Journal of the ACS Advances in Computer Science
Volume 15, Issue 1, 2024

Smart Tour Guide Service System Using Deep Learning and Voice Chatbot

Marwa M. Abd El Hamid
The Higher Institute of Computers and Information Technology, Computer Science Department,
El Shorouk Academy, Cairo, Egypt
marwa.mustafa@sha.edu.eg

ABSTRACT

One of the challenges facing the tourism sector is the lack of accuracy in information provided by ordinary tour guides or their unfamiliarity with these places and their history. This can result in poor experiences, high costs, limited options, and a lack of awareness about relevant attractions for tourists. Mobile phones are now widely available, and almost everyone owns one. An application for mobile phones has been developed using the Flutter framework. Tourist attractions related to the type of tourism that the tourist is visiting are displayed using GPS and Maps, integrated into the application. The exceptional progress in Artificial Intelligence (AI), like deep learning, plays a significant role. It is powerful in image recognition, algorithms such as convolutional neural networks (CNNs). The Horus dataset, consisting of 1025 images, has been collected, processed, and categorized into ten classes, each representing a specific statue. CNNs has been implemented on the data, achieving an accuracy of 99.88%. The evaluation of the model's performance using precision, recall, and F1 score is all 99.5%. The system uniquely integrates several advanced technologies to enhance the tourism experience, including a QR code scanner for locations where photography is prohibited, and an AI-powered voice chatbot for providing accurate and reliable information. The system uniquely combines convolutional neural networks (CNNs) for precise image recognition with real-time, interactive assistance through voice chat, addressing challenges such as accessibility, accuracy, and cultural sensitivities. This innovative approach offers personalized and up-to-date information, significantly improving the overall tourism experience. The application combines user-friendly features and adaptable AI technologies, making it accessible to non-technical users and scalable for diverse regions and tourism types worldwide. The goal is to provide an enjoyable, easy, and knowledge-rich experience in line with Egypt's Vision 2030, with Egypt's focus on technological improvement in the tourism sector.

Keywords— Artificial Intelligence (AI), Deep Learning, convolutional neural networks (CNNs), Voice Chatbot

I. INTRODUCTION

Tourism is one of the most important sources of national income for Egypt. It is essential to give this sector the attention and seriousness it deserves. We showcase our Egyptian antiquities and history in the best possible way, providing accurate and reliable information about our landmarks and history in the most effective manner. Our goal should be to make their visit rich in knowledge and enjoyment. In recent years, artificial intelligence has become widespread and widely applied in various fields [1]. The development in this field has become unprecedented, and deep learning plays a major role in it [2]. Convolutional

10.21608/asc.2025.318690.1023

neural network (CNN) technique is powerful in image analysis and recognition algorithms [3].

It is known that people rely on mobile phone technology because of its ease of use, reliability, and availability wherever they go. Additionally, it has become easier to create mobile applications simply and efficiently using Flutter Framework [4]. To facilitate easy integration of the model into the application and to ensure fast communication between the application and the model, the model was hosted on a provider called Render [5]. This also helps keep the application size small. It is also clear that AI voice chatbot has played an effective role in recent times, providing accurate and reliable information, and many people use it in various fields [6].

Due to its accessibility from any device and its ability to provide real-time updates, a web application was built to serve as a guide for the app and the project, and to be used in its marketing. However, there is no denying that the human element can sometimes make mistakes that lead to major problems. For example, in the tourism sector, a human tour guide can make mistakes such as providing incorrect information, lacking information, having a limited tourism program, or not being available in certain places, all of which negatively affect the tourist experience and lack of awareness among tourists of all the places related to the type of tourism they are visiting for.

AI-integrated tourism applications have numerous pros, including advanced features like CNN-based image recognition, multilingual AI chatbots for real-time assistance, and personalized recommendations, making them highly scalable and adaptable to various regions and tourism types. However, they also have notable cons, such as dependence on internet connectivity, and potential challenges in addressing specialized queries or ensuring seamless user experiences for non-technical audiences without careful design.

This system stands out in the tourism domain by seamlessly integrating advanced technologies like convolutional neural networks (CNNs) for accurate image recognition, a QR code scanner for restricted areas where photography is prohibited, and an AI-powered voice chatbot for interactive, hands-free assistance. Unlike many existing solutions, it provides a comprehensive, personalized experience by offering real-time, context-aware information about historical and cultural sites. These features address key challenges such as accessibility, accuracy, and cultural sensitivities, while supporting broader goals to enhance the tourism sector through innovative, technology-driven solutions.

The main aim of this paper is to display all the tourist attractions related to the type of tourism that the tourist is visiting using GPS and Maps and deep learning techniques and algorithms, especially convolutional neural networks (CNNs) to recognize images of statues or tourist attractions. After the recognition process, AI-based chatbots are applied to communicate and provide accurate and reliable information. All of this is integrated into a mobile application to leverage the technologies and capabilities of mobile phones, their ease of use, the best presentation of information, and their availability anytime and anywhere, a web application was built to serve as a guide for the app and the project, and to be used in its marketing.

II. RELATED WORKS

In [7], a mobile application was applied for image detection in the City of Makkah for Pilgrims from all over the world, using 950 images with a size of 240x240 pixels. CNN model achieved an accuracy of 84%. In [8], the authors created a mobile application that uses multiple deep-learning models, trained on specific datasets, to capture and classify a photo taken by a tourist. The authors used a dataset of 846 images and different algorithms. The accuracy rates of various deep learning algorithms, including (SVM, KNN, LSTM, RNN, and CNN) were compared. CNN attained the highest accuracy at 97%.

In [9], the authors applied a mobile application that takes a photo of a tourist in Pakistan, classifies it, and then provides information about the tourist attraction. The authors used a massive dataset of 8000 images and applied CNN to achieve an accuracy of 77 %. In [10], a tourism image classification application was applied based on a convolutional neural network using SqueezeNet. The authors aimed to achieve the best accuracy using a dataset of 3,740 images, each with a size of 227x227. That research achieved an accuracy of 85.75% using CNN. In [11], the authors applied a deep learning-based real-time tourist spot detection and recognition mechanism using 15,000 images for training the data and achieved an accuracy of 70%.

However, the proposed work applied the Horus model using CNN-MobileNet V2. The main aim of this proposed work is to enhance accuracy using a dataset of 1,025 images, each with a size of 224x224. This paper achieved an accuracy of 99.88% using CNN in image prediction by applying fine-tuning on CNN-MobileNet V2. Table 1 shows the reached results of the proposed work compared to other research papers [7-11]. The achieved results revealed that the proposed model achieved promising results with high accuracy based on Horus-Dataset.

Reference index	Model	Dataset name	Accuracy
[7]	CNN	Makkah	84%
[8]	LSTM ,CNN	Bangladesh	87.35 ,97.40 %
[9]	CNN	Pak JEY	78.4%
[10]	Sqeeneezenet	Slender west lake	85.75%
[11]	R-CNN	Region	70%
Proposed Work	MobileNet V2	Horus-Dataset	99.88%

Table (1) Comparison of proposed work with other studies

III. MATERIALS AND METHODS

This paper aims to facilitate tourists' visits to be rich in knowledge and enjoyment and detect ancient Egyptian kings' statues. Therefore, this paper used deep learning

10.21608/asc.2025.318690.1023

methodology harnessing the power of the CNN-MobileNetV2 model [12]. CNN-MobileNet's efficacy across a spectrum of computer vision tasks renders it an ideal candidate for the main goal of this paper. Figure 1 shows the main steps of the workflow diagram of this paper. The application is highly accessible for non-technical users, with intuitive features like GPS, maps, QR code scanners, and an AI voice chatbot, ensuring ease of use across diverse user groups. Its modular design and adaptable technologies, such as CNNs and AI chatbots, offer significant scalability potential for application in other regions and various types of tourism, making it a versatile tool for global tourism enhancement.

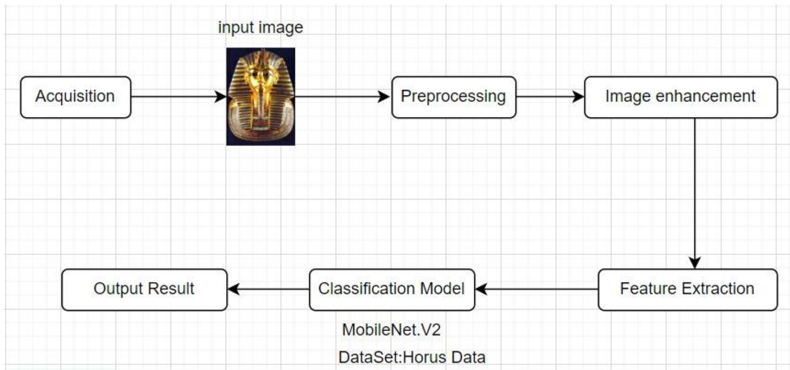

Figure (1) Work flow diagram

Data collection and preprocessing

1- Data Collection [14]:

Figure 2 shows the used dataset in this paper which comprises a collection of images describing ancient Egyptian kings' statues called Horus-Dataset [15]. The used images have been organized into ten folders. Each folder contains pictures of a specific Egyptian king's statue. The number of images in each folder is about 100 images. The used image format is JPG. Figure 3 presents samples of Horus dataset. Horus-Dataset consists of 1025 images of ancient Egyptian kings' statues, collected from diverse sources such as museums and online archives. Its limitations include a small image set per statue and a focus solely on kings' statues, but it is crucial for training the CNN to recognize and categorize these statues for enhancing tourist experiences.

Name	Category	Located in	Images
Tutankhamun	Mask	Cairo	100
Khafre	Statue	Cairo	101
Hatshepsut	Statue	Luxor	103
Tuhotmus III	Statue	Cairo	101
Nevrtity	Statue	Berlin	103
Ramesses II	Statue	Cairo	104
Sphinx	Statue	Giza	100
Horus	Statue	Aswan	103
Djoser	Statue	Cairo	108
Akhenaten	Statue	Cairo	102
10			1025

Figure (2) Horus Dataset

Figure (3) Horus dataset samples

2- **Data Preprocessing:**

To ensure consistency and optimize model performance, we applied essential data preprocessing steps [16]. This involves resizing the collected images to a fixed resolution of 224 x 224 pixels and normalizing the pixel values. Data augmentation techniques such as rotation, scaling, and flipping are also applied to enhance dataset diversity, facilitating robust model training. Resizing images is a fundamental preprocessing step in image classification. Different models require images of a specific size and resizing helps in standardizing the input dimensions. Therefore, all images were resized to a uniform dimension of (224, 224) pixels. By resizing all images to the same dimensions, consistency in the input data was ensured which is crucial for CNN commonly used in image classification.

3- **Train-Test-Validation Split** [17]**:**

Splitting data into training, validation, and test sets is essential for building robust deep-learning models. This process ensures that the model's performance is evaluated fairly and prevents overfitting. In this work, a training Set of 80% was used to train the model and allow it to learn patterns and features. A validation Set of 10% was used to tune hyperparameters and evaluate the model during training without exposing it to the test

10.21608/asc.2025.318690.1023

dataset. And 10% for testing the model used to evaluate the final performance of the model, remaining unseen during training and validation for unbiased evaluation.

4- **Image Data Generator** [18]**:**

Converting Images to arrays of pixels during working with images in deep learning is essential to convert images into arrays of pixels. Because mathematical models cannot directly process raw images. Digital images consist of a grid of pixels in which each pixel is represented by a color value. This conversion allows the model to understand and process the images effectively.

Colored images contain three channels red, green, and blue and each pixel is represented by three values indicating the intensity of each color. Normalizing pixel values to a specific range such as [0, 1] or [-1, 1] is an essential step to help in improving the model's performance and accelerates the training process.

Model Development

The proposed method for detecting ancient Egyptian kings' statues is described emphasizing the key components and the rationale behind it, delving into data collection. Then, preprocessing, model training, evaluation, and deployment phases were applied. Finally, accurate and reliable results were achieved with high accuracy. CNN-MobileNetV2 is a pre-trained convolutional neural network architecture that has been extensively trained on large-scale image datasets.

Figure 4 shows the architecture of the proposed model. CNN-MobileNetV2's pre-trained features offer a significant advantage. Transfer learning involves taking this pre-trained model [13] and adapting it to a new task like classifying different types of statues by fine-tuning its parameters on a smaller, specific dataset. This approach leverages the pre-learned features of CNN-MobileNetV2, optimizing efficiency and accuracy in the new classification task.

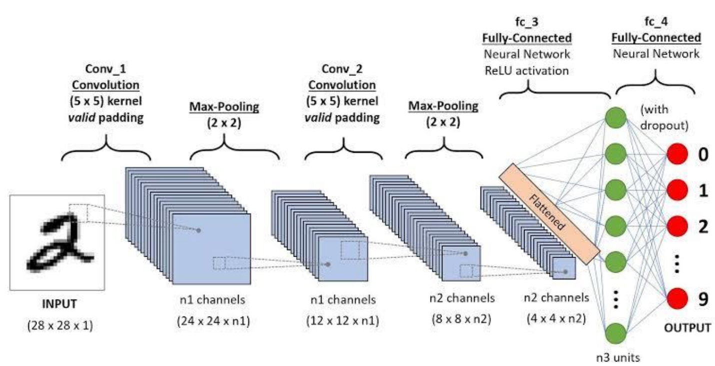

Figure (4) Model Architecture [3]

IV. EXPERIMENTAL SETUP

Table 2 shows the environmental specifications that were used to apply the proposed model

OS	Windows 10
Memory	16.0 GB
Processor	Intel(R) Xeon(R)
GPU	NVIDIA TeslaV100 card 16 GB
Development tool	Python 3.10.13
Used libraries	TensorFlow and Keras
Data set Location	Cloud

Table (2) Development Environment

- **Python Language**:
  Python provides many useful libraries for data processing and model building.
- **Python Libraries**:
  1- TensorFlow Library [19]: TensorFlow is a deep learning library developed by Google. This library provides many tools for building, training, and evaluating machine learning models.

  2- Keras [19] is a high-level API built on top of TensorFlow. Keras allows to build machine learning models and makes them more user-friendly.
- **Training Process** [20]**:** the dataset is partitioned into training, validation, and test subsets.
  Convolutional Neural Network (CNN) is trained on the partitioned dataset. During training, CNN-MobileNet model learns to extract relevant features from images and categorize them.
- **Batch Size**: A batch size of 16 was used to balance computational efficiency and model convergence.
- **Number of Epochs:** The model was trained over 35 epochs, allowing it to iteratively learn from the dataset.
- **Optimizer**: Adamax [21] optimizer was chosen for its effectiveness in handling large datasets and complex architectures.
- **Learning Rate** [22]**:** A learning rate schedule was implemented, starting at 0.001 and decaying by a factor of 0.5 (50% reduction) every 10 epochs. The choice of these hyperparameters was based on preliminary experimentation to ensure optimal model performance.
- **Dataset Splitting:** The dataset was strategically divided into three subsets: training, validation, and test sets, ensuring a fair representation of data across all subsets.
- **Training Set:** Comprising 80% of the dataset, the training set served as the foundation for training the CNN model.

10.21608/asc.2025.318690.1023

- **Validation Set:** A randomly selected 10% of the dataset was allocated to the validation set to monitor model performance during training and prevent overfitting.

- **Test Set:** The remaining 10% of the dataset formed the test set, which remained unseen during training and was used to assess the model's ability to generalize to unseen data. The dataset splitting was performed with careful consideration to maintain a balanced distribution of classes across all subsets, ensuring a representative evaluation of the model's performance.

Table 3 shows CNN's performance metrics. The model's performance evaluation is further enhanced by additional metrics such as precision, recall, and F1 score, complementing its impressive accuracy of 99.88%. Precision, which measures the proportion of correctly identified instances among all predicted instances, is 99.5%. Similarly, recall, representing the proportion of actual instances correctly identified, also reaches approximately 99.5%. As a result, the F1 score, which balances precision and recall, is also 99.5%.

Metric	Value (%)
Accuracy	99.88
Precision	99.5
Recall	99.5
F1 Score	99.5

Table (3) CNN's Performance Evaluation

V. RESULTS AND DISCUSSIONS

The dataset used in this study is called Horus Dataset. This dataset underwent an 80-20 split ratio for training, validation, and test sets. The distribution of images across these sets is as follows: 820 images for training, 164 images for testing, and 41 images for validation. Horus Data comprises images categorized into 10 classes.

We employed CNN-MobileNetV2 model, implemented using the TensorFlow deep learning framework in Python. Performance evaluation utilized metrics such as accuracy, precision, recall, and F1-score. The proposed CNN-MobileNetV2 model showcased promising performance on the Horus Data dataset and demonstrated notable proficiency across the specified evaluation metrics. The AI voice chatbots, trained with advanced NLP techniques, provide accurate and context-aware responses, ensuring users receive reliable real-time assistance. With multilingual capabilities, they enhance accessibility and user satisfaction, making the application interactive and effective for diverse audiences.

Horus Data dataset was partitioned into training, validation, and test sets following an 80-20 split ratio. The distribution of images across these sets is as follows: 820 images for

Journal of the ACS Advances in Computer Science
Volume 15, Issue 1, 2024

training, 164 images for testing, and 41 images for validation. The dataset comprises images classified into 10 classes.

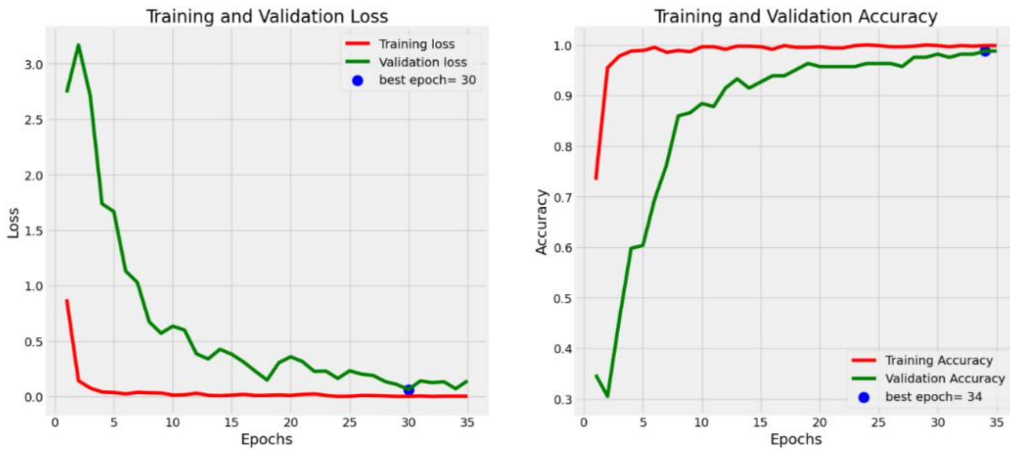

Figure (5) Model's Loss and Accuracy

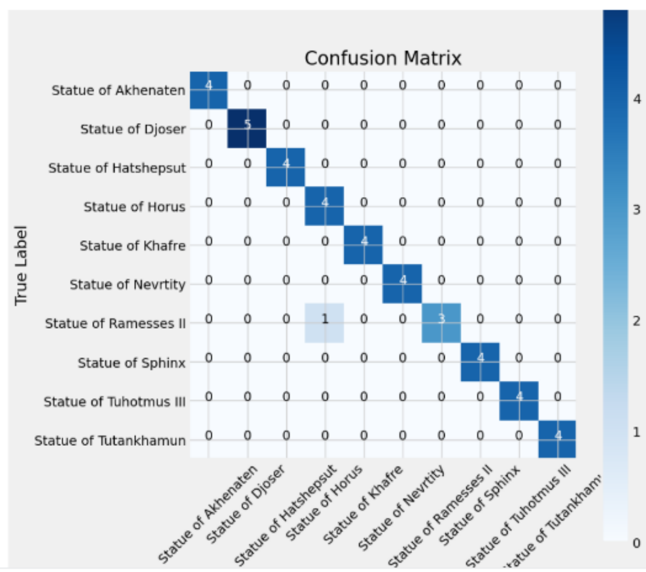

Figure (6) Confusion Matrix

CNN-MobileNetV2 model achieved an accuracy of 99.88%, showing its efficiency and reliability in classifying statues. The evaluation of the model's performance using additional metrics, such as precision, recall, and F1 score is all 99.5%. Figure 5 illustrates the changes in the model's loss and accuracy on both the training and validation sets across epochs. Figure 6 shows the confusion matrix to demonstrate the performance of the classification model by showing the true labels versus the predicted labels. The promising results that were reached help in reliably knowing the name of the statue using this name and passing it to the chatbot.

10.21608/asc.2025.318690.1023

APP DEVELOPMENT

- **App User Interface**

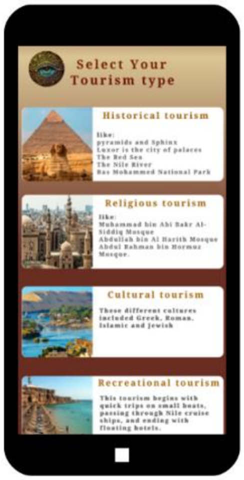

Figure (7) Select Type

Figure (8) Detection Process

Figure (9) Chatbot

- Select the type of tourism as shown in figure 7 to display all associated tourist attractions.
- Use the model to identify the statue and determine its name as shown in figure 8.

  After identifying the statue in figure 8, the statue's name was passed to the API chatbot for further conversation as shown in figure 9. This achieves the goal of providing accurate and reliable information and free from errors.
- Main Challenges
  1. Model Capacity

     Deep learning models typically have a large number of parameters, resulting in large model files. These files may exceed the available storage capacity of many smartphones, making on-device storage impractical.
  2. Computational Resources

     Real-time execution of deep CNN models requires significant processing power, which may not be available on all smartphones.
- Adopted Solutions: Client-Server Architecture
- Server-Side Model Deployment :

  MobileNet-based CNN model has been deployed on a powerful server with sufficient computational resources and memory. This avoids exhausting mobile device storage with large model sizes. Tourists can send images of tourist destinations and relevant data through the mobile application to this server, where they are efficiently analyzed to identify destinations.

- Client-Side Mobile Application :

  A user-friendly mobile application has been developed for tourists on the client side. This app simplifies the process of capturing images of tourist destinations and entering related information while focusing on user interaction and efficient data collection.

- Server-Side Model Execution :

  When a tourist sends an image via the mobile app, it is forwarded to the server for model inference. Leveraging its computational power, the server analyzes the image, ensuring even smartphones with limited capabilities can effectively utilize deep learning models.

- Real-Time Response :

  Enhance the user experience by providing immediate feedback, which is crucial for maintaining user engagement and satisfaction. Additionally, real-time capabilities can be extended to provide personalized recommendations based on user preferences and behaviors.

1. User Interaction: The user captures or selects an image of a tourist destination using the mobile app built with Flutter.
2. Image Upload: The image is sent from the mobile app to the server via RESTful API endpoints.
3. Preprocessing: The server preprocesses the image to ensure it is in the correct format for analysis.
4. Inference: The server runs the deep learning model (MobileNet) to analyze the image and recognize the tourist destination.
5. Response Generation: The server prepares the response which includes detailed information about the recognized destination.
6. Data Display: The mobile app receives the response and displays the information to the user in real-time using Flutter widgets.

- Tourists can capture and upload photos to the application, where they are analyzed using a server-side model as shown in figure 10. After analyzing the photos, the application provides a brief description of each tourist attraction, helping tourists quickly and reliably obtain information about the places they visit, making their journeys more enriching and beneficial.

10.21608/asc.2025.318690.1023

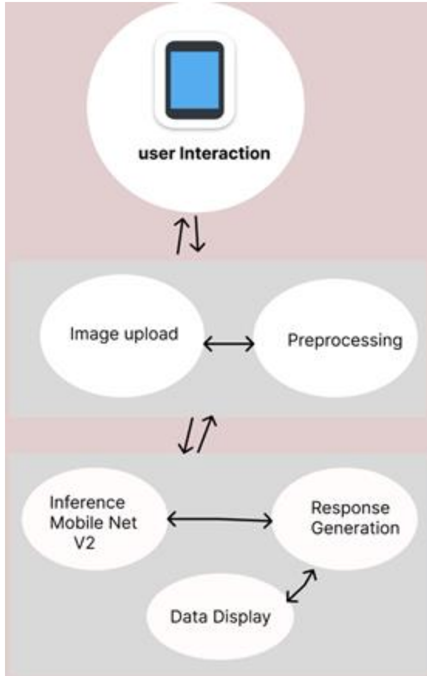

Figure (10) Mobile App Architecture

WEP DEVELOPMENT

The idea behind the website is to serve as a comprehensive description platform for a mobile application. It aims to provide detailed explanations of its features, including computer vision, chatbot capabilities, and text-to-speech functionality. These tools are designed to deliver information quickly and efficiently to users, primarily through voice communication for enhanced accessibility. Additionally, the site utilizes Splide.js to incorporate animations and dynamic movements, creating a user-friendly interface that enhances user engagement as shown in figure 11.

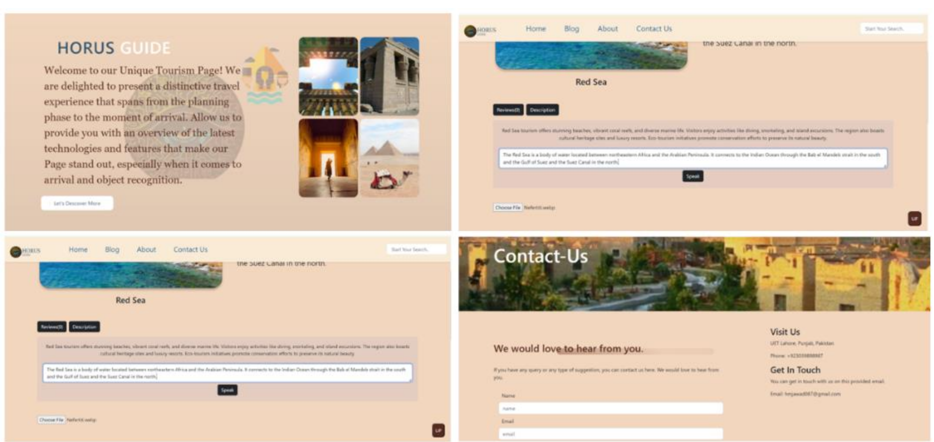

Figure (11) WEP User Interface

- **Key Components of the Website:**
  1- **Application Description**: The description includes a thorough explanation of the mobile application, highlighting features such as computer vision, chatbot interaction, and text-to-speech conversion. It should be clear and informative to illustrate how the application can benefit users.

  2- **Text-to-Speech**: This technology converts text into speech, allowing users to access information through audio instead of reading text. This makes interacting with the application faster and more intuitive.

  3- **Splide.js for Animation**: Splide.js is utilized to add dynamic movement and animations within the website. This may include animated images, smooth slide transitions, and other effects that make the user experience more interactive and enjoyable.

  Development Recommendations:

- **User Experience**: Ensure optimized user experience with an easy-to-use and simplified interface, focusing on fast and clear information access.
- **Responsiveness**: Ensure the design is responsive, adapting well to various screen sizes and devices, including mobile phones and tablets.
- **Security**: Prioritize application security and user data protection, especially when handling sensitive data such as voice and images.
- **Performance Testing**: Regularly test performance to ensure the website operates efficiently and smoothly, particularly with animations and dynamic effects.

The proposed AI-integrated tourism applications offer several advantages, such as CNN-based image recognition, multilingual AI chatbots for real-time support, and personalized recommendations, making them highly adaptable and scalable across different regions and tourism sectors. However, they also come with drawbacks, including reliance on internet connectivity and potential difficulties in handling specialized queries or providing a seamless experience for non-technical users without thoughtful design. The tested results are achieved due to the robust integration of deep learning, real-time data processing, and AI-driven decision-making establishes the system as a highly effective and innovative solution in the tourism sector, reinforcing the reliability and precision of the tested results.

VI. CONCLUSION

In this research paper, the proposed work revealed that Convolutional Neural Network (CNN) technique using MobileNet V2, achieved an accuracy of 99.88%. The evaluation of the model's performance using additional metrics, such as precision, recall, and F1 score, complements its remarkable accuracy of 99.88%. The precision, which measures the proportion of correctly identified instances among the total predicted instances, is 99.5%. Recall, indicating the proportion of actual instances correctly identified, also achieves values around 99.5%. Consequently, the F1 score, a harmonic mean of precision and recall is 99.5%. This result shows high accuracy and reliability in statue recognition. Mobile

10.21608/asc.2025.318690.1023

technologies were used for their availability, reliability, and ease of use by integrating the MobileNet V2 model into a mobile application using the Flutter framework. To facilitate easy integration of the model into the application and to ensure fast communication between the application and the model, the model was hosted on a provider called Render. Additionally, a QR code was included in the application for places where photography is not allowed. An AI voice chatbot was also integrated into the application, providing accurate and reliable information about the statue and its history, and answering any questions the tourist may have. Moreover, tourists can choose the type of tourism they are interested in and see all related places. A web application was provided to serve as a guide for the app and the project and to be used in its marketing. This ensures the tourist has an enjoyable, informative journey, free from any errors or lack of knowledge. The proposed method combines deep learning, AI-driven chatbots, and mobile technology to provide an accurate, scalable, and user-friendly tourism solution. Using CNNs for image recognition, GPS for location-based recommendations, and multilingual AI chatbots, it overcomes challenges like inaccurate information and limited accessibility. A QR code scanner further enhances usability in areas where photography is restricted, ensuring a seamless and culturally sensitive experience.

REFERENCES

[1] J. H. Fetzer and J. H. Fetzer, *What is artificial intelligence?* Springer, 1990.

[2] A. Saxe, S. Nelli, and C. Summerfield, 'If deep learning is the answer, what is the question?', *Nat Rev Neurosci*, vol. 22, no. 1, pp. 55–67, 2021.

[3] P. Picton and P. Picton, *What is a neural network?* Springer, 1994.

[4] L. Dagne, 'Flutter for cross-platform App and SDK development', 2019.

[5] P. Borra, 'An Overview of Cloud Computing and Leading Cloud Service Providers', *INTERNATIONAL JOURNAL OF COMPUTER ENGINEERING & TECHNOLOGY*, vol. 15, pp. 122–133, 2024, doi: 10.17605/OSF.IO/5HQ4M.

[6] Y. Sharma, V. Sharan, and J. Arora, 'The Study Of Voice Bots And Its Impact On Human Welfare', in *Proceedings of the International Conference on Advances in Management Practices (ICAMP 2021)*, 2021.

[7] A. Al Luhaybi *et al.*, 'Smart Tour Guide in Makkah City based on Image Recognition', in *ACI@ ISIC*, 2022, pp. 353–360.

[8] P. Roy, J. Setu, A. Binti, F. Koly, and N. Jahan, 'Tourist Spot Recognition Using Machine Learning Algorithms', 2023, pp. 99–110. doi: 10.1007/978-981-19-1844-5_9.

[9] R. Enam, M. Tahir, S. M. N. Mustafa, R. Qureshi, and H. Shahid, 'Tourist Guide via Image Processing Techniques', *Pakistan Journal of Engineering and Technology*, vol. 4, p. 2021, 2021, doi: 10.51846/vol4iss2pp144-148.

[10] L. Xu, X. Chen, and X. Yang, 'Tourism image classification based on convolutional neural network SqueezeNet——Taking Slender West Lake as an example', *PLoS One*, vol. 19, no. 1, p. e0295439, 2024.

[11] Y.-C. Chen, K.-M. Yu, T.-H. Kao, and H.-L. Hsieh, ‘Deep learning based real-time tourist spots detection and recognition mechanism’, *Sci Prog*, vol. 104, no. 3_suppl, p. 00368504211044228, 2021.

[12] K. Dong, C. Zhou, Y. Ruan, and Y. Li, ‘MobileNetV2 model for image classification’, in *2020 2nd International Conference on Information Technology and Computer Application (ITCA)*, IEEE, 2020, pp. 476–480.

[13] X. Han *et al.*, ‘Pre-trained models: Past, present and future’, *AI Open*, vol. 2, pp. 225–250, 2021.

[14] H. Taherdoost, ‘Data collection methods and tools for research; a step-by-step guide to choose data collection technique for academic and business research projects’, *International Journal of Academic Research in Management (IJARM)*, vol. 10, no. 1, pp. 10–38, 2021.

[15] https://www.kaggle.com/datasets/mohamedwalid008/horus-data-v3 .

[16] J. Chaki and N. Dey, *A beginner's guide to image preprocessing techniques*. CRC Press, 2018.

[17] S. Oymak, M. Li, and M. Soltanolkotabi, ‘Generalization guarantees for neural architecture search with train-validation split’, in *International Conference on Machine Learning*, PMLR, 2021, pp. 8291–8301.

[18] J. G. Pauloski, Z. Zhang, E. Gates, D. Schellingerhout, J. D. Hazle, and D. Fuentes, ‘Optimizing Deep Learning Methods for Image Segmentation with Distributed Training’.

[19] F. J. John Joseph, S. Nonsiri, and A. Monsakul, ‘Keras and TensorFlow: A Hands-On Experience’, 2021, pp. 85–111. doi: 10.1007/978-3-030-66519-7_4.

[20] C. Tennant, M. Boonkrong, and P. A. B. Roberts, ‘The design of a training programme measurement model’, *Journal of European industrial training*, vol. 26, no. 5, pp. 230–240, 2002.

[21] R. Llugsi, S. El Yacoubi, A. Fontaine, and P. Lupera, ‘Comparison between Adam, AdaMax and Adam W optimizers to implement a Weather Forecast based on Neural Networks for the Andean city of Quito’, in *2021 IEEE Fifth Ecuador Technical Chapters Meeting (ETCM)*, IEEE, 2021, pp. 1–6.

[22] E. Even-Dar, Y. Mansour, and P. Bartlett, ‘Learning Rates for Q-learning’, *Journal of machine learning Research*, vol. 5, no. 1, 2003.

10.21608/asc.2025.318690.1023