption of the YOLO model, an advanced convolutional neural network recognized for its effectiveness in analyzing images and detecting objects in real time. We propose the development of a mobile application based on the YOLO model that leverages its capabilities to provide tourists with instant recognition of points of interest. This application enables users to receive immediate and pertinent information about the landmarks they encounter, thereby enhancing their travel experience through real-time insights and contextual data. In this study, we created a specialized dataset specifically tailored for tourist attractions. This dataset serves as a proof-of-concept for our application. Tourists use their mobile phones to capture pictures of the environment, which are then used as input to identify various tourist attractions within the visual frame.

Tourism organizations and agencies can utilize this application as an effective tool for acquiring a comprehensive knowledge of tourist behavior. By examining the data derived from user engagement with the application, these entities can enhance their understanding of visitor patterns, inclinations, and attractions, leading to informed strategic decision-making and efficient resource distribution. The capacity to monitor and assess tourist itineraries and activities immediately enables a more adaptable management of tourism flows, potentially improving overall visitor satisfaction. Additionally, the application provides enhanced promotional opportunities for lesser-known attractions and local businesses, helping diversify tourist activities and boost local economies. Furthermore, the app offers educational opportunities by providing rich contextual information about destinations, which can deepen visitors' understanding and appreciation of cultural, historical, and natural sites, thus enriching the overall travel experience.

This study advances modern tourism strategies by evaluating YOLO models for real-time detection of tourist points of interest (PoI), demonstrating the selected model's high accuracy and practical application. We developed a specialized dataset to enhance the model's adaptability in recognizing emerging PoIs. The resulting mobile app uses this model to identify tourist spots and deliver information in the user's preferred language, while also collecting visitor behavior analytics. These insights can drive personalized tourism experiences and strategic development, particularly in enhancing tourism strategies in Serbia to meet global tourism demands.

The remainder of this paper is organized as follows. Related Work offers a review of pertinent studies that have served as the foundation for the present research. The methods section presents the proposed methodology, including assessments of YOLO models and custom dataset development. In the Discussion section, we evaluate the effectiveness of the selected YOLO model, trained on our proof-of-concept dataset, through experimental results, and assess its accuracy and usability. We also explored potential improvements and obstacles to implementation, in addition to identifying areas for future research. Finally, the Conclusion summarizes the key findings and contributions of our research.

RELATED WORK

Existing real-time solutions and technologies in tourism-related AI applications for tourist spot detection are multifaceted and incorporate various AI systems and applications. AI has been instrumental in enhancing the experience of travelers by facilitating the identification of destination points (Jamaluddin & Rahmat, 2023). These technologies include personalization and recommender systems that leverage big data and machine learning algorithms to suggest tourist spots to travellers based on their preferences (Bulchand-Gidumal, 2022). In addition, AI-powered search and booking engines, which rank high in adoption susceptibility, are pivotal in helping tourists discover and plan visits to attractions (Huang *et al.*, 2022). The recent literature indicates a growing interest in the development of applications that assist in the real-time identification of tourist points of interest. These applications leverage various technological advancements, including Ambient Geographic Information, deep learning, and Mobile GIS technology, to enhance the experience of tourists by providing them with immediate and relevant information about attractions (Celardo *et al.*, 2023) (Zhang *et al.*, 2021) (Dai, 2013).



Interestingly, while some approaches focus on analyzing user-generated content to inform tourists and industry stakeholders (Celardo *et al.*, 2023), others emphasize the use of image recognition algorithms to identify locations and personalize recommendations (Chen *et al.*, 2021) (Zhang *et al.*, 2021). Additionally, the application of these technologies extends beyond tourism, as demonstrated by the use of RGB-D sensors for natural landmark recognition in mobile robotics (Souto *et al.*, 2017). Moreover, the integration of conversational recommender systems into tourism reflects an innovative direction for facilitating user interaction and preference determination (Auliarahman *et al.*, 2021). For instance, the integration of artificial intelligence (AI) in systems within museums allows for the analysis of visitor behavior, providing data that can help optimize exhibition layouts and monitor attendance (Dyk *et al.*, 2023). In the context of smart cities and urban development, GeoAI can process large volumes of geospatial data, potentially including visitor movements, to optimize urban planning and management (Hosen *et al.*, 2023).

Real-time object detection using AI models includes methods such as the Single Shot MultiBox Detector (SSD) and Region-based Convolutional Neural Networks (Faster R-CNN). These models are known for their accuracy and use of anchor boxes, feature pyramid networks, and non-maximum suppression to achieve a balance between precision and processing speed (Alsharabi, 2023). The You Only Look Once (YOLO) algorithm, on the other hand, is recognized for its speed and ability to process the entire image simultaneously, predicting bounding boxes and class probabilities for multiple objects, which is particularly advantageous in time-sensitive applications (Lakhotiya *et al.*, 2023) (Lavanya & Pande, 2023). YOLO's ability to process images in a single evaluation makes it well-suited for applications in which immediate object recognition is crucial, such as identifying landmarks or points of interest in a tourism context. In addition, continuous improvements in YOLO architecture, such as the introduction of YOLOv5 with a feature pyramid network and anchor boxes, suggest ongoing enhancements in accuracy and speed (Lakhotiya *et al.*, 2023). The convergence of user-generated content analysis, image processing, and real-time data retrieval technologies represents a significant step in the development of applications for the real-time identification of tourist points of interest. These applications not only promise to improve the tourist experience by providing instant access to information but also offer valuable insights for the tourism industry to tailor services and marketing strategies effectively (Chary, 2023) (Lavanya & Pande, 2023) (Lee & Hwang, 2022) (Guan, 2023). The continued development and application of AI in tourism are expected to further refine and enhance the detection and promotion of tourist spots, offering travelers more tailored and efficient planning tools (Any *et al.*, 2024) (Xiong & Zhang, 2024).

Despite extensive discussions on object detection technologies and their applications across various sectors, there remains a notable deficiency in studies that specifically focus on integrating these technologies into the tourism industry, particularly for enhancing image-based search capabilities to recognize tourist points of interest. Therefore, we propose this application as a potential solution to bridge this gap by leveraging the advanced capabilities of the YOLO object detection technology. This approach aims to enhance visitor experiences and provide valuable insights into tourist behavior by enabling the real-time, accurate identification of tourist attractions. The immediate availability of contextual and relevant information directly influences tourist satisfaction by enabling more informed and enriched experiences and allowing tourists to interact with their environment in a more meaningful manner (Pawlak, 2023). Additionally, tourism businesses benefit from the ability to dynamically manage resources and respond to tourist needs in real time, which can lead to improved service delivery and increased operational efficiency (Gupta *et al.*, 2024) (Lee *et al.*, 2019) (Verhun, 2022). The absence of such real-time capabilities highlights critical gaps in current tourism infrastructure, underscoring the need for continued investment in digital technologies. Therefore, investing in systems that provide real-time information and data analysis represents a compelling opportunity for the tourism industry to evolve and adapt to the expectations of modern travelers, ensuring that destinations remain competitive and capable of delivering exceptional experiences.



METHODS

YOLO Models

The You Only Look Once (YOLO) model, originally developed using the Darknet framework, is a series of AI models designed for efficient real-time object detection. This series uses convolutional neural networks (CNNs) that enable the system to detect objects by looking only once at the image, which is a significant improvement over traditional models that require separate proposal and classification steps (Lavanya & Pande, 2023). The YOLO model operates by dividing the image into a grid and predicting the bounding boxes and the probabilities of each grid cell. A single neural network predicts multiple bounding boxes per grid cell and assigns class probabilities to these bounding boxes. This unified approach enables the model to achieve remarkable speed and accuracy, making it particularly suitable for applications that require real-time processing.

Since its inception, the YOLO model has evolved through several iterations, each improving upon the previous model in terms of accuracy, speed and capability to handle different object sizes and complexities:

- YOLOv1 to YOLOv3: These early versions pioneered real-time object detection capabilities, with YOLOv1 establishing the framework but struggling with localization errors and recall. YOLOv2 improves the precision and recall by introducing anchor boxes, thereby enhancing the detection of smaller objects. YOLOv3 further refines this approach using multiscale predictions.
- YOLOv4 made substantial gains in both speed and accuracy by incorporating new technologies such as CSP and MiSH activation, enhancing contextual understanding essential for complex scenes.
- YOLOv5, developed in PyTorch, offers the best in terms of integration ease and operational efficiency. It also optimizes the balance between speed and accuracy, making it particularly effective for real-time applications that require rapid and reliable performance.
- YOLOv6 to YOLOv8: These later versions of the YOLO model series present advanced improvements in accuracy and detection speeds. However, they pose challenges in practical implementation, owing to their increased complexity and higher computational demands.

After evaluating various models for object detection and conducting experiments with different YOLO models, YOLOv5, specifically the YOLOv5m variant, emerged as the most suitable for our application because of several factors. YOLOv5 provides a balance between speed and accuracy with substantial improvements in processing speed, making it suitable for applications requiring efficient real-time object detection without significant resource expenditure. It is built on a PyTorch framework, which simplifies model training and deployment compared with its predecessors. The YOLOv5 model offers various scaled versions from small to large, allowing it to be tailored according to the computational resources and accuracy requirements. Through extensive experimentation, YOLOv5m was selected because of its optimal balance of performance and efficiency. This version demonstrated a superior capability in handling real-time processing demands while maintaining high detection accuracy, making it the preferred choice for our project objectives. This adaptability makes it ideal for server-side processing in which images uploaded by tourists can be analyzed efficiently without overloading mobile devices. PyTorch makes the codebase more accessible and easier to modify for specific needs, which is critical for adapting the model to recognize diverse tourist attractions across different environments.

Another advantage of YOLOv5 is its ability to effectively apply learning from natural image data across various domains. It can be trained to identify various categories, making it adaptable to diverse tourist attractions including natural landmarks, sculptures, and historical buildings. However, running advanced AI models typically requires significant computational resources, which can be costly, especially when scaled to handle multiple simultaneous users across various geographic locations. The computational cost increases further when the application needs to process high-resolution images in real-time, which is often necessary for accurate object detection in diverse environments. Furthermore, keeping the AI model effective and relevant requires continuous updates. This could involve retraining the model periodically with new data, a process that can be resource-intensive. Implementing transfer learning strategies can help alleviate some of these challenges



by allowing the model to adapt better to new types of data or environments. However, it is important to manage the process carefully to prevent any degradation in the performance of the model in previously learned scenarios.

The experiments were conducted on Google Colab utilizing its robust cloud-based platform to ensure consistent computational performance and accessibility. The Colab notebook, which provides a comprehensive environment for running and experimenting with the code, details the setup and execution of our experiments, can be accessed through the following link: https://colab.research.google.com/drive/1pYpC648QIfuA2bPoK98u_ndPiQtzQWl-. This notebook includes all necessary dependencies, pre-configured settings and step-by-step instructions to help you explore and replicate the findings of this study. The results and further analysis of these experiments are discussed in the Discussion section of our paper. This setup allowed us to systematically evaluate the performance and applicability of the YOLOv5m model under controlled conditions by leveraging Google Colab's powerful GPUs to effectively simulate real-world usage scenarios.

Custom Dataset

Creating a high-quality dataset is essential for effectively training the YOLO model to identify the tourist points of interest. This section discusses the process of creating such a dataset, which involves data collection, annotation and preprocessing. Tools, such as RoboFlow, can be used to streamline these processes. The initial step in creating a dataset for identifying tourist PoIs is to gather a diverse range of images depicting the various PoIs in which the application is to be used. This process involves ensuring geographic diversity by including images of PoIs from different locations, which helps to ensure the effectiveness of the model across diverse regions. It is also important to collect images captured at different times of the day and under varied weather conditions because this variety makes the model resistant to changes in lighting and weather. Finally, it is essential for the dataset to contain images captured from various angles and distances from each PoI, allowing the model to recognize them from any standpoint. Images can be obtained from open datasets available online, partnerships with tourism boards or agencies, crowdsourced contributions by volunteers, or by hiring professional photographers specifically tasked with capturing distinct PoIs.

Once the images have been obtained, the subsequent task involves adding descriptive labels, which help to train the YOLO model. RoboFlow offers a user-friendly platform for annotating images and uses various tools to streamline the process. Each image was labelled by creating bounding boxes around the points of interest. It is essential to ensure that these boxes tightly enclose the object to enhance model accuracy. Furthermore, each bounding box is assigned a label corresponding to the PoI, and it is crucial to maintain consistency in the labelling, particularly when similar PoIs appear in different images. Before training the YOLO model, it is essential to preprocess the images to improve the efficiency and performance of the learning process. In our experiments, we implemented specific preprocessing steps to enhance the efficiency and accuracy of the YOLO model. Images were first auto-oriented and resized to 640×640 pixels to reduce the file size and quicken training. The contrast was auto-adjusted using Contrast Stretching, which optimized the pixel intensity for better edge detection. To further enhance the diversity of the dataset and improve model robustness, specific image augmentations were applied during preprocessing. Horizontal flipping was used to make the model insensitive to the subject orientation. The brightness was adjusted by 20% to make the model more resilient to variations in lighting and camera settings, which helps maintain the detection accuracy under different environmental conditions. Additionally, we applied 20% hue augmentation by randomly adjusting the colours in each image. This augmentation makes the model less sensitive to colour variations, enhancing its robustness and reducing its sensitivity to minor discrepancies in image capture conditions.

The dataset used in our experiments was created using Roboflow tools, an advanced platform designed for building and optimizing machine-learning datasets. This tailored dataset is publicly available to facilitate further research and development in the field. It can be accessed via the following URL address: <https://universe.roboflow.com/serbia-tourist-points-of-interest/tourist-points-of-interest/dataset/5>. This resource includes all preprocessing and augmentation configurations applied, ensuring that others can replicate our setup or extend the dataset with additional annotations and modifications.



Real-Time Object Detection

The first step involves resizing the captured image to meet the input requirements of the YOLO model. Once processed, the YOLO model was used to analyze the PoI images using a trained neural network to compute their locations and assess probabilities. When a PoI is identified, the relevant information is retrieved from the backend database or API. This database was populated with data during the initial dataset creation and could be updated periodically with new findings or corrections. The final stage includes presenting the retrieved information on the user's device, which may consist of a PoI name, brief description, historical facts, and links to more detailed content in a format that is easy to understand and often enhanced with multimedia elements, such as images, audio descriptions, or augmented reality overlays.

The dataset must be updated regularly to include new attractions or changes in existing attractions. This process involves continuous data collection, annotation and validation to ensure the high quality and relevance of the data, which can be both time-consuming and expensive. Maintaining an updated database that provides detailed and accurate information about the recognized points of interest is crucial. This database must store information and efficiently handle dynamic updates. Challenges include ensuring data consistency, managing data redundancy and securing the database against potential breaches while providing fast access to the information for real-time application needs. When the system encounters an image it cannot initially recognize, this provides an opportunity not just for technical improvement but also for enriching the tourist's experience. Once an image is added to the system, tourist organizations can notify visitors who initially encounter an unrecognized point. This process also opens avenues for tourist organizations to offer additional products or services. By linking notifications about the recognized image to offers for commercial products, souvenirs, or promotions of other tourist spots, organizations can capitalize on increased engagement to boost local commerce. Such targeted suggestions not only improve the tourist experience but also support local businesses, creating a symbiotic relationship that benefits all stakeholders involved.

Multilanguage Support

We propose integrating multilanguage support to enhance the utility and accessibility of the application further. This feature can cater to international tourists by providing information in their preferred language, thereby broadening the reach and usability of the application. To achieve this, we propose implementing multilanguage support using advanced AI translation tools, such as Google's real-time text translator. The integration of Google's real-time text translator will allow the application to instantly translate text related to points of interest (PoIs) into a variety of languages. This feature first detects the user's language preference, either through explicit user selection or by automatically detecting the language settings of the user's device. Once the preferred language is identified, all text outputs, including descriptions of PoIs, historical facts and navigational instructions, are translated in real-time. The use of an AI-powered translation tool ensures that translations are not only swift but also contextually accurate, enhancing the user's understanding and engagement. This approach leverages ongoing advancements in natural language processing (NLP) technologies, which have significantly improved in terms of accuracy and speed owing to continuous updates and training on diverse language datasets. Incorporating multilanguage support in this manner would not only make the application more user-friendly for a global audience but also increase its educational value by making information accessible in multiple languages. Moreover, this feature supports inclusivity and promotes wider adoption of the application among tourists from various linguistic backgrounds, making the touring experience more enjoyable and informative.

DISCUSSION

The YOLOv5m model, employing pre-trained COCO weights, underwent training for 100 epochs. Leveraging these pre-trained weights facilitated faster convergence, enhanced accuracy and robust feature extraction capabilities in our experiments. Our dataset introduces new classes into the model through transfer learning. The benefits of using transfer learning with pre-trained weights include significantly reduced training time



and the need for less training data to achieve high precision, making it a practical strategy for adapting deep learning models to specialized tasks. This medium-sized model configuration utilizes a deeper and wider architecture than the smaller YOLOv5s, incorporates 212 layers and approximately 20.86 million parameters, and operates at a computational complexity of 47.9 GFLOPs. Upon validation, the model demonstrated excellent performance metrics: a precision of 0.948 indicated that 94.8% of the model's predictions were accurate, whereas a recall rate of 0.905 confirmed that the model successfully identified nearly all relevant objects within the dataset. The model achieved a mean Average Precision of 0.957 at 50% Intersection over Union (IoU) and a Mean Average Precision of 0.653 from 50% to 95% IoU, indicating how well the model performs under increasingly strict criteria for matching the predicted bounding boxes with the ground truth. This confirms that the YOLOv5m configuration was selected for its optimal balance between speed and accuracy, making it particularly suitable for scenarios demanding real-time performance without significantly compromising the quality of detections. The trained model showed excellent performance on the validation dataset, with high precision and recall. The high mAP scores indicate that the model is very accurate at detecting the classes on which it has been trained, although the mAP50-95 suggests some potential room for improvement under more stringent IoU thresholds. This could be addressed by further tuning the model, augmenting the training dataset, or adjusting the IoU thresholds for evaluation.

The tourism industry is increasingly embracing the integration of intelligent tourism applications with burgeoning research focusing on the potential of these technologies. Employing visual sensors alongside artificial intelligence, computer vision and machine learning offers reliable and autonomous solutions that can significantly enhance smart tourism in various aspects. Within this context, our use case specifically explores the enhancement of tourist engagement and experience by leveraging real-time information about points of interest. Through the development of this system, Serbia can enhance its tourism infrastructure that directly supports the "Tourism Development Strategy". To maximize the utility and educational value of the application, the system should be capable of recognizing a wide array of objects that enrich the tourist experience. This includes various architectural structures and landmarks ranging from iconic statues and historical buildings to modern government buildings and sports stadiums, enhancing both urban and historical tours. Additionally, it should identify specific artwork, sculptures and historical artefacts, making it an indispensable educational resource within museums by delivering rich contextual information directly to visitors' mobile devices. Moreover, to support eco-tourism, the system should also be equipped to identify natural landmarks, such as mountains, lakes and parks, offering an interactive guide to nature lovers and adventurers exploring natural wonder.

This application offers a multifaceted platform that significantly benefits tourists, tourist organizations and overall engagement. For tourists, it provides a seamless, enriched experience by enabling easy recognition of points of interest, guided tours, interactive museum visits and real-time information on nearby attractions and local businesses. The inclusion of features like personalized routes, voice descriptions and educational tools transforms the app into a personal tour guide, enhancing convenience and learning while promoting a deeper connection with the destination. For tourist organizations, the app serves as a powerful tool for managing visitor flow, promoting lesser-known attractions and gathering valuable analytics on tourist behaviour. This data can be used to optimize city planning, improve marketing strategies and drive sustainable tourism initiatives. Increased engagement is driven by the app's ability to offer interactive and customizable experiences, such as virtual field trips, gamified learning and real-time notifications. The app's potential to integrate with smart city infrastructure further enhances its utility, making it not only a guide for tourists but also a resource for urban development and cultural preservation. Overall, this application fosters a more informed, enjoyable and sustainable tourism experience, benefiting both users and the broader tourism ecosystem.

As we look towards future developments in AI-driven tourism applications, several promising research avenues have emerged that could further enhance the efficacy and impact of these technologies. Future research into dynamic guidance systems that recommend new points of interest based on the locations a tourist has already visited could significantly enhance personalized travel experiences. Such research could explore various data sources and modelling techniques to refine the recommendation accuracy, such as integrating real-time data from social media and other platforms to capture current events or special exhibits. By providing dynamic guidance, the application aligns with the strategy's goals of promoting lesser-known sites and distributing tourist traffic more evenly, thus aiding in reducing seasonality and boosting local economies. There



is also the potential for future research to examine the economic impact of such applications on local businesses and communities. Understanding how AI-driven tourist applications influence spending patterns and visitor footfall can provide valuable insights for local economies. This study could also explore how targeted recommendations and promotions within the application could drive traffic to lesser-known attractions and businesses, thereby promoting a more sustainable and inclusive tourism industry.

This comprehensive support for the strategy's objectives demonstrates how technology can be leveraged to meet broader economic, cultural and educational goals, making its application a vital tool in the evolution and growth of Serbia's tourism sector. The UNDP has recognized sustainable tourism as a significant opportunity for development in Serbia, emphasizing the need to integrate sustainability into the tourism sector to benefit the local economy and align with global sustainability goals. These efforts include enhancing educational programs, improving digital and business skills among local entrepreneurs and promoting tourism outside large urban centers. Our application supports these initiatives, thereby aligning with UNDP's sustainable tourism strategies.

CONCLUSION

Tourists often face significant challenges owing to a lack of readily available information about their destination spots. This gap in accessible, timely and detailed information can hinder their ability to fully engage with and appreciate the attractions they visit, potentially leading to a less-satisfying overall experience. The absence of an insightful context regarding cultural, historical, or natural spots can leave visitors with unanswered questions and unfulfilled expectations, particularly in areas where signage is minimal or only available in a local language. In this study, we explored a real-time object detection application aimed at enhancing the tourist experience by providing instant, accurate information about points of interest. After evaluating various AI models, we selected YOLOv5m (Jocher, 2020) because of its suitability based on the balance of speed, accuracy and ease of integration, providing a substantial increase in detection capability over YOLOv5s, which is designed for ultrafast processing but at the cost of lower accuracy. Compared to larger variants such as YOLOv5l or YOLOv5x, YOLOv5m offers a more balanced approach, ensuring that it remains sufficiently computationally efficient for use in environments with moderate hardware capabilities. This makes YOLOv5m an ideal choice for applications that require a good trade-off between speed and detailed detection, such as in moderate surveillance systems, consumer-grade hardware applications, or when there is a need to detect objects with reasonable accuracy in near real-time. This suitability was further confirmed by our experiments, which demonstrated the effectiveness of the model under these specific conditions.

This study outlines a comprehensive methodology for the application of mobile tourist information systems, employing a robust method based on the AI model for detection and a custom dataset specifically tailored for this purpose. This integrated approach ensures that tourists receive real-time relevant data that enhance their visiting experience, demonstrating the effectiveness of combining advanced AI technology with user-centered design in the development of tourism-oriented applications. We also examined the potential enhancements for the app, such as personalized recommendations that tailor the tourist experience based on individual preferences and past behaviors. Other suggested features include interactive learning tools for museums and historical sites and integration with local tourism and business promotions to support local economies. The challenges and limitations of implementing such a system are also discussed, including computational demands and the need for continuous updates and maintenance to ensure the effectiveness of the application. We detail how tourist organizations could benefit significantly from the deployment of this technology, particularly in regions like Serbia, where comprehensive visitor analytics are not currently available. From an educational perspective, the application empowers tourist organizations to deliver enriched content directly to visitors' mobile devices, facilitating a more interactive and engaging learning experience at cultural, historical and natural sites. Moreover, the application can be utilized to promote local businesses, eco-tourism and general tourist attractions, thereby effectively expanding the economic benefits of tourism. Highlighting lesser-known and eco-friendly options may help diversify tourist activities and distribute tourist traffic more evenly. Although there are challenges to overcome, the potential benefits of this AI-driven real-time object detection application for the tourism industry are substantial, offering a transformative approach to rethinking



strategies for modern tourism. By enhancing visitor experiences through personalized, real-time information delivery and facilitating data-driven insights into tourist behavior, this technology has the potential to reshape how destinations are marketed, managed and explored, driving innovation and sustainable growth within the sector.

REFERENCES

- Alsharabi, N. (2023). Real-Time Object Detection Overview: Advancements, Challenges, and Applications. *Amran University Journal*, 3(6), 12. <https://doi.org/10.59145/jaust.v3i6.73>
- Any, B., Four, S., & Tariazela, C. (2024). Technology Integration in Tourism Management: Enhancing the Visitor Experience. *Startupreneur Business Digital (SABDA Journal)*, 3(1), 81–88. <https://doi.org/10.33050/sabda.v3i1.508>
- Auliarahman, R. M., Baizal, Z. K. A., & Ikhsan, N. (2021). Conversational Recommender Systems Based on Criticism for Tourist Attractions using TF-IDF. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 5(4), 1242.
- Bulchand-Gidumal, J. (2022). Impact of Artificial Intelligence in Travel, Tourism, and Hospitality. In *Handbook of e-Tourism* (pp. 1943–1962). Springer International Publishing. https://doi.org/10.1007/978-3-030-48652-5_110
- Celardo, L., Misuraca, M., & Spano, M. (2023). Georeferencing sentiment scores to map and explore tourist points of interest. *5th International Conference on Advanced Research Methods and Analytics (CARMA 2023)*, 115–122.
- Chary, P. S. (2023). Real Time Object Detection Using YOLOv4. *International Journal for Research in Applied Science and Engineering Technology*, 11(12), 1375–1379. <https://doi.org/10.22214/ijraset.2023.57602>
- Chen, Y.-C., Yu, K.-M., Kao, T.-H., & Hsieh, H.-L. (2021). Deep learning based real-time tourist spots detection and recognition mechanism. *Science Progress*, 104(3_suppl), 003685042110442. <https://doi.org/10.1177/00368504211044228>
- Dai, J. P. (2013). Research of Tourist Attractions Real-Time Query System Based on Mobile GIS. *Advanced Materials Research*, 756–759, 1948–1952. <https://doi.org/10.4028/www.scientific.net/AMR.756-759.1948>
- Dyk, T., Viskupič, M., Dražanský, M., & Macek, I. (2023). Significance of Museum Visitor Behavior Analysis. *Muzeum Muzejní a Vlastivedná Práce*, 62(1), 8–11. <https://doi.org/10.37520/mmvp.2024.003>
- Guan, Z. (2023). Real time object recognition based on YOLO model. *Theoretical and Natural Science*, 28(1), 137–143.
- Gupta, J., Srivastava, S., & Asrani, D. . K. (2024). E-TOURISM: A TOURIST SATISFACTION OF ONLINE TRAVEL PURCHASING. *International Journal of Innovative Research in Computer Science and Technology (IJIRCST)*, 140–144.
- Hosen, B., Rahaman, M., Kumar, S., Sagar, L., & Akhtar, Md. N. (2023). LEVERAGING ARTIFICIAL INTELLIGENCE AND BIG DATA FOR ADVANCED SPATIAL ANALYTICS AND DECISION SUPPORT SYSTEMS IN GEOGRAPHY. *Malaysian Applied Geography*, 1(2), 62–67. <https://doi.org/10.26480/magg.02.2023.62.67>
- Huang, A., Chao, Y., de la Mora Velasco, E., Bilgihan, A., & Wei, W. (2022). When artificial intelligence meets the hospitality and tourism industry: an assessment framework to inform theory and management. *Journal of Hospitality and Tourism Insights*, 5(5), 1080–1100. <https://doi.org/10.1108/JHTI-01-2021-0021>
- Jamaluddin, Z., & Rahmat, A. K. (2023). Artificial Intelligence Technology in Travel, Tourism and Hospitality: Current and Future Developments. In *Technology Application in Aviation, Tourism and Hospitality* (pp. 169–177). Springer Nature Singapore. https://doi.org/10.1007/978-981-19-6619-4_12
- Jocher, G. (2020). *Ultralytics YOLOv5*. <https://doi.org/10.5281/zenodo.3908559>
- Kumar, R. R., Said Al Shekaili, R. S., Al-Sulaimi, B. D. S., & Sulaiman Al-Alawi, R. K. (2021). An Evaluation of Tourism Attributes related to Satisfaction and Challenges by Foreign Tourists in Sultanate of Oman. *International Business Research*, 14(8), 55. <https://doi.org/10.5539/ibr.v14n8p55>
- Lakhotiya, R., Chavan, M., Divate, S., & Pande, S. (2023). Image Detection and Real Time Object Detection. *International Journal for Research in Applied Science and Engineering Technology*, 11(5), 2785–2790.
- Lavanya, G., & Pande, S. D. (2023). Enhancing Real-time Object Detection with YOLO Algorithm. *EAI Endorsed Transactions on Internet of Things*, 10. <https://doi.org/10.4108/eetiot.4541>
- Lee, J., & Hwang, K. (2022). YOLO with adaptive frame control for real-time object detection applications. *Multimedia Tools and Applications*, 81(25), 36375–36396. <https://doi.org/10.1007/s11042-021-11480-0>



- Lee, J., Kang, Y., Kim, N., Kim, D., & Park, Y. (2019). Analysis of Travel Patterns of Seoul Tourists by Trajectory Data Mining. *Abstracts of the ICA*, 1, 1–2. <https://doi.org/10.5194/ica-abs-1-209-2019>
- Pawlak, K. (2023). City tourism application functionality from the perspective of tourist experience and technology acceptance models. Scientific Papers of Silesian University of Technology. *Organization and Management Series*, 2023(174), 131–152.
- Souto, L., Castro, A., Gonçalves, L., & Nascimento, T. (2017). Stairs and Doors Recognition as Natural Landmarks Based on Clouds of 3D Edge-Points from RGB-D Sensors for Mobile Robot Localization. *Sensors*, 17(8), 1824.
- Verhun, A. (2022). Digital Tools for the Development of the Hospitality and Tourism Industry in the Context of a Digitized Economy. *Economic Affairs*, 67(4s). <https://doi.org/10.46852/0424-2513.4s.2022.20>
- Xiong, S., & Zhang, T. (2024). Enhancing tourist loyalty through location-based service apps: Exploring the roles of digital literacy, perceived ease of use, perceived autonomy, virtual-content congruency, and tourist engagement. *PLOS ONE*, 19(1), e0294244. <https://doi.org/10.1371/journal.pone.0294244>
- Zhang, Q., Liu, Y., Liu, L., Lu, S., Feng, Y., & Yu, X. (2021). Location Identification and Personalized Recommendation of Tourist Attractions Based on Image Processing. *Traitement Du Signal*, 38(1), 197–205. <https://doi.org/10.18280/ts.380121>